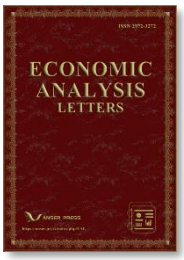
Economic Analysis Letters

Homepage: https://anser.press/index.php/EAL



COVID-19 and non-communicable diseases: GMM/IV Panel VAR evidence from US states

Lotanna E. Emediegwu [a,*]

[a] *Department of Economics, Policy and International Business, Manchester Metropolitan University, Manchester, UK*

ABSTRACT

This paper estimates panel vector autoregressions to analyze the endogenous connection between COVID-19 and non-communicable diseases (NCDs). Using weekly, US state-level data, the study finds evidence of a significant positive effect of NCD-related mortality on deaths due to COVID-19. I find this effect to be higher for males than females. Results are robust to several sensitivity checks, so large deviations are unexpected.

KEYWORDS

COVID-19; Mortality; NCDs; Panel VAR

* Corresponding author: Lotanna E. Emediegwu
E-mail address: l.emediegwu@mmu.ac.uk

ISSN 2972-3272
doi: 10.58567/eal02010006

Received 13 March 2023; Accepted 16 April 2023; Available online 18 April 2023

1. Introduction

Since the discovery of COVID-19 (henceforth covid or coronavirus) in 2019, health services have been disrupted, thus impacting the ability of countries to address and respond to non-communicable diseases (NCDs). For example, new cancer diagnoses have fallen by 25% since the pandemic lockdown began in the Netherlands (Dyer, 2020). NCDs, especially cardiovascular diseases, cancers, diabetes, and chronic respiratory diseases, are the leading causes of death and disability globally, affecting more people annually than all other causes combined (WHO, 2020).

Medical studies (e.g., Jordan et al., 2020) suggest a potential link between covid and NCDs since those with pre-existing NCDs appear more vulnerable to contracting or even dying from the covid virus. Nikoloski et al. (2021), through a systematic review, find that patients with certain chronic illnesses such as diabetes, hypertension (and other cardiovascular diseases), chronic respiratory illnesses, chronic kidney and liver conditions are more likely to be affected by covid. Similarly, Gaur et al. (2021) use state-level Indian covid data to show a positive correlation between deaths per million with NCD risk factors such as obesity, hypertension, diabetes, etc. They also show analogous findings using state-level covid cases per million. Pécout et al. (2021), on their part, use simple statistical analysis to describe the significant impact of the pandemic on NCD patients in Europe and the US. They find that 50% of NCD patients reported worsening conditions during the pandemic, while 17% developed a new disease.

It is important to understand how NCD-related mortality has been impacted by the Covid pandemic in the US since it has the largest share of infection worldwide. Given the lack of detailed analysis in this area, this paper adds to the evidence by undertaking a detailed and robust analysis that explores the bi-directional relationship between covid and NCDs in the US. While few existing studies consider a subset of US cities using raw death counts, this study takes a holistic approach by considering all covid-affected US states. I carry out this activity by estimating panel-data vector autoregressions (PVARs) using weekly datasets covering all covid-affected states in the continental US for 2020-2021, allowing for substantial spatial and temporal variation in the model.

Since the seminal contribution by Holtz-Eakin et al. (1988), PVAR is increasingly becoming a popular econometric tool for estimating multivariate time-series data in a panel setting. This class of models is built on the achievements of traditional panel models (Baltagi, 2001) in terms of allowing the use of large dataset, and VAR models (Sims, 1980), that control for endogeneity by allowing endogenous interactions of the variables within the system. I use Arellano–Bond's dynamic panel equations to show and provide reliable estimates for the dynamic relationship between covid and NCDs. More importantly, this strategy allows the estimation of orthogonalized impulse response functions (IRFs), which helps isolate the response of a variable to an orthogonal shock in another variable of interest.

The results show evidence of a significant positive relationship between mortality due to COVID-19 and NCDs. Most impact peak after two weeks, corresponding with the latent period between covid infection and mortality. These results differ by gender and are robust to several sensitivity checks, so large deviations are unexpected.

The remainder of the paper is subdivided as follows: Section 2 describes the methodology and data, while the various results are discussed in Section 3. The paper ends with some concluding remarks in Section 4.

2. Methodology and Data

I employ a panel VAR approach to investigate the endogenous interactions between deaths due to covid (COV) and NCDs in US states. NCDs are proxied here by respiratory (RES) and circulatory (CIR) diseases. The following model is estimated

$$y_{it} = \alpha_i + A(L)y_{it} + BX_{it} + \varepsilon_{it} \quad (1)$$

where y_{it} is a three-variable vector (COV, RES, CIR) in state *i* at time *t* (in weeks); α_i is a diagonal matrix of state-specific intercepts (fixed effects), which capture time-invariant factors that affect mortality (nation-wide lockdowns, for example). $A(L)$ is a matrix polynomial of lagged coefficients with $A(L) = A_1L_1 + A_2L_2 + \ldots + A_PL_P$, and X_{it} is a vector of exogenous (weather) covariates as these have been cited as important influencers of mortality (e.g., Emediegwu, 2021, Deschênes and Greenstone, 2011). *B* are parameters to be estimated, and ε_{it} is a vector of idiosyncratic errors. The autoregressive order (*p*=2) of the VAR is selected using the Bayesian Information Criterion.

In the spirit of Emediegwu and Nnadozie (2023), I transform death cases into their week-on-week (WoW) log-differenced values to ease the interpretation of the impulse-responses in percentage terms and for policy relevance. I calculate the growth rate of mortality (g_{it}) as

$$g_{it} = \log(y_{it}) - \log(y_{it-1})$$

where y_{it} refers to cumulative deaths from each mortality cause in state *i* at time *t*. In principle, $y_{it} - y_{it-1}$ refers to the number of new deaths in the last week within each state in the US.

Given the dynamic nature of the model, the fixed effects are likely correlated to the lags of the dependent variable, meaning that the common method of eliminating fixed effects (mean-differencing) would produce biased results. To overcome this weakness, I employ the forward mean-differencing or orthogonal deviation (Helmert transformation) approach proposed in Arellano and Bover (1995) as an alternative transformation.[1] This transformation allows the use of lagged covariates as instruments since it retains the orthogonal structure between the lagged covariates and the transformed variables (Baltagi, 2008). Hence, the model coefficients can be jointly estimated using system GMM.

To compute the impulse-response functions (IRFs) and forecast-error variance decompositions (FEVDs), I apply Cholesky decomposition to the residuals to orthogonalize them. In the Cholesky ordering, I allow COV to have a contemporaneous impact on RES and CIR, while the latter two are not allowed to have such impact on the former. This arrangement, by construction, implies that the variable that appears earlier (covid) is weakly exogenous with respect to the rest of the covariates in the short run. However, with no strong theoretical basis for this ordering, I present results with alternative orderings as robustness.

Finally, the IRFs and FEVDs are estimated using the method described in Love and Zicchino (2006), where the confidence intervals are estimated using Monte-Carlo simulations.[2] The IRFs describe the response of a variable over time to shocks to another variable within the system, while the variance decomposition measures the percent of the change in a variable that is explained by the innovation in another variable at a given forecast horizon (6 weeks, in our case).

2.1. Mortality and Population Data

I collate weekly mortality data from the US Centers for Disease Control and Prevention (CDC) WONDER Online Database.[3] The study's sample period begins from January 2020 to December 2021 to allow for the relevant period

[1] According to Love & Zicchino (2006), this "forward mean-differencing" approach removes the mean of all future observations for each state-week instead of using deviations from past observations.

[2] Practically, I re-estimated the IRFs by randomly building a draw of coefficients A of equation (1) using the estimated coefficients and the associated variance-covariance matrix. This process is repeated 1000 times to obtain the 5th and 95th percentiles of the distribution used as confidence intervals of IRFs.

[3] Respiratory diseases are made up of 20 underlying causes of death related to the respiratory system, while circulatory diseases comprise 30 underlying causes of death related to the circulatory system. More information on how the data are compiled can be found in the Technical Appendix domiciled on the CDC's website.

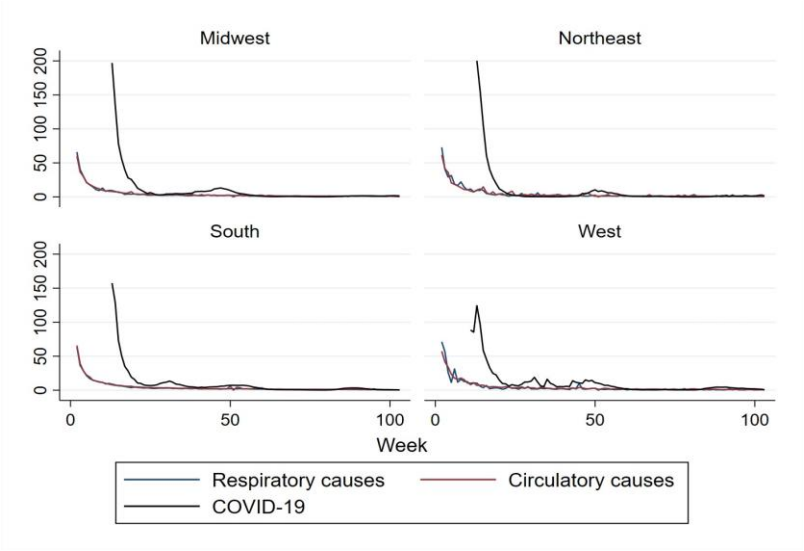


Figure 1. Trends of growth rates of deaths (in %) from COVID-19, respiratory, and circulatory diseases in the US (by Region).

that captures the inception of the covid pandemic but excludes the vaccine era (2022). The sample contains the continental 48 US states, including the District of Columbia, totaling 4,627 state-week observations.[4] Further, state-level projected 2020 population information was obtained from the US Census Bureau's website. I use this state-level population figure to derive the first difference in the natural logarithm of weekly cumulative deaths per 1 million people. Figure 1 suggests some comovement in the growth rates of mortality from the three diseases considered.

2.2. Weather Data

The historical weather dataset is obtained from the ERA5 reanalysis product from European Centre for medium-range weather Forecasts (ECMWF), which provides daily gridded weather variables at 0.25^{o} resolution. Specifically, I collected hourly downward UV radiation at the surface (in J/m^{2}hour), 2-meter temperature (in ^{o}C), 2-meter dewpoint temperature (in ^{o}C), and 10-metre U and V wind components (in m/s) for January 01, 2020, to December 31, 2021. The weather measures were averaged across hours in a week to obtain weekly averages. Thereafter, I link the weather data to state-level mortality cases by overlaying a US polygon with state boundaries on the gridded weather dataset for each grid cell and then taking a simple average across all grid cells per state.

3. Results and Discussion

3.1. Benchmark results

Figure 2 depicts that the estimated GMM panel VAR (Eq. (1)) is stable since the modulus of each eigenvalue of the fitted model is strictly less than one: inside the unit circle. The stability of the estimated model implies that shocks will eventually zero-converge: hence, the PVAR is invertible, making the estimated IRFs and FEVDs interpretable.

In what follows, I present the IRFs graphs and their associated 95% confidence intervals generated *via* Monte Carlo simulations with 1,000 repetitions in Figure 3. The Figure reports the impact of a shock in any variable on

[4] I excluded Alaska and Hawaii, and also dropped observations from unspecified states.

other endogenous variables for 26 weeks (6 months) after the introduction of the shock. The results show that shocks to RES, CIR, and COV are transitory as the effect of a shock fades out almost immediately as seen in the diagonal panels (a, e, and i). However, this paper's main interest lies in the impact of shocks from any of the disease categories on the others – the off-diagonal panels. A positive shock to COV appears to have a positive impact on both RES and CIR as seen in panels g and h in Figure 3. Similarly, RES and CIR respective shocks also impact COV positively, as shown in panels c and f. These results confirm that mortality from covid and NCDs are closely linked. Further, most impacts peak after two weeks (corresponding with the latent period between when one contracts covid and mortality) before they start dying out. Additionally, this study finds that a shock to CIR does not impact RES significantly, whereas a positive shock to RES affects CIR positively, as seen in panels d and b, respectively. These results suggest that the perceived correlation between respiratory and circulatory disease-related deaths is due to the effect of RES on CIR, rather than the other way around.

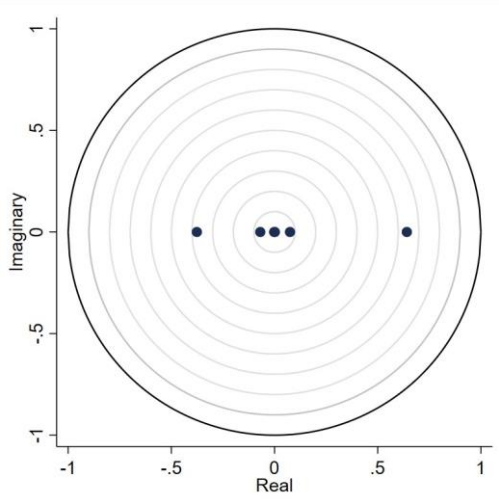


Figure 2. Roots of companion matrix.

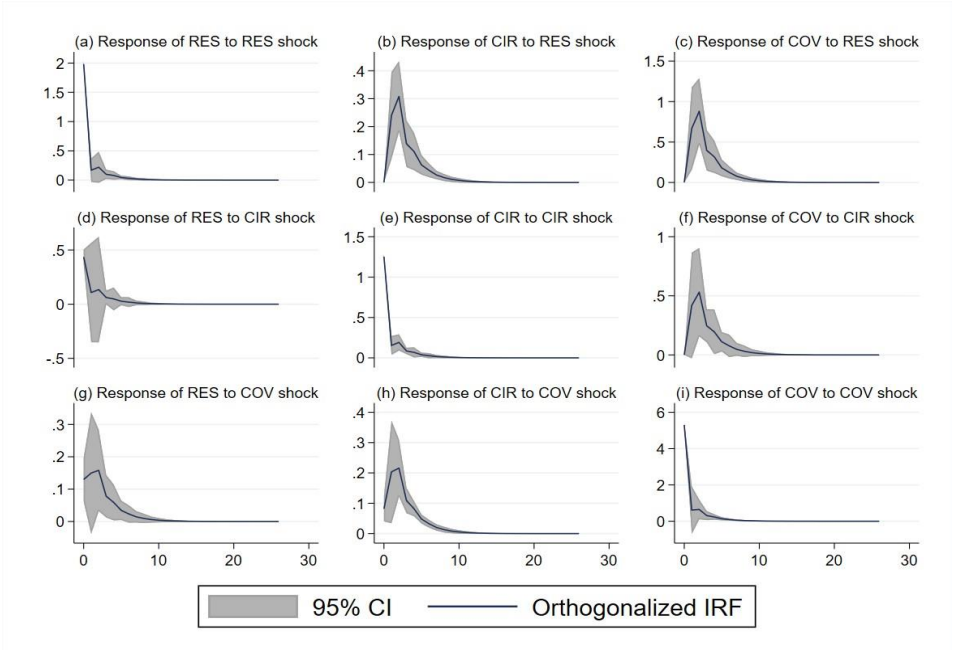


Figure 3. Impulse-response functions.

I present the cumulative IRF in Figure 4 to show the effects in levels rather than in log-differences by accumulating the impacts over time (26 weeks). Although the appearance of Figure 4 differs from Figure 3, they have similar interpretations and findings. As discussed earlier, the impact of a positive shock to COV on RES and CIR is positive. In like manner, a shock from either CIR or RES affects COV positively. Also, CIR shock does not have a significant impact on RES.

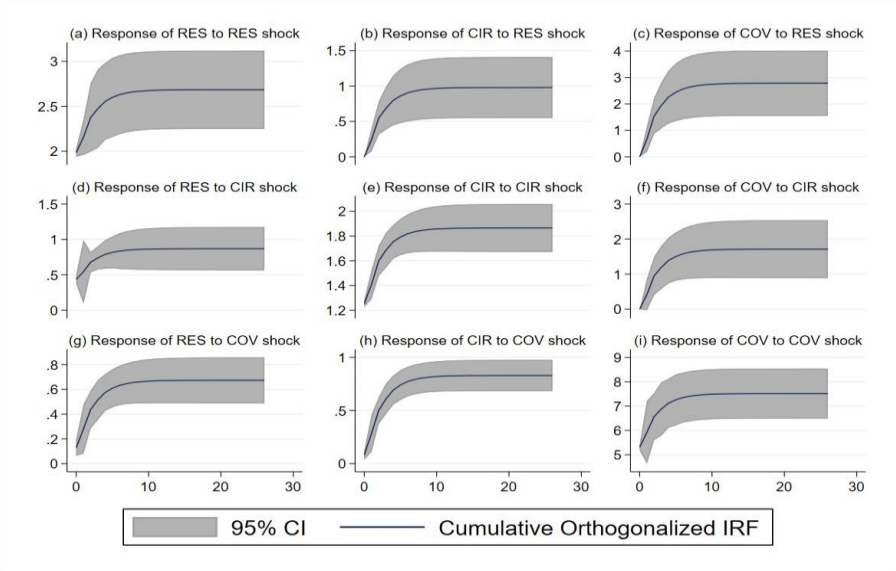


Figure 4. Cumulative IRFs.

Advances in medical sciences have shown that mortality rate differs across sub-population. Since the GMM approach imposes homogeneous dynamics across spatial units, I divide the sample by gender. The gendered IRFs in Figure 5 show heterogeneous dynamics across the sub-sample. While the overall pattern remains qualitatively similar to the main results, the IRFs for both genders are below what is observed in the entire sample for responses to RES shock. On the other hand, the impact of COV shock is higher in males than females, suggesting that men with NCDs are more predisposed to die from covid than women.

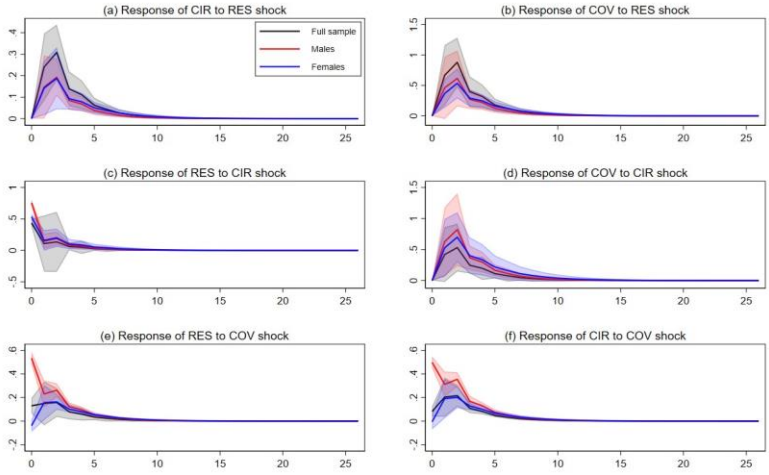


Figure 5. Gendered IRFs. 95% confidence bounds are represented by shaded areas.

To understand the contribution of changes in each variable to changes in other variables, I turn to the variance decomposition results presented in Table 1. The Table shows that RES explains more of the variation in COV and

CIR for 25 periods ahead: nevertheless, the size of the effect is rather small, 4.9% and 9.7%, respectively.

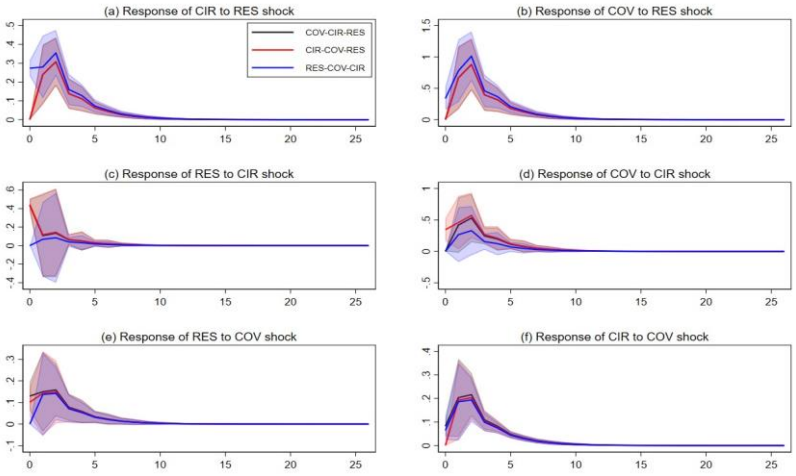
Table 1. Variance decompositions.

	COV	CIR	RES
COV	0.932	0.018	0.049
CIR	0.059	0.842	0.097
RES	0.017	0.052	0.930

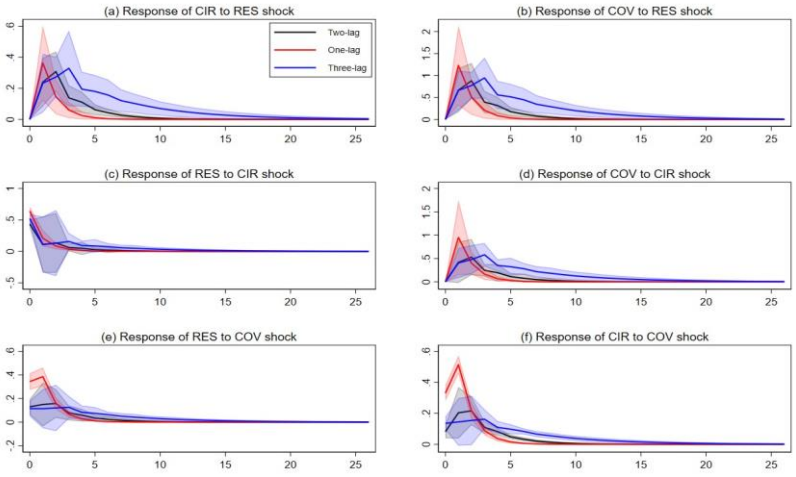
Note: Variation in the row variable explained by column variable (25 periods ahead).

3.2. Robustness test

Here, I use four alternative specifications to test the robustness of the results. First, I re-arrange the recursive order of the endogenous variables – allowing both RES and CIR to assume weak exogeneity in turns. The results in Figure 6a produce a similar pattern as the main result. Also, I show in Figure 6b that the results are robust to lag choices by replacing the selected lag used in the primary model with lags 1 and 3, respectively. For most of the specifications, the IRF from 3-lag model appears higher than others; however, such marginal differences do not alter the overall interpretation of the results.

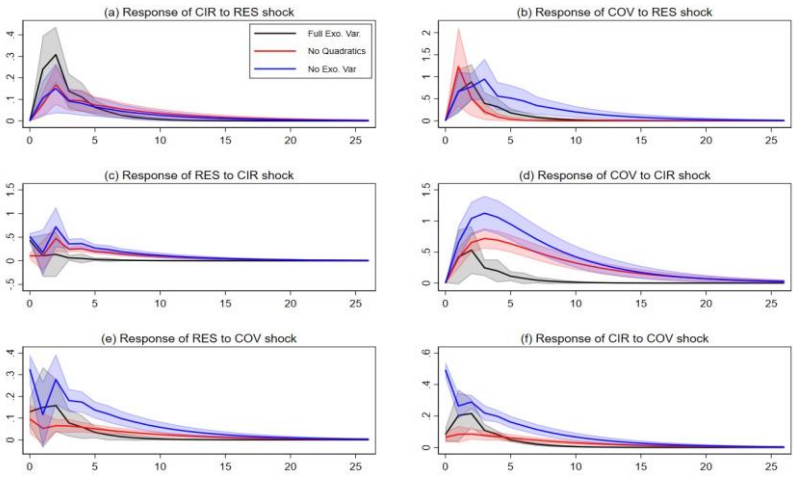


(a) Recursive Ordering

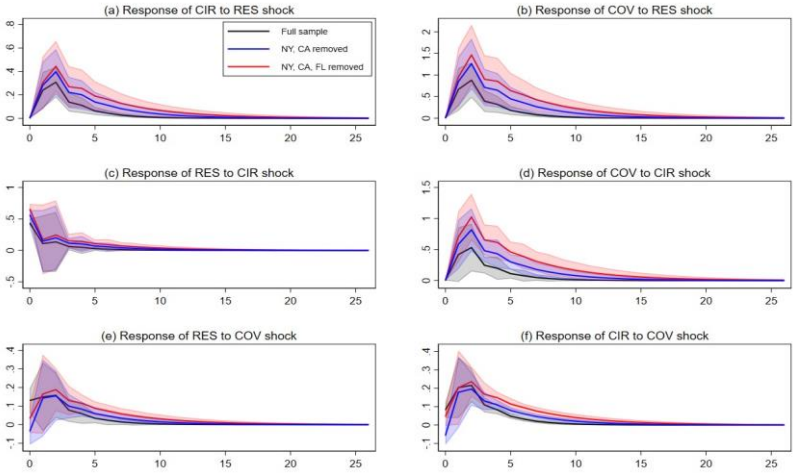


(b) Lag Adjustment

Figure 6. IRFs from Alternative Specifications (Recursive Re-ordering and Lag Adjustment). 95% confidence bounds are represented by shaded areas.



(a) No exogenous variables



(b) Outliers

Figure 7. IRFs from Alternative Specifications (Weather Variables and Outliers). 95% confidence bounds are represented by shaded areas.

Next, the baseline model is re-estimated with the exogenous variables (weather) excluded. Figure 7a shows that omitting weather variables increases the impact of a shock on one variable on another variable significantly. This result highlights the importance of weather variables in explaining mortality. Lastly, to confirm that the results are not driven by outliers, I purged out observation in the three states with the highest cumulative covid cases as of December 31, 2021 – New York (NY), California (CA), and Florida (FL) in that order. Figure 7b shows that the states with large cases do not principally drive the results, as the estimates are still very similar to the baseline results.

4. Conclusion

This paper reports that COVID-19 and NCDs reinforce themselves. After controlling for time-invariant location effects and important (exogenous) weather variables, I find a bi-directional relationship between most

combinations of death causes considered. Specifically, I find that covid mortality in the US positively affects mortality from respiratory and circulatory diseases. On the other hand, the study does not find a statistically significant effect of mortality due to respiratory diseases on circulatory diseases. Further, the paper finds that these impacts are higher in the male sub-population than in the females. Finally, the results are robust to alternative VAR specifications and several sampling alterations, suggesting that large deviations are unexpected.

Using robust dataset and established empirical approach, the findings in this paper add to evidence of the impact of coronavirus on NCD-related deaths. Therefore, urgent health policies should be implemented to ensure that people dying from non-covid causes are not rising due to the enormous attention shifted towards the pandemic. Innovative health measures (e.g., having cancer detection scanners in supermarkets) should be developed to ensure the continuation of conventional health services. Although with the introduction and push towards compulsory vaccination in countries like the US, the coronavirus-NCD nexus is expected to be weakened. However, further research is required to investigate if this assertion holds and to what degree.

Some limitations of the study are as follows. The paper could not disentangle the endogenous interactions between covid and NCDs by age and race due to data unavailability. Besides. the SARS-CoV-2 genome has mutated into more than a thousand strands. These strands may interact differently with NCDs.

Funding Statement

The author benefited from the Manchester Metropolitan University Internal Seed Funding. The usual disclaimer applies.

Conflict of interest

The author claims that the manuscript is completely original. The author also declares no conflict of interest.

Acknowledgement

Much appreciation goes to Thomas Goodrick for providing an excellent research assistantship on this project.

References

Arellano, M., Bover, O. (1995), 'Another look at the instrumental variable estimation of error-components models', *Journal of Econometrics* 68(1), 29–51. https://doi.org/10.1016/0304-4076(94)01642-D

Baltagi, B. H. (2001), *Econometric analysis of panel data*, John Wiley, Sons, London.

Baltagi, B. H. (2008), *Econometric analysis of panel data*, Vol. 4, Springer.

Deschênes, O., Greenstone, M. (2011), 'Climate change, mortality, and adaptation: Evidence from annual fluctuations in weather in the US', *American Economic Journal: Applied Economics* 3(4), 152–85. https://www.aeaweb.org/articles?id=10.1257/app.3.4.152

Dyer, O. (2020), 'Covid-19: Pandemic is having "severe" impact on non-communicable disease care, WHO survey finds'. https://doi.org/10.1136/bmj.m2210

Emediegwu, L. E. (2021), 'Health impacts of daily weather fluctuations: Empirical evidence from COVID-19 in US counties', *Journal of Environmental Management* 291, 112662. https://doi.org/10.1016/j.jenvman.2021.112662

Emediegwu, L. E., Nnadozie, O. O. (2023), 'On the effects of COVID-19 on food prices in India: a time-varying approach', *European Review of Agricultural Economics*, *50*(2), 232-249. https://doi.org/10.1093/erae/jbac015

Gaur, K., Khedar, R. S., Mangal, K., Sharma, A. K., Dhamija, R. K., Gupta, R. (2021), 'Macrolevel association of COVID-19 with non-communicable disease risk factors in India', *Diabetes, Metabolic Syndrome: Clinical Research, Reviews* 15(1), 343–350. https://doi.org/10.1016/j.dsx.2021.01.005

Holtz-Eakin, D., Newey, W., Rosen, H. S. (1988), 'Estimating vector autoregressions with panel data', *Econometrica:*

Journal of the econometric society pp. 1371–1395. https://doi.org/10.2307/1913103

Jordan, R. E., Adab, P., Cheng, K. (2020), 'Covid-19: risk factors for severe disease and death', *British Medical Journal* 368. https://doi.org/10.1136/bmj.m1198

Love, I., Zicchino, L. (2006), 'Financial development and dynamic investment behavior: Evidence from panel var', *The Quarterly Review of Economics and Finance* 46(2), 190– 210. https://doi.org/10.1016/j.qref.2005.11.007

Nikoloski, Z., Alqunaibet, A. M., Alfawaz, R. A., Almudarra, S. S., Herbst, C. H., El-Saharty, S., Alsukait, R., Algwizani, A. (2021), 'Covid-19 and non-communicable diseases: evidence from a systematic literature review', *BMC Public Health* 21(1), 1–9. https://doi.org/10.1186/s12889-021-11116-w

Pécout, C., Pain, E., Chekroun, M., Champeix, C., Kulak, C., Prieto, R., van Vugt, J., Gilchrist, K., Lainé-Pellet, A.-F. (2021), 'Impact of the COVID-19 Pandemic on Patients Affected by Communicable Diseases in Europe and in the USA', *International Journal of Environmental Research and Public Health* 18(13), 6697. https://doi.org/10.3390/ijerph18136697

Sims, C. A. (1980), 'Macroeconomics and reality', *Econometrica: journal of the Econometric Society* pp. 1–48.

WHO (2020), 'Information note on COVID-19 and NCDs'. https://doi.org/10.2307/1912017