
New Zealand Economic Papers

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/rnzp20

# Government mandated lockdowns do not reduce Covid-19 deaths: implications for evaluating the stringent New Zealand response

John Gibson

NEW ZEALAND ECONOMIC PAPERS
https://doi.org/10.1080/00779954.2020.1844786

Routledge
Taylor & Francis Group



# Government mandated lockdowns do not reduce Covid-19 deaths: implications for evaluating the stringent New Zealand response

John Gibson

Department of Economics, University of Waikato, Hamilton, New Zealand

**ABSTRACT**
The New Zealand policy response to Coronavirus was the most stringent in the world during the Level 4 lockdown. Up to 10 billion dollars of output ($\approx$3.3% of GDP) was lost in moving to Level 4 rather than staying at Level 2, according to Treasury calculations. For lockdown to be optimal requires large health benefits to offset this output loss. Forecast deaths from epidemiological models are not valid counterfactuals, due to poor identification. Instead, I use empirical data, based on variation amongst United States counties, over one-fifth of which just had social distancing rather than lockdown. Political drivers of lockdown provide identification. Lockdowns do not reduce Covid-19 deaths. This pattern is visible on each date that key lockdown decisions were made in New Zealand. The apparent ineffectiveness of lockdowns suggests that New Zealand suffered large economic costs for little benefit in terms of lives saved.

**ARTICLE HISTORY**
Received 18 August 2020
Accepted 25 October 2020

**KEYWORDS**
Covid-19; deaths; impact evaluation; lockdown; response stringency

**JEL CODES**
C21; I18

## 1. Introduction

On 23 March 2020 New Zealand's Prime Minister announced a nationwide lockdown for four weeks, to start on 25 March. On 20 April the lockdown was extended until 27 April. The lockdown was Level 4 of the Coronavirus alert system – the 'eliminate' level. The levels had been introduced just two days earlier, first starting at Level 2 – the 'reduce' level – and jumping to Level 3 – the 'restrict' level – during the Prime Ministerial statement. With these decisions, between 25 March and 27 April New Zealand had the most stringent settings in the world for containing Coronavirus.

In the top panel of Figure 1 the timing of these announcements, and key changes in the alert levels, are overlaid on the time-series of the OxCGRT stringency index for New Zealand, which is based on eight indicators of containment and closure (Hale, Webster, Petherick, Phillips, & Kira, 2020). In the bottom panel of Figure 1 the stringency index for New Zealand is compared with that for several other countries. This comparison shows that from 25 March the New Zealand stringency index exceeded that for countries like Italy, Spain and France who by then had thousands of Covid-19 deaths.

The New Zealand Treasury assume that output at Level 4 was reduced by 40%, at Level 3 by 25%, and at Level 2 by 10–15% (Treasury, 2020). So even with a V-shaped shock and recovery rather than a U or L shape, 33 days of Level 4 and 19 of Level 3 (that ended 13 May) would reduce output by 10 billion dollars (ca. 3.3% of GDP) compared to staying in Level 2 throughout. The purpose of the current study is to see what health benefits – in terms of lives saved – were likely achieved to balance against this cost in terms of lost output.

**CONTACT** John Gibson jkgibson@waikato.ac.nz Department of Economics, University of Waikato, Private Bag 3105, Hamilton 3240, New Zealand

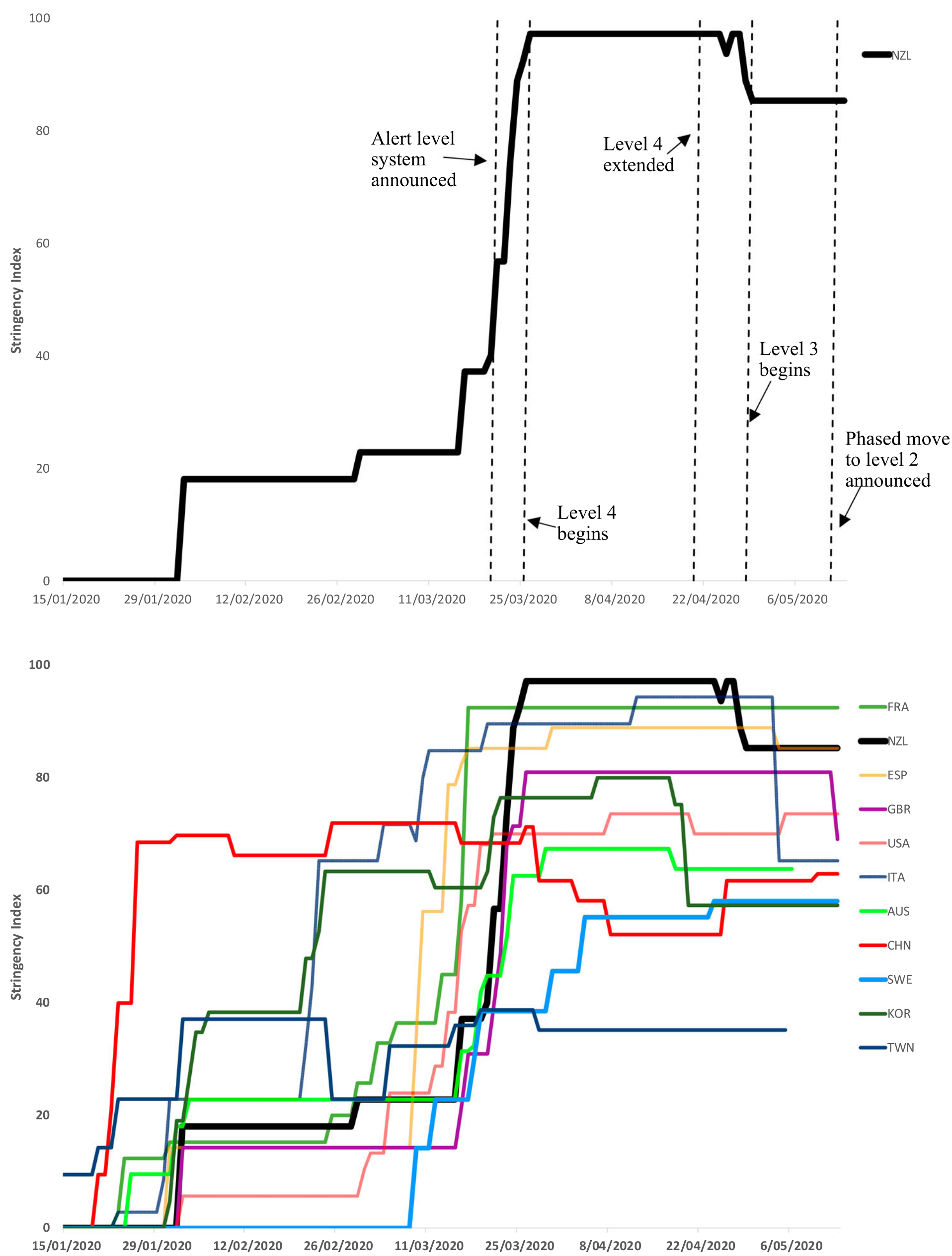


**Figure 1.** New Zealand had the most stringent government response to coronavirus.
Source: Author's calculation from Oxford Coronavirus Government Response Tracker data

One would assume that rigorous cost–benefit analyses accompanied the decision to set the most stringent policy response in the world. Yet Cabinet papers released six weeks later suggest not: the government ignored advice from the Ministry of Health to stay at Level 2 for 30 days, instead jumping to Level 3 after just two days, then Level 4 two days later (Daalder, 2020).[1] Two epidemiological simulations seem to have played a key role; the Imperial College forecast of 0.5 million Covid-19 deaths in the U.K. and 2.2 million in the U.S. if no changes in individual behaviour or in control measures occurred (Ferguson et al, 2020), and forecasts by University of Otago academics with an on-line simulator (http://covidsim.eu) that ranged from seven Covid-19 deaths (assuming low infectiousness, $R_0 = 1.5$ and 50% general contact reduction for nine months) to 14,400 (highly infectious, $R_0 = 3.5$, just 25% contact reduction for six months), with a mean across the six forecasts of 8300 deaths (Wilson, Barnard, Kvalsvig, & Baker, 2020).

Even though the Imperial College forecast was not for New Zealand, it seemed to shift local strategy away from 'flatten the curve' mitigation to one where:

> … you want to have a series of small peaks over a longer period of time and you amplify up quite stringent controls … then as it goes down again, you can ease those and be prepared to ramp them up again. (Director-General of Health, March 19)

This description matches a chart in the Imperial College forecast, for a suppression strategy in place for two years (Daalder, 2020). The highest death forecasts from the Otago academics may have influenced comments made by the Prime Minister in announcing the lockdown:[2]

> If community transmission takes off … our health system will be inundated, and tens of thousands New Zealanders will die … it is the reality we have seen overseas … We can stop the spread by staying at home and reducing contact … That's why … effective immediately we will move to Alert Level 3 … after 48 hours we will move to Level 4.

It is unfortunate that epidemiological simulations had such impact. The Susceptible, Infected, Recovered (SIR) epidemiological model, and variants with Exposed and Dead (SEIRD), have infectious people mixing (homogeneously) with others; each person has equal chances to meet any other, regardless of their health status. Yet in reality, people engage in preventative behaviour to reduce the risk of exposure; allow for this, and some public actions designed to reduce disease spread may do more harm (Toxvaerd, 2019). These models also have too many degrees of freedom, so are poorly identified from short-run data on cases. For example, Korolev (2020) shows long-run forecasts of U.S. COVID-19 deaths from observationally equivalent SEIRD models ranged from about 30,000 to over a million. Forecast deaths depend on arbitrary choices by researchers, and data at the time cannot show which forecast is right as so many models are observationally equivalent in the short-run. Elsewhere, Swedish researchers using the Imperial College approach forecast (in mid-April) 80,000 Covid-19 deaths by mid-May (Gardner *et al.*, 2020). In fact, just 3500 died by 15 May, with the forecast more than 20-times too high. A final example is the Otago forecasts, which had assumed no case tracing and isolation; using the same simulation model, Harrison (2020) set tracing and isolation success at 50% and forecast deaths fell by 96%.

Rather than using poorly identified simulation models, I use data on Covid-19 deaths, as of each date key lockdown decisions were made in New Zealand. Deaths data are more reliable than cases data (Homburg, 2020).[3] My research design exploits variation among U.S. counties, over one-fifth of which just had social distancing rather than lockdown. Political drivers of lockdown provide identification. If the Prime Ministerial claim, that *sans* lockdown tens of thousands of New Zealanders would die, is correct then one would expect to see more deaths in places without a lockdown. This may explain global fascination with Sweden, as a country without lockdown. However a within-country research design has two benefits; less variation in measuring Covid-19 deaths than for between-country comparisons, and it better suits the highly clustered nature of Covid-19. For example, Lombardy's Covid-19 death rate was 1500 per million versus 300 per million elsewhere in Italy. The New York death rate (by 15 May) was 1410 per million but just 190 per million in the other 49 states. Taking China's data at face value, Hubei's death rate was 76 per million versus 0.12 per million elsewhere. With such clustering, analyses using national averages may mislead.[4]

Whether a county had a lockdown has no effect on Covid-19 deaths; a non-effect that persists over time. Cross-country studies also find lockdowns superfluous and ineffective (Homburg, 2020). This ineffectiveness has several causes: real-time activity indicators suggest the threat of Covid-19, rather than lockdown *per se*, drives behaviour (Chetty, Friedman, Hendren, & Stepner, 2020). Just one-tenth of the 60% fall in consumer mobility in the U.S. was from legal restrictions, with the rest from people voluntarily staying home to avoid infection (Goolsbee & Syverson, 2020). Likewise, Cronin and Evans (2020) find that more than three-quarters of the decline in foot traffic was due to private behaviour, with mobility falling before state or local regulations were in place. Economic theory also shows that public health interventions can paradoxically increase infection rates due to risk compensation effects (Dasaratha, 2020; Toxvaerd, 2019). Notably, lockdown is not historically used to deal with epidemics, which is why some epidemiologists (e.g. Giesecke, 2020) remain opposed. A review, prompted by the 2006 U.S. Pandemic Influenza Plan, argued against confining large groups like entire cities:

> There are no historical observations or scientific studies that support the confinement by quarantine of groups of possibly infected people for extended periods in order to slow the spread of influenza … . The negative consequences of large-scale quarantine are so extreme … that this mitigation measure should be eliminated from serious consideration. (Inglesby, Nuzzo, O'Toole, & Henderson, 2006, p. 371)

Instead, isolation of infected individuals was historically relied upon – and eventual use of this in Wuhan, rather than lockdown, was key to breaking the disease spread (Stone, 2020).

## 2. County-level evidence from the United States

The U.S. provides useful variation for estimating impacts of lockdowns because the Tenth Amendment to the Constitution gives police powers to states, which limits the federal response to epidemics (Inglesby *et al.*, 2006). The first-shelter-in-place or stay-at-home orders were issued on 14 March for San Francisco-area counties, followed by a California-wide lockdown from 19 March. Many other governors quickly issued state-wide lockdowns, but in others (e.g. Texas), weaker 'state of disaster' notices let cities and counties adopt local lockdown rules, albeit with federal social distancing guidelines in the background.[5]

The varied situation that resulted is seen in Figure 2, which shows counties subject to lockdown orders (technically, government-ordered community quarantine) and those with just social distancing. Data are from American Red Cross reporting on emergency regulations for each county, from 14 March onward. The map was first posted by ESRI (of ArcGIS fame) on 3 April, updating through early April if rules changed.

With such a dynamic situation, care must be taken in defining the treatment variable. One could use the timing and duration of lockdown orders but with many orders still in place by mid-May durations were incomplete at the time of key decisions in New Zealand and so provide weaker evidence. Instead, I use the binary treatment of being subject to lockdown as of early April; the situation seen in Figure 2. All data sources were available to inform New Zealand policymakers from mid-March (the map data were available from the Red Cross, ESRI later made them more conveniently available). See Appendix 1 for details.

The number of Covid-19 deaths per county is highly skewed, with standard deviations over eight times the mean (as of mid-May). Therefore, the log of the number of deaths is the outcome variable for the regressions, reducing the coefficient of variation (CoV) to 1.3.[6] Death rates could be used (CoV $= 2.5$), but are less flexible than log deaths with log population as a covariate (rates force the coefficient on log population to 1.0). To get percentage impacts of lockdown from the log outcome, I use $100 \times (e^{\hat{\beta}-0.5\hat{V}(\hat{\beta})} - 1)$ with confidence intervals from the approximate unbiased variance estimator of Van Garderen and Shah (2002).

The regressions use 22 control variables, including county population and density, the elder share, the share in nursing homes, nine other demographic and economic characteristics and a set of regional fixed effects.[7] These controls explain about two-thirds of variation in log deaths (as of mid-May). Even with these controls, the errors for the log death equations may correlate with treatment status, if selection into the treatment group (77% of counties) is due to unobservables.[8] Political drivers of lockdown are plausible instruments; counties without lockdown are all in states with Republican governors (overall, 26 states have Republican governors) and lockdown was more likely if a

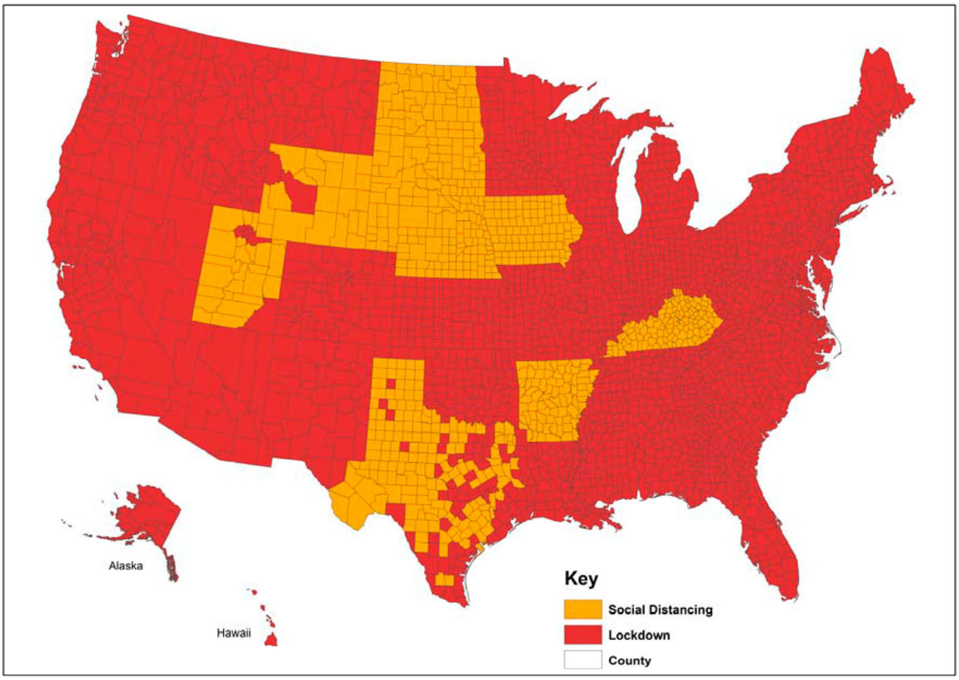


**Figure 2.** County-level variation in lockdown orders as of early April, 2020.

gubernatorial election is set for November 2020 (these elections are due in 11 states). Conditional on the state-level factors, the extent that a county became more partisan between the 2012 and 2016 Presidential elections, relative to the state-level change, affects odds of lockdown. This county-level relative change in partisanship has a mean of zero, by construction, but the standard deviation swing in partisanship was 7.7 percentage points.[9] It is hard to think of other paths for these variables to affect Covid-19 deaths than via political calculations about lockdown. I use a control function version of IV, with first stage residuals added to OLS outcome equations, because the percentage impact estimator (and its variance) is based on OLS.

A final issue on estimators is possible spatial autocorrelation. Counties neighbouring a county with unexplainably more deaths may have more deaths, given epidemic spread of Covid-19. I cluster at state level, to allow correlations in errors for counties in the same state. Clustered errors can be conservative, in not letting intra-cluster correlations vary and in not allowing between-cluster correlations (Gibson, Kim, & Olivia, 2014). As a variant I also use a spatially autoregressive model with autoregressive errors (SARAR), estimated by generalized spatial two-stage least squares (Drukker, Prucha, & Raciborski, 2013). This lets errors correlate across neighbouring counties (and neighbours of neighbours), and allows for spatial spill-overs in deaths (see Appendix 3 for details).

The main regression results are in Table 1. The first column has the first-stage results, for which counties have lockdown. The $F$-test for excluding the instruments is 4.1 ($p < 0.02$) using clustered standard errors or 46 ($p < 0.01$) using the spatial error model. The remaining columns have OLS and IV results for cumulative deaths at three dates matching key decisions made in New Zealand: 23 March when Level 3 was announced with the two-day warning for Level 4; 20 April when Level 4 was extended; and, 11 May when a staged move to Level 2 over 10 days was announced. The aim of showing results for these dates is to see how any evidence evolved for whether lockdowns reduce Covid-19 deaths; the data were available to New Zealand decision-makers at the time so it is not a question of being wise in hindsight.

**Table 1.** County-level impacts of lockdowns on Covid-19 deaths.

| | First stage model | ln (deaths, by March 23) | | ln (deaths, by April 20) | | ln (deaths, by May 11) | |
|---|---|---|---|---|---|---|---|
| | (Lockdown = 1) | OLS | IV/CF | OLS | IV/CF | OLS | IV/CF |
| Lockdown (= 1, otherwise social distancing) | | −0.028 | 0.112 | 0.017 | −0.062 | −0.007 | −0.154 |
| | | (0.024) | (0.154) | (0.127) | (0.667) | (0.155) | (0.665) |
| Residuals (from first stage for lockdown) | | | −0.150 | | 0.085 | | 0.159 |
| | | | (0.163) | | (0.675) | | (0.658) |
| ln (county population, 2019) | 0.078** | 0.035 | 0.024 | 0.506*** | 0.512*** | 0.666*** | 0.678*** |
| | (0.037) | (0.024) | (0.026) | (0.067) | (0.089) | (0.070) | (0.091) |
| ln (county population density) | −0.015 | 0.048* | 0.051* | 0.120 | 0.119 | 0.112 | 0.109 |
| | (0.029) | (0.027) | (0.027) | (0.072) | (0.076) | (0.073) | (0.078) |
| Share of county age 75 years or older | −1.241 | 1.924*** | 2.054*** | 6.415*** | 6.341*** | 6.908*** | 6.771*** |
| | (1.229) | (0.666) | (0.718) | (1.965) | (1.986) | (2.011) | (1.939) |
| Share of county population in nursing homes | −1.616* | 0.854** | 1.151** | 3.352** | 3.184 | 4.369** | 4.055 |
| | (0.944) | (0.373) | (0.498) | (1.597) | (2.053) | (1.921) | (2.420) |
| Male share of county population | 0.265 | 0.878*** | 0.836*** | 3.262*** | 3.286*** | 3.187*** | 3.231*** |
| | (0.414) | (0.268) | (0.264) | (0.963) | (1.014) | (1.030) | (1.083) |
| White share of county population | −0.087 | −0.337** | −0.325** | −1.119** | −1.126** | −1.524*** | −1.536*** |
| | (0.373) | (0.159) | (0.155) | (0.423) | (0.432) | (0.457) | (0.477) |
| Black share of county population | 0.585 | −0.110 | −0.181 | 0.750 | 0.790 | 1.243* | 1.318** |
| | (0.389) | (0.160) | (0.181) | (0.663) | (0.648) | (0.719) | (0.638) |
| ln (median income for county, 2010 census) | 0.059 | 0.328*** | 0.319*** | 1.574*** | 1.579*** | 1.691*** | 1.701*** |
| | (0.111) | (0.111) | (0.109) | (0.291) | (0.297) | (0.315) | (0.323) |
| Gini coefficient | 0.166 | 0.668*** | 0.657*** | 3.215*** | 3.221*** | 3.022*** | 3.033*** |
| | (0.381) | (0.179) | (0.179) | (0.625) | (0.617) | (0.576) | (0.567) |
| Unemployment rate | −0.154 | −0.210 | −0.306 | −2.011 | −1.956 | −2.548 | −2.446 |
| | (1.139) | (0.307) | (0.301) | (1.465) | (1.576) | (1.645) | (1.781) |
| Share of housing units with rental occupier | −0.691*** | 0.392 | 0.489 | 0.362 | 0.307 | 0.245 | 0.143 |
| | (0.239) | (0.273) | (0.309) | (0.905) | (0.891) | (1.012) | (0.938) |
| Share of population without health insurance | 2.359** | 1.072*** | 0.746 | 3.976*** | 4.160* | 4.697*** | 5.041** |
| | (1.064) | (0.398) | (0.526) | (1.427) | (2.303) | (1.626) | (2.451) |
| Adult smoking rate | −0.205 | −0.207* | −0.206* | −1.248** | −1.248** | −1.380** | −1.380** |
| | (0.559) | (0.107) | (0.106) | (0.526) | (0.527) | (0.586) | (0.587) |
| Gubernatorial election set for 2020 | 0.147 | | | | | | |
| | (0.094) | | | | | | |
| Governor is Democratic | 0.170 | | | | | | |
| | (0.127) | | | | | | |
| Change in county partisanship, 2012–16 | 0.358* | | | | | | |
| | (0.216) | | | | | | |
| Constant | −0.633 | −4.514*** | −4.427*** | −23.164*** | −23.213*** | −25.044*** | −25.136*** |
| | (1.310) | (1.233) | (1.216) | (3.201) | (3.261) | (3.439) | (3.510) |
| *R*-squared | 0.430 | 0.238 | 0.240 | 0.600 | 0.600 | 0.639 | 0.639 |

Notes: Models also include nine fixed effects for US regions. Robust standard errors in () clustered at state level, ***, **, * denote statistical significance at 1%, 5% and 10% levels, $N = 3109$ US counties.

There is no evidence that counties with a lockdown have fewer deaths. For all three dates, the coefficient on lockdown is statistically insignificant.[10] Given the strength of the instruments (e.g. an $F$-test of 46 for excluding them, with the spatial model), the insignificant effects of lockdown are unlikely due to weak instruments. A test of over-identifying restrictions also reveals no concerns ($p < 0.18$).

It typically takes three weeks or more for a SARS-CoV-2 infection to cause Covid-19 death (Homburg, 2020) so early April lockdowns should show effects by May. To monitor this, Figure 3 shows percentage impacts (and 95% confidence intervals) of lockdown on Covid-19 deaths (cumulative), from models estimated every Monday from 23 March until 1 June.[11] On just two of 44 test occasions (11 Mondays over four models) do 95% confidence intervals exclude zero (25 May and 1 June, for the spatial control function approach). Adjusting for multiple hypothesis testing, using a bonferroni correction, requires significance at the $\alpha/n$ level, which is .0011. This is 12-times smaller than the actual

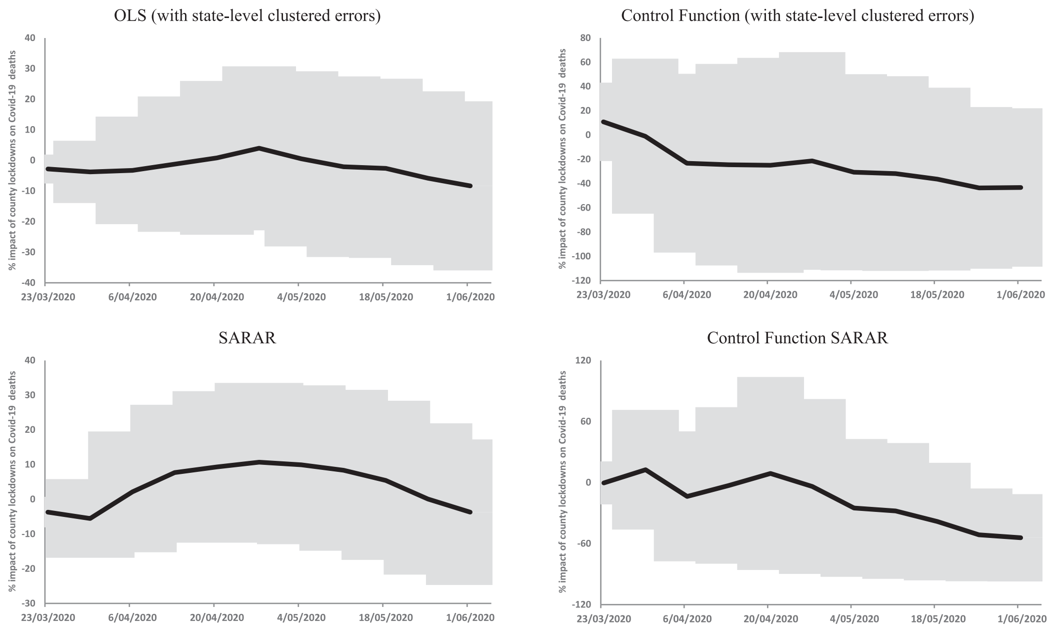


**Figure 3.** Evolving estimates of the impact of county lockdowns on Covid-19 deaths.
Note: Shaded regions show 95% confidence intervals.

*p*-value. So the firmest conclusion is that over more than two months after New Zealand's 23 March lockdown decision, there was no evidence of more Covid-19 deaths in counties without lockdowns. In terms of the confidence intervals, for the OLS results these range from −2% to 8% on 23 March, −24% to 26% on 20 April and −32% to 27% on 11 May, and with the SARAR estimator they tend to have larger positive values (more deaths with lockdowns) and negative values closer to zero (−13% and −15% on 20 April and 11 May).

This statistical insignificance of the lockdown treatment variable is not from the failure of the models to explain cross-county death patterns; about two-thirds of variation is explained by early May. The models show deaths are higher if the elderly or those in nursing homes are more of the population; patterns noted in popular discussion of Covid-19. Deaths are higher if whites are a lower share and blacks a higher share of the population, as noted by Millett *et al.* (2020). Counties with higher inequality and more people without health insurance experience more deaths. Fewer deaths occur if the smoking rate is higher, similar to what is found in the U.K. for 17 million NHS patients, where Williamson et al (2020) find current smokers less likely than others to die (as hospital in-patients) with confirmed COVID-19.

Five sensitivity analyses confirm the result that lockdowns are ineffective at reducing Covid-19 deaths. The first weights by county population; the 5th percentile county has under 3000 people while the 95th percentile has 450,000 so a case can be made for more weight on populous counties. The second uses death rates (by 11 May). The third uses IV-Poisson count data models, and the fourth uses LIML which may be preferred if there are weak instruments. In all four of these alternative approaches, lockdowns have no impact on Covid-19 deaths. The last sensitivity analysis is just for Texas, which had a more even split of 89 counties with lockdown and 165 with social distancing. The IV results show no effect of lockdown but with OLS it seems that counties with a lockdown have more deaths – a pattern strengthening over time (e.g. lockdown counties have 37.1% (SE = 18.6%) more deaths by 11 May).[12]

## 3. Summary and implications for New Zealand

Lockdowns are ineffective at reducing Covid-19 deaths. Variation amongst counties in the United States, where over one-fifth had no lockdown, shows no impact of lockdowns. Specifically, one cannot reject the hypothesis of zero difference in deaths between lockdown and non-lockdown counties. Using these results to inform a counterfactual of what would have happened if New Zealand had not gone into a Level 4 lockdown faces the criticism that the setting is different. Yet it is a universal force of human nature – privately taking steps to reduce exposure to a new risk – that likely makes lockdown superfluous. Moreover, evidence from elsewhere suggests that lockdowns were either superfluous (Homburg, 2020; Stone, 2020) or cause total deaths to rise because of non-Covid mortality (Williams, Crookes, Glass, & Glass, 2020).

A non-economist might say 'what difference does it make?' If people would reduce interactions anyway, due to perceived Covid-19 risks, having government force them to stay home would seem costless. Yet as economists know, a government *diktat* approach runs into the central planning problem; no central planner has all the information (collectively) held by parties involved in voluntary exchange (Hayek, 1945). For example, absent lockdown, if a butcher felt they could operate safely and if customers felt they could safely shop at this butchery, voluntary and beneficial exchange could occur. Instead, under the central planning approach applied in New Zealand, butchers were shut but supermarkets selling meat were not. Potentially, much economic surplus (for both consumers and producers) was lost.

In terms of implications for the future, these results add to the evidence that lockdowns are ineffective. This was also the prior view in public health; for example, Inglesby *et al.* (2006, p. 371) noted: 'It is difficult to identify circumstances in the past half-century when large-scale quarantine has been effectively used in the control of any disease.' So when the next pandemic occurs, the Covid-19 lockdowns should not be considered a success that should be replicated. In terms of the (recent) past, the ineffectiveness of lockdowns implies that New Zealand suffered large output losses, of 10 billion dollars or more according to Treasury figures, for no likely benefit in terms of lives saved as a result of the decision to move almost immediately from Level 2 to Level 4. Notably, this decision went against Ministry of Health advice to stay at Level 2 for 30 days. If decision-making from March and April is reviewed, any claim that lockdown was necessary to save lives can be treated with strong scepticism. It is especially concerning that there were data available, on the dates of those key decisions, to show that lockdowns are ineffective at reducing Covid-19 deaths.

## Notes

1. One cost-benefit analysis subsequently published by a government agency concerns the April 20 decision to extend Level 4 by five days. Heatley (2020) calculates that this extension provided a benefit of 239 quality-adjusted life years (QALYs), at a cost of 22,453 QALYs. In other words, the costs were over 90 times larger than the benefits. In line with the current analysis, Heatley (2020) restricts attention to using data that were available to decision makers at the time of making decisions about lockdowns.
2. https://www.rnz.co.nz/news/political/412403/all-of-new-zealand-must-prepare-to-go-in-self-isolation-now-prime-minister
3. Nevertheless, deaths data have some problems of over-counting especially when positive Covid-19 tests are linked with subsequent deaths from any cause (Loke & Heneghan, 2020; Williams *et al.*, 2020). For the county level data used here, reporting is mostly from state public health agencies who all follow the same set of CDC guidelines for Covid deaths which are available here: https://www.cdc.gov/nchs/data/nvss/vsrg/vsrg03-508.pdf
4. Even U.S. state-level data may mislead; 75% of the variance in death rates is within rather than between states.
5. One issue with using U.S. data is whether lockdown means the same thing as in New Zealand. Fortunately, the OxCGRT stringency index based on containment and closure regulations has been extended to cover U.S. states (Hale, *et al.*, 2020). While not as stringent as New Zealand, which is inherent in New Zealand's response being the most stringent in the world, several states (e.g. Maryland) had stringency index values that exceeded those recorded for the United Kingdom (GBR) in Figure 1.

6. Many counties have zero deaths so the inverse-hyperbolic-sine transformation is used. This is identical to using logarithms for non-zero observations, but let zeros be used without resorting to crude adjustments like adding one to all values before logging (Gibson, Datt, Murgai, & Ravallion, 2017).
7. The 10 Standard Federal Administrative Regions (SFARs). With some instrumental variables defined at state level, using state fixed effects introduces a collinearity problem.
8. In other words, the treatment of lockdown is potentially endogenous. Some evidence for this is from Lurie, Silva, Yorlets, Tao, and Chan (2020) who show that Covid-19 case numbers were increasing faster during March, with shorter doubling times, in states that subsequently had a lockdown compared to those states that did not subsequently have a lockdown.
9. The descriptive statistics on the instruments and other variables are in Appendix 2.
10. Statistical insignificance of the coefficient on the first-stage residuals implies (via the added-variable form of the Hausman test) that potential selection on unobservables (in terms of which counties have lockdown) may not cause significant bias in OLS results.
11. Deaths after June 1 likely reflect changed treatment since the lockdowns in early April, including restrictions being relaxed from early May in many states and the large Black Lives Matter protests from 26 May. The falling trend in 7-day averages of cases, and in deaths (lagged 23 days, following Homburg, 2020) both reversed in early June, suggesting that factors from early April that caused differences in Covid-19 deaths by county were being supplanted by more recent driving forces. Perhaps for this reason, over-identification tests are statistically significant in June even though they were not in March and April so less weight should be placed on the control function results from later dates.
12. Only 7% of Texas counties with (or soon to) lockdown had a Covid-19 death by 23 March (6% nationally) so it was not deaths driving lockdown. Two months later, by 18 May, the risk a county had any Covid-19 deaths, conditional on having no deaths by 23 March, had increased by significantly more for lockdown counties compared to those that did not lockdown, further suggesting ineffectiveness of lockdowns.

## Acknowledgements

Helpful comments from the editor and an anonymous referee and assistance with the mapping from Geua Boe-Gibson are acknowledged. These are the views of the author.

## Disclosure statement

No potential conflict of interest was reported by the author(s).

## References

Chetty, R., Friedman, J. N., Hendren, N., & Stepner, M. (2020). Real-time economics: A new platform to track the impacts of COVID-19 on people, businesses, and communities using private sector data. https://opportunityinsights.org/wp-content/uploads/2020/05/tracker_paper.pdf

Cronin, C., & Evans, W. (2020). Private precaution and public restrictions: What drives social distancing and industry foot traffic in the COVID-19 era? *Working Paper* No. w27531, National Bureau of Economic Research.

Daalder, M. (2020). Long read: The month that changed New Zealand. https://www.newsroom.co.nz/2020/05/13/1168837/long-read-the-month-that-changed-new-zealand

Dasaratha, K. (2020). Virus dynamics with behavioural responses. *mimeo* arXiv:2004.14533.

Drukker, D., Prucha, I., & Raciborski, R. (2013). Maximum likelihood and generalized spatial two-stage least-squares estimators for a spatial-autoregressive model with spatial-autoregressive disturbances. *Stata Journal*, *13*(2), 221–241.

Ferguson, N., Laydon, D., Nedjati-Gilani, G., Imai, N., Ainslie, K., Baguelin, M., ... Ghani, A. (2020). Impact of non-pharmaceutical interventions (NPIs) to reduce COVID-19 mortality and healthcare demand. *Mimeo* Imperial College.

Gardner, J., Willem, L., van der Wijngaart, W., Kamerlin, S., Brusselaers, N., & Kasson, P. (2020). Intervention strategies against COVID-19 and their estimated impact on Swedish healthcare capacity. *medRxiv*, doi:10.1101/2020.04.11.20062133

Gibson, J., Datt, G., Murgai, R., & Ravallion, M. (2017). For India's rural poor, growing towns matter more than growing cities. *World Development*, *98*(1), 413–429.

Gibson, J., Kim, B., & Olivia, S. (2014). Cluster-corrected standard errors with exact locations known: An example from rural Indonesia. *Economics Bulletin*, *34*(3), 1857–1863.

Giesecke, J. (2020). The invisible pandemic. *The Lancet*, *6736*(20), 31035–7.

Goolsbee, A., & Syverson, C. (2020). Fear, lockdown, and diversion: Comparing drivers of pandemic economic decline. *Working Paper* No. w27432, National Bureau of Economic Research.

Hale, T., Atav, T., Hallas, L., Kira, B., Phillips, T., Petherick, A., ... Pott, A. (2020). Variation in US states' responses to Covid-19. *Working Paper* No. 2020/034, Blavatnik School of Government, University of Oxford, Oxford UK.

Hale, T., Webster, S., Petherick, A., Phillips, T., & Kira, B. (2020). *Oxford COVID-19 Government response tracker*. Oxford: Blavatnik School of Government, University of Oxford.
Harrison, I. (2020). The Ministry of Health's modelling of the impact of the Coronavirus on New Zealand: A look behind the headlines. *mimeo* Tailrisk Economics, Wellington. http://www.tailrisk.co.nz/documents/Corona.pdf
Hayek, F. A. (1945). The use of knowledge in society. *American Economic Review*, *35*(4), 519–530.
Heatley, D. (2020). A cost benefit analysis of 5 extra days at Covid-19 alert level 4. *Research Note* 2020/02, New Zealand Productivity Commission.
Homburg, S. (2020). Effectiveness of Corona Lockdowns: Evidence for a Number of Countries. *Hannover Economic Papers* No. dp-671. Leibniz Universität Hannover, Wirtschaftswissenschaftliche Fakultät.
Inglesby, T., Nuzzo, J. B., O'Toole, T., & Henderson, D. A. (2006). Disease mitigation measures in the control of pandemic influenza. *Biosecurity and Bioterrorism: Biodefense Strategy, Practice, and Science*, *4*(4), 366–375.
Korolev, I. (April 20, 2020). Identification and Estimation of the SEIRD Epidemic Model for COVID-19. http://dx.doi.org/10.2139/ssrn.3569367
Loke, Y., & Heneghan, C. (2020). Why no-one can ever recover from Covid-19 in England – a statistical anomaly. Centre for Evidence-Based Medicine, University of Oxford. https://www.cebm.net/covid-19/why-no-one-can-ever-recover-from-covid-19-in-england-a-statistical-anomaly/
Lurie, M., Silva, J., Yorlets, R., Tao, J., & Chan, P. (2020). Coronavirus disease 2019 epidemic doubling time in the United States before and during stay-at-home restrictions. *The Journal of Infectious Diseases*, *222*(10), 1601–1606.
Millett, G., Jones, A., Benkeser, D., Baral, S., Mercer, L., Beyrer, C., … Sullivan, P. 2020. Assessing differential impacts of COVID-19 on Black communities. *Annals of Epidemiology*, *47*(1), 37âĂŞ44.
Stone, L. (2020). Lockdowns don't work. *Public Discourse* (April 21, 2020) https://www.thepublicdiscourse.com/2020/04/62572/
Toxvaerd, F. (2019). Rational disinhibition and externalities in prevention. *International Economic Review*, *60*(4), 1737–1755.
Treasury. (2020). *Treasury Report T2020/973: Economic scenarios* - 13 April 2020. https://treasury.govt.nz/publications/tr/treasury-report-t2020-973-economic-scenarios-13-april-2020
Van Garderen, K., & Shah, C. (2002). Exact interpretation of dummy variables in semilogarithmic equations. *The Econometrics Journal*, *5*(1), 149–159.
Williams, S., Crookes, A., Glass, K., & Glass, A. (2020). An improved measure of deaths due to COVID-19 in England and Wales. *Working Paper* No 3635548, SSRN. http://doi.org/10.2139/ssrn.3635548
Williamson, E., Walker, A., Bhaskaran, K., Bacon, S., Bates, C., Morton, C., … Goldacre, B. (2020). OpenSAFELY: Factors associated with COVID-19-related hospital death in the linked electronic health records of 17 million adult NHS patients. *Nature*, *584*(7821):430âĂŞ436. http://doi.org/10.1038/s41586-020-2521-4.
Wilson, N., Barnard, L., Kvalsvig, A., & Baker, M. (2020). Potential Health Impacts from the COVID-19 Pandemic for New Zealand if Eradication Fails: Report to the NZ Ministry of Health.

# Appendices

## Appendix 1

**Data sources**

County-level daily data on cumulative deaths related to Covid-19 are obtained from the aggregation site: https://usafacts.org/visualizations/coronavirus-covid-19-spread-map/ who source the raw data from the Centers for Disease Control and Prevention (CDC), and from state- and local-level public health agencies. The data file starts from 22 January 2020 and updates daily. The Wayback Machine digital archive indicates that the website aggregating the data has been available at least since March 17, 2020. In a few cases, deaths are attributed to a state but not to a county (e.g. from the Grand Princess cruise ship). Where local agencies changed their methodology in reporting deaths due to COVID-19 the county-level counts are retroactively adjusted, so the most recent data file should always be used.

Data on which counties had Government Ordered Community Quarantine (*aka* 'Lockdown') and which had Government Directed Social Distancing are from the Disaster Response Program of ESRI (the supplier of ArcGIS and related software and data): https://coronavirus-disasterresponse.hub.arcgis.com/app/ebe29d4c1fca4ac292d00dbd54ed37e9. The data in the map are compiled and periodically updated by the American Red Cross based on documents made publicly available by State, Tribal, Territorial and municipal governments. These data have been compiled by the Red Cross since at least March 14, 2020.

The politics data used as the instrumental variables come from two sources. The county level variable that measures changes in the Republican-Democratic gap in the vote share from the 2012 to the 2016 Presidential elections comes from the MIT Election Data and Science Lab: 'County Presidential Election Returns 2000-2016', https://doi.org/10.7910/DVN/VOQCHQ The data on States having gubernatorial elections in 2020 is from the National Governors Association: https://www.nga.org/governors/elections/ and the data on the party affiliation of the incumbent governor for each State is from: https://www.nga.org/governors/

The control variables come from four sources. The estimated population of each county in 2019 is from the same source as the number of Covid-19 deaths (https://usafacts.org) in order to use a population denominator as close in time to the deaths as possible. The population density and demographic ratios (shares of population who are: age 75 or older; white; black; male; and, renters) are originally from the 2010 census, reported at the ArcGIS Hub for USA Counties (http://hub.arcgis.com/datasets/48f9af87daa241c4b267c5931ad3b226_0). The ratios use the 2010 population counts as the denominator, rather than the 2019 population estimates reported at usafacts.org in order to be internally consistent. The median earnings, the Gini coefficient, the unemployment rate, the share uninsured and the smoking rate are from the MIT Election Data and Science Lab, whose URL is given above. These variables are also originally from the 2010 Census. Data on rest homes are from the Skilled Nursing Facilities Quality Reporting Program, covering all Medicare and Medicaid-certified nursing homes, (available here: https://www.medicare.gov/nursinghomecompare/search.html). For each of the $n = 15{,}436$ nursing homes, the certified number of beds and the total number of residents is reported, along with the facility address (including ZIP code). A few lacking resident counts get given an imputed value based on the number of beds. The estimated count of residents is aggregated to ZIP code level and then to county level using the ZIP code to FIPS crosswalk provided here: https://www.huduser.gov/portal/datasets/usps_crosswalk.html. The number of nursing home residents is expressed as a ratio to the 2019 estimate of the county population.

## Appendix 2

Summary statistics

| | Mean | Std Dev | Min | Max |
|---|---|---|---|---|
| Deaths (as of May 11) | 25.148 | 205.144 | 0.000 | 6024.000 |
| Death rate, deaths per million (as of May 11) | 89.327 | 223.659 | 0.000 | 3098.259 |
| Lockdown ( = 1, otherwise social distancing) | 0.770 | 0.421 | 0.000 | 1.000 |
| ln (county population, 2019) | 10.288 | 1.490 | 5.130 | 16.122 |
| ln (county population density) | 3.815 | 1.730 | −2.110 | 11.199 |
| Share of county age 75 years or older | 0.072 | 0.024 | 0.013 | 0.203 |
| Share of county population in nursing homes | 0.011 | 0.012 | 0.000 | 0.233 |
| Male share of county population | 0.488 | 0.035 | 0.238 | 0.744 |
| White share of county population | 0.812 | 0.166 | 0.095 | 0.992 |
| Black share of county population | 0.088 | 0.147 | 0.000 | 0.879 |
| ln (median income for county, 2010 census) | 10.126 | 0.193 | 8.576 | 10.945 |
| Gini coefficient | 0.432 | 0.036 | 0.207 | 0.645 |
| Unemployment rate | 0.077 | 0.028 | 0.008 | 0.283 |
| Share of housing units with rental occupier | 0.233 | 0.075 | 0.042 | 0.763 |
| Share of population without health insurance | 0.179 | 0.054 | 0.031 | 0.389 |
| Adult smoking rate | 0.212 | 0.059 | 0.031 | 0.511 |
| Gubernatorial election set for 2020 | 0.182 | 0.386 | 0.000 | 1.000 |
| Governor is Democratic | 0.436 | 0.496 | 0.000 | 1.000 |
| Relative change in county partisanship 2012–16 | 0.000 | 0.077 | −0.413 | 0.466 |

Notes: Based on $N = 3109$ U.S. counties.

## Appendix 3: Sensitivity analyses using spatial autoregressive models with spatial errors

County-level impacts of lockdowns on Covid-19 deaths – total of the direct and indirect impacts allowing for spill-overs.

| | First stage model | ln (deaths, by March 23) | | ln (deaths, by April 20) | | ln (deaths, by May 11) | |
|---|---|---|---|---|---|---|---|
| | (Lockdown = 1) | SPREG | IV/CF | SPREG | IV/CF | SPREG | IV/CF |
| Lockdown ( = 1, otherwise social distancing) | | −0.034 | −0.001 | 0.095 | 0.224 | 0.088 | −0.167 |
| | | (0.022) | (0.104) | (0.102) | (0.471) | (0.110) | (0.504) |
| Residuals (from first stage for lockdown) | | | −0.037 | | −0.136 | | 0.268 |
| | | | (0.108) | | (0.483) | | (0.517) |
| ln (county population, 2019) | 0.025*** | 0.047*** | 0.048*** | 0.774*** | 0.774*** | 0.923*** | 0.931*** |
| | (0.007) | (0.014) | (0.015) | (0.065) | (0.066) | (0.066) | (0.068) |
| ln (county population density) | 0.000 | 0.048*** | 0.050*** | 0.163*** | 0.163*** | 0.178*** | 0.180*** |
| | (0.007) | (0.013) | (0.014) | (0.047) | (0.048) | (0.051) | (0.051) |
| Share of county age 75 years or older | −0.610** | 1.973*** | 2.095*** | 12.743*** | 12.875*** | 12.358*** | 12.274*** |
| | (0.238) | (0.495) | (0.536) | (1.866) | (1.895) | (1.878) | (1.898) |
| Share of county population in nursing homes | −0.063 | 1.488** | 1.571** | 5.380** | 5.423** | 5.871*** | 5.833** |
| | (0.303) | (0.660) | (0.700) | (2.141) | (2.154) | (2.271) | (2.281) |
| Male share of county population | −0.011 | 0.997*** | 1.052*** | 3.638*** | 3.661*** | 3.308*** | 3.309*** |
| | (0.132) | (0.295) | (0.315) | (0.938) | (0.943) | (0.983) | (0.987) |
| White share of county population | 0.001 | −0.344*** | −0.363*** | −2.138*** | −2.148*** | −2.285*** | −2.294*** |
| | (0.056) | (0.109) | (0.116) | (0.406) | (0.408) | (0.423) | (0.424) |
| Black share of county population | 0.109 | −0.102 | −0.110 | 0.783* | 0.774* | 1.154** | 1.171** |
| | (0.069) | (0.105) | (0.112) | (0.440) | (0.442) | (0.470) | (0.472) |
| ln (median income for county, 2010 census) | 0.010 | 0.377*** | 0.392*** | 1.602*** | 1.608*** | 1.535*** | 1.542*** |
| | (0.028) | (0.069) | (0.074) | (0.210) | (0.211) | (0.214) | (0.214) |
| Gini coefficient | 0.146 | 0.854*** | 0.893*** | 3.324*** | 3.331*** | 3.130*** | 3.177*** |
| | (0.122) | (0.273) | (0.290) | (0.872) | (0.878) | (0.912) | (0.918) |
| Unemployment rate | 0.187 | −0.327 | −0.381 | −5.702*** | −5.782*** | −5.663*** | −5.597*** |
| | (0.221) | (0.404) | (0.435) | (1.530) | (1.548) | (1.598) | (1.612) |
| Share of housing units with rental occupier | −0.161** | 0.413*** | 0.442*** | 0.609 | 0.638 | 0.432 | 0.395 |
| | (0.076) | (0.157) | (0.168) | (0.514) | (0.523) | (0.546) | (0.553) |
| Share of population without health insurance | 0.681*** | 1.326*** | 1.359*** | 4.214*** | 4.142*** | 5.624*** | 5.834*** |
| | (0.138) | (0.276) | (0.302) | (0.932) | (0.993) | (1.005) | (1.075) |
| Adult smoking rate | 0.037 | −0.120 | −0.133 | −1.825*** | −1.843*** | −1.446*** | −1.429*** |
| | (0.075) | (0.139) | (0.148) | (0.509) | (0.513) | (0.532) | (0.535) |
| Gubernatorial election set for 2020 | 0.181*** | | | | | | |
| | (0.019) | | | | | | |
| Governor is Democratic | 0.111*** | | | | | | |
| | (0.016) | | | | | | |
| Change in county partisanship, 2012–16 | 0.045 | | | | | | |
| | (0.063) | | | | | | |
| Spatial lag of the error term | 0.053*** | 0.025*** | 0.023*** | 0.033*** | 0.033*** | 0.034*** | 0.033*** |
| | (0.001) | (0.006) | (0.006) | (0.002) | (0.002) | (0.002) | (0.002) |
| Spatial lag of the dependent variable | | 0.014*** | 0.016*** | 0.020*** | 0.020*** | 0.017*** | 0.017*** |
| | | (0.005) | (0.005) | (0.002) | (0.002) | (0.002) | (0.002) |
| Pseudo *R*-squared | 0.343 | 0.232 | 0.230 | 0.575 | 0.575 | 0.623 | 0.623 |

Notes: Models also include nine fixed effects for US regions and an intercept. In the first column and in the spatial lag rows, cell values are coefficient estimates, otherwise they are the average total impacts, taking into account direct and indirect impacts from spill-overs operating through the spatially lagged dependent variable. For those cells, the standard errors in () are calculated from delta method, otherwise they are from a heteroscedasticity-robust GMM variance estimator. Estimation uses generalized spatial two-stage least squares, with a second-order contiguity weighting matrix, ***, **, * denote statistical significance at 1%, 5% and 10% levels, $N = 3104$ US counties.