EERI

Economics and Econometrics Research Institute

<div>

Does Testing for Coronavirus reduce Deaths?

Weshah Razzak

</div>

Does Testing for Coronavirus reduce Deaths? [*]

W A Razzak

School of Economics and Finance, Massey University, New Zealand

April 2020

Abstract

We examine the effect of testing for Coronavirus on deaths in eight countries over the month of March 2020 by estimating a fixed-effect regression model using the Generalized Method of Moments (GMM). On average, the data reject the hypothesis that "testing" for the virus *does not* affect death. By country, however, we reject the hypothesis in two countries at the 5 percent level, in three countries at the 10 percent level, and could not reject it in three other countries. On average, testing for the virus is an important element of the health policy.

. * All correspondence to W A Razzak, Honorary Research Fellow at the School of Economics and Finance, Massey University, New Zealand at razzakw@gmail.com or w.razzak@massey.ac.nz

1. Introduction

The World Health Organization (WHO) emphasizes that testing for Coronavirus is an essential pillar in the strategy for fighting against the virus. Testing determines with some accuracy the number of people who need treatments and isolations, thus reduces death. In general, there are large variances in the testing data across countries. Some countries have done more tests than others have. Others began testing earlier than later and therefore have more data. Some countries indicated that they might stop testing altogether. There are reported shortages of testing kits. Some countries have followed more efficient testing strategies than others, thus more tests.

Does more testing for the Coronavirus in people reduce death? This is the main question we try to answer in this paper. We use two data sets, one includes deaths and infections, which is published by the EU and covers the entire world. There are missing data because countries vary in reporting statistics. The other is a smaller data set published by Oxford University on tests. It reports cumulative tests per million people. To have a balanced panel we could only find eight countries, Austria, Belgium, Iceland, Italy, Japan, South Korea, U.K. and the U.S. to have data from March 1 to March 31, 2020.

Visually, deaths and tests are negatively correlated. Figure (1) plots the data. We test the hypothesis that tests for Coronavirus do not affect deaths by estimating a linear State Dependence first-order dynamic model using the Generalized Method of Moments (GMM). The relationship between deaths and tests, which results from this Instrumental Variable method could be interpreted as causal. However, the lack of data on the results of the tests, i.e., negative and positive, and on treatments makes the causal interpretation difficult. Such difficulty, notwithstanding, estimation of a robust GMM coefficients between tests and deaths is informative.

We estimate semi-elasticity across the panel and reject the null hypothesis on average. Then we allow the country slope to vary, and were able to reject the null hypothesis on different statistical levels. The hypothesis is strongly rejected in the cases of Italy and the U.S. i.e., tests significantly reduce deaths. We can reject the hypothesis in the cases of Belgium and the U.K., and possibly in Japan but not in Austria, Iceland, and South Korea.

Next we describe the data first because the data determined our estimation methodology then we present the model, estimation, and results. Section 3 is a conclusion.

2. Hypothesis, data, methodology, and results

2.1 Hypothesis

We are unaware of any empirical examination of the efficacy of "testing for Coronavirus". Does it reduce death seems like a reasonable question to ask, and by how much? Therefore, the objective of this paper is to test this hypothesis that testing for the Coronaviris has no effect on deaths.

2.2 Data

We begin by describing the data because the data determine our methodology. We have data for deaths and infections for more than 200 countries and territories, the EU Open Data Portal reports daily data. Countries reported infection cases at different dates, but the last observation for all of them is April 4, 2020.[1] Some countries and territories have reported less data than others have and some have literally one observation. For these reason we removed a few of these countries and territories. We identify 172 countries with reasonable amount of data. This data set does not report testing data. The data source and the countries are listed in the data appendix.

---

[1] This is the date of the beginning of writing this paper. The reporting of the data continues.

Another set of data by Oxford University, which report the cumulative number of tests for Coronavirus per millions of people by country is much smaller than the above data set. The data do not report the results of the tests so we do not know if some people tested positive or negative. There is no information about the methods, the institutions, etc. This data set is much smaller than the earlier one because fewer countries have tested systematically, and some countries tested much later in time, e.g., end of March or early April. We identified only eight countries only that have reported complete and *continuous* time series data for the cumulative number of tests from March 1 to March 31, 2020. Thus, we choose these countries in order to have a balanced panel *without missing values*. These countries are Austria, Belgium, Iceland, Italy, Japan, South Korea, U.K., and the U.S.

Figure (1) is a scatter plot of the percentage change in the number tests and deaths. All the correlations are negative, except for South Korea where there is no correlation. The cumulative number of tests per million people increased in all countries in March 2020. The challenge is to confirm these visually observed correlations (and perhaps causations) in regressions. We also want to measure the magnitude of the change in deaths due to testing.

2.3 Methodology

We fit a linear State-Dependence first-order dynamic model with an unobserved heterogeneity:

$$y_{it} = \alpha y_{it-1} + x'_{it}\beta + \varepsilon_{it}, \varepsilon_{it} = \delta_i + u_{it} \tag{1}$$

The dependent variable y_{it} is death. The regressors include a first-order dynamic term y_{it-1}, and a vector x_{it}, which denotes tests for the Coronavirus, δ_i denotes a fixed effect correlated with the regressors, and u_{it} are serially uncorrelated errors, which is a testable hypothesis.

The OLS coefficient estimates of equation (1), whether a fixed-effect model or a first-differenced transformation are biased and inconsistent, Therefore, we estimate equation (1) using the GMM to estimate a fixed-effect model with White cross-section instrument weighting matrix; and White cross-section standard errors and covariance. (see for example, Wooldridge (2002), Matyas and Sevestre (1996), Hyslop (1999) Gary (1984), Baltagi (1995), Arellano and Bover (1995), and Anderson and Hsiao (1982)).[2]

The instruments include a constant term, Δy_{it-3}, Δx_{it-2}, and infections and lagged infections.[3] Infections could lead to death, thus they could be used as instruments. Figure (2) is a scatter plot of infections and death. The correlation is positive in all cases except in the case of South Korea, where the correlation is weakly negative. Not all infected people die of course. More people recover. Nonetheless, death is proportional to infections, and if infection is zero, death must be zero too. We include infections contemporaneously and at lags 1, 3 and 4 in the instruments. We report two sets of results in table (1) and table (2). The first is an average estimate across all countries, and the second allows for the slope coefficient of tests to vary across countries.

In table (1), the lagged death variable (i.e., the first-order dynamic) is close to one, and highly significant, which indicates a significant persistence. The variable "Tests" has a significant negative coefficient so the null hypothesis that tests do not affect (or cause, in the case of GMM) deaths is rejected. We interpret the coefficient as a semi-elasticity.[4] It implies that on *average*, a one percent increase in testing reduces deaths by 4 a day across the 8 countries.

---

[2] Fixed-effect (i.e., taking the deviations from group means) is equivalent to first-difference transformation of y_{it} and x_{it} . This first differencing makes the differenced error term Δu_{it} negatively correlated with Δy_{it-1} which renders the estimated coefficients biased downwards, and inconsistent. $E(\Delta y_{it-1}\Delta u_{it}) = E\big(y_{it-1}(-u_{it-1})\big) = E(-u_{it-1}^2) = \sigma_u^2$.

[3] The instruments cannot have y_{it-1} and y_{it-2} because they would be correlated with the error term.

[4] We measured the variable "tests" in log-difference because the reported data were cumulative daily tests so the relevant variable is the first difference or log first difference. The log first difference implies that the coefficient β is a semi-elasticity). Further, this formulation is useful in case some of the countries under-report deaths.

Table (2) allows the slope coefficient β to vary across countries. The estimated semi-elasticity is significant in Italy, and the U.S. The semi-elasticity indicates that a one percent increase in testing for the virus reduces deaths by 68.5 and 12.6 a day in these two countries respectively. However, at the 10 percent level, the Belgium and U.K. data suggest that testing has a significant negative effect on deaths, more so in the case of Belgium. The estimated semi-elasticity is -1.65 and -31.8 thus a one percent increase in daily testing reduces deaths by 1.65 and 31.8 a day in Belgium and the U.K. respectively. The estimated semi-elasticity in Japan could be considered significant at a higher than 10 percent level, albeit much less significant than in Belgium and the U.K. It implies that a one percent increase in daily tests reduces deaths by 25. The results for Austria, Iceland, and South Korea are statistically insignificant.

3. Conclusions

We examined the effect of "testing" for Coronavirus on deaths in eight countries (Austria, Belgium, Iceland, Italy, Japan, South Korea, U.K., and the U.S.). We chose this panel only because a balanced panel exist for the period from March 1, 2020 to March 31, 2020. We estimated a State Dependence – a linear first-order dynamic model – using GMM, where by deaths depends on lagged deaths, and tests for the Coronavirus. Our instruments included infections, lagged infections, and appropriate (distanced) lags of deaths and tests. *On average and across the panel*, a one percent increase in tests reduces death by about 4 a day. When we allowed the effect of tests on deaths to vary across countries, we found that tests reduce deaths in Italy and the U.S. at the 5% significance level, in Belgium and the U.K. at the 10%

Suppose that a country reports a fraction ω of the deaths such that we observe $\omega_i y_{it}$ instead of y_{it}. Taking the log-difference eliminates the constant ω_i.

level, and at a lower significance level in Japan. The hypothesis that tests do not reduce deaths is rejected in the case of Austria, Iceland, and South Korea. We conclude that testing for the Coronavirus could be a useful pillar of the strategy to deal with the pandemic.

New Zealand is not included in the sample because the data were not reported for the full month of March. The implication of the results is that testing for COVID-19 is an important element of the health policy, and New Zealand should consider it and increase the number of tests in the population.

References

Anderson, T. W. and C. Hsiao. (1982). The Foundation and Estimation of Dynamic Models Using Panel Data, Journal of Econometrics, 18, 67-82.

Arellano, M. and O. Bover. (1995). Another Look at Instrumental Variables Estimation of Error Component Models, Journal of Econometrics, 68, 29-51.

Baltagi, B. (1995). <u>Econometric Analysis of Panel Data</u>, John Wiley and Sons.

Hyslop, D. (1999). State Dependence, Serial Correlation and Heterogeneity in International Labor Force Participation of Married Women, Econometrica, 67 (6), 1255-1294.

Matyas, L. and P. Sevestre (eds.,) (1996). <u>The Econometrics of Panel Data: Handbook of Theory and Applications</u>, Kluwer-Nijoff, Dordrecht.

Wooldridge, J. (2002). <u>Econometric Analysis of Cross Section and Panel Data</u>, MIT Press. MA.

Tables and Figures

<div align="center">

Table (1)
</div>

Dependent Variable: y_{it} (DEATH)
Method: Panel GMM EGLS (Cross-section weights)
Periods included: 27
Cross-sections included: 8
Total panel (balanced) observations: 216
White cross-section instrument weighting matrix
Linear estimation after one-step weighting matrix
White cross-section standard errors & covariance (d.f. corrected)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	70.84714	22.43690	3.157617	0.001*
y_{it-1}	0.962212	0.081554	11.79849	0.000*
x_{it}	-4.567294	1.619102	-2.820882	0.005*

<div align="center">

Weighted Statistics
</div>

Root MSE	64.69200	R-squared	0.710428
Mean dependent variable	58.03890	Adjusted R-squared	0.697777
S.D. dependent variable	109.1907	S.E. of regression	66.24359
Sum squared residuals	903971.8	Durbin-Watson stat	1.658484
J-statistic	5.275447	Instrument rank	13
Prob(J-statistic)	0.152704		

<div align="center">

Unweighted Statistics
</div>

R-squared	0.893153	Mean dependent var	78.16667
Sum squared residuals	782367.6	Durbin-Watson stat	2.174279

The instruments are y_{it-3}, infections, lags 1, 3, and 4 of infections.
Asterisk denotes significant At the 5% level. The J statistics P values
indicates that we cannot reject the validity of the over-identifying
restrictions.

Table (2)

Dependent Variable: y_{it}(DEATH)
Method: Panel GMM EGLS (Cross-section weights)
Sample (adjusted): Mar 5, 2020 – Mar 31, 2020
Periods included: 27
Cross-sections included: 8
Total panel (balanced) observations: 216
White cross-section instrument weighting matrix
Linear estimation after one-step weighting matrix
White cross-section standard errors & covariance (d.f. corrected)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	223.9346	111.0194	2.017076	0.0450*
y_{it-1}	0.497655	0.319181	1.559163	0.1205#
x_{it}				
Austria	-0.203774	0.328276	-0.620740	0.5355
Belgium	-1.657207	0.988440	-1.676588	0.0952#
Iceland	-0.003510	0.004648	-0.755221	0.4510
Italy	-68.49904	35.64869	-1.921502	0.0561*
Japan	-0.252101	0.218778	-1.152315	0.2506#
South Korea	0.064096	0.290564	0.220592	0.8256
U.K.	-31.80554	23.56964	-1.349428	0.1787#
U.S.	-12.63421	5.056308	-2.498701	0.0133*

Weighted Statistics

Root MSE	123.8680	R-squared	0.457505
Mean dependent variable	110.0778	Adjusted R-squared	0.413887
S.D. dependent variable	130.9937	S.E. of regression	129.0504
Sum squared residuals	3314147.	Durbin-Watson stat	2.495593
J-statistic	21.88009	Instrument rank	34
Prob(J-statistic)	0.189374		

Unweighted Statistics

R-squared	0.689874	Mean dependent var	78.16667
Sum squared residuals	2270842.	Durbin-Watson stat	1.996665

The instruments include Δy_{it-3} for each cross-section; infections and lags of infections 1, 3 and 4. Asterisk denotes statistically significant at the 5% level; #denotes statistically Significant at the 10% level. The J statistics P values indicates that we cannot reject the Validity of the over-identifying restrictions.
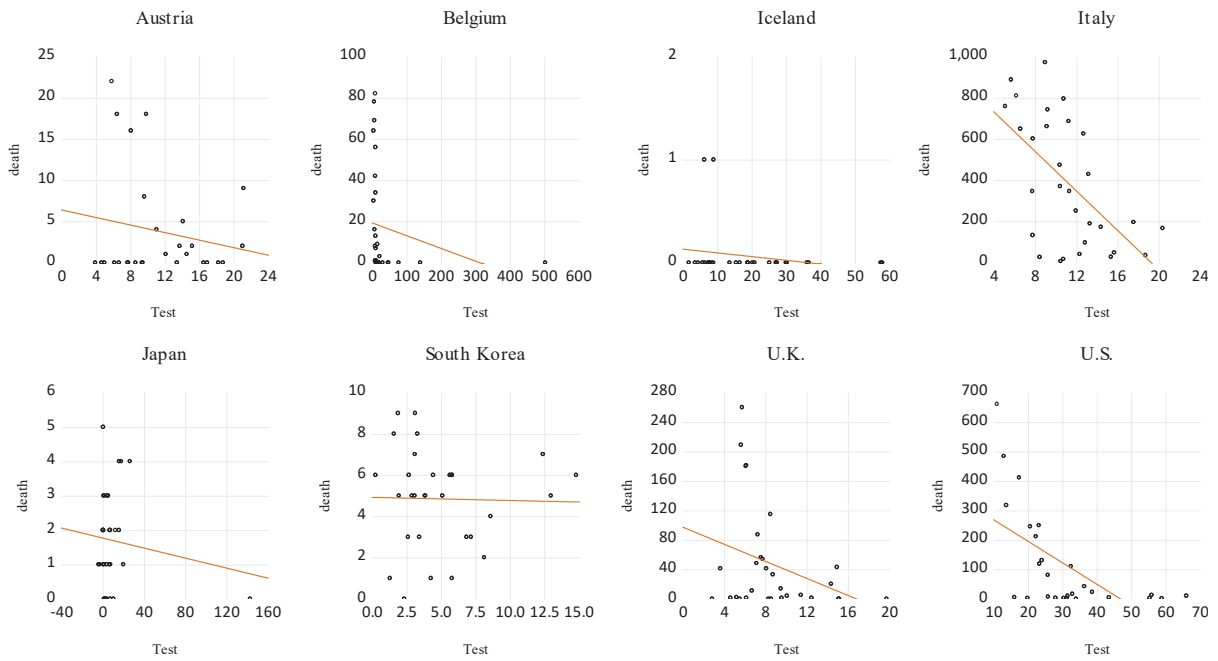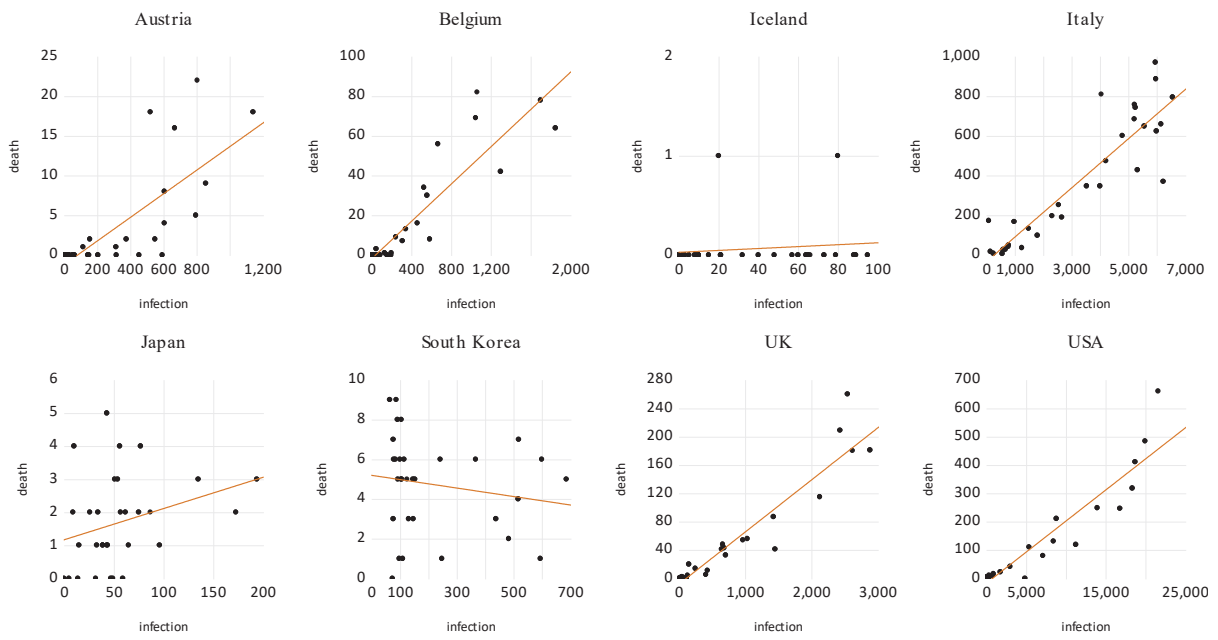
Figure (1)

Figure (2)

Data Appendix

The source is https://www.ecdc.europa.eu/sites/default/files/documents/COVID-19-geographic-disbtribution-worldwide.xlsx

The source for the tests is the University of Oxford
https://www.oxfordmartin.ox.ac.uk/global-development, and the 8 countries are Austria, Belgium, Iceland, Italy, Japan, South Korea, The U.K., and the U.S.