

**UNIVERSIDAD SAN FRANCISCO DE QUITO
USFQ**

Colegio de Administración y Economía

Factores que influyen en la mortalidad por COVID-19

Carlos Matthew Sagnay Ramírez

Economía

Trabajo de fin de carrera presentado como requisito
para la obtención del título de
Economista

Quito, 24 de Diciembre de 2020

**UNIVERSIDAD SAN FRANCISCO DE QUITO
USFQ**

Colegio de Administración y Economía

**HOJA DE CALIFICACIÓN
DE TRABAJO DE FIN DE CARRERA**

Factores que influyen en la mortalidad por COVID-19

Carlos Matthew Sagnay Ramírez

Carlos Uribe, PH.D

XXXXXXXXXXXXXXXX, XXXXXXXXXXXXXXXX

Quito, 24 de Diciembre de 2020

© DERECHOS DE AUTOR

Por medio del presente documento certifico que he leído todas las Políticas y Manuales de la Universidad San Francisco de Quito USFQ, incluyendo la Política de Propiedad Intelectual USFQ, y estoy de acuerdo con su contenido, por lo que los derechos de propiedad intelectual del presente trabajo quedan sujetos a lo dispuesto en esas Políticas.

Asimismo, autorizo a la USFQ para que realice la digitalización y publicación de este trabajo en el repositorio virtual, de conformidad a lo dispuesto en la Ley Orgánica de Educación Superior del Ecuador.

Nombres y apellidos: Carlos Matthew Sagnay Ramírez

Código: 00203845

Cédula de identidad: 1722713748

Lugar y fecha: Quito, 24 de Diciembre de 2020

ACLARACIÓN PARA PUBLICACIÓN

Nota: El presente trabajo, en su totalidad o cualquiera de sus partes, no debe ser considerado como una publicación, incluso a pesar de estar disponible sin restricciones a través de un repositorio institucional. Esta declaración se alinea con las prácticas y recomendaciones presentadas por el Committee on Publication Ethics COPE descritas por Barbour et al. (2017) Discussion document on best practice for issues around theses publishing, disponible en <http://bit.ly/COPETHeses>.

UNPUBLISHED DOCUMENT

Note: The following capstone project is available through Universidad San Francisco de Quito USFQ institutional repository. Nonetheless, this project – in whole or in part – should not be considered a publication. This statement follows the recommendations presented by the Committee on Publication Ethics COPE described by Barbour et al. (2017) Discussion document on best practice for issues around theses publishing available on <http://bit.ly/COPETHeses>.

RESUMEN

En este Reporte de Árbitro se analiza la investigación de Liang, Tseng, Ho & Wu con el título de “*Covid-19 mortality is negatively associated with test number and government effectiveness*”. El principal objetivo de este reporte es profundizar en los factores que influyen en la mortalidad por COVID-19 a nivel país, lo que es estudiado en la investigación mencionada. Como comentarios principales se incluye la mención de que existen subestimaciones en los errores estándar, así como comentarios acerca de la robustez y calidad de las variables. Además, al reconstruir la regresión con datos actualizados y variables más aptas, se encontró que la robustez encontrada en la investigación de base se perdió en los resultados de este reporte. Por lo tanto, si se utilizará este modelo, para profundizar en este tema, se lo debería usar cautelosamente.

Palabras clave: COVID-19, reporte de árbitro, mortalidad por COVID-19, ratio de positividad, robustez.

ABSTRACT

This Referee Report analyzes the research by Liang, Tseng, Ho & Wu, titled “*Covid-19 mortality is negatively associated with test number and government effectiveness*”. The main objective of this report is to delve into the factors that influence, at country level, COVID-19 mortality rate, which is studied in the aforementioned research. The main comments include the existence of underestimates in standard errors, as well as comments about the robustness and quality of some variables. Furthermore, when reconstructing the regression with updated data and more suitable variables, it was found that the robustness in the baseline investigation was lost in the results of this report. Therefore, as a final suggestion this model should be treated cautiously if it is to be used to delve into this topic.

Key words: COVID-19, referee report, COVID-19 mortality rate, positivity rate, robustness.

TABLA DE CONTENIDO

Introducción.....	9
Desarrollo del Tema Reporte De Réferi	10
Resumen	10
Comentarios Principales	10
Comentarios Secundarios	12
Conclusiones.....	16
Referencias Bibliográficas.....	17

ÍNDICE DE TABLAS

Tabla 1 Regresión comparación investigación contra recreación	14
Tabla 2 Regresiones más robustas	15

INTRODUCCIÓN

A pesar de que existe la percepción de que el desarrollo económico y el desempeño de los estados para manejar la pandemia por COVID-19 están relacionados, los datos no arrojan una respuesta clara. La actual pandemia causada por el COVID-19 ha impactado en diferentes niveles a los países alrededor del mundo. Sin embargo, no existe un gran número de investigaciones que intenten analizar los factores que determinen esta variación a nivel país. Investigar acerca de este tema, puede ayudar a mejorar la política pública en el Ecuador y en el mundo. En la investigación realizada por Liang, Tseng, Ho & Wu encuentran que el número de pruebas realizadas y la efectividad del gobierno tienen una relación negativa con la tasa de mortalidad. A pesar de que se puede observar cierta validez en las conclusiones de la investigación, también se pueden realizar ciertas observaciones. Por lo tanto a través de este reporte de referí se busca hacer una revisión integral de la investigación y realizar comentarios con respecto a posibles mejoras.

Entonces, en este reporte de referí se pretende demostrar que el modelo original tiene espacio para mejoras en la medición de los errores estándar y, por lo tanto, de la calidad de los resultados, a través de la implementación de bootstrap. Para esto, primero se reconstruyó la base de datos con la información existente. Después, se reconstruyó la regresión original y se propusieron regresiones con mejores variables. Finalmente, se aplicó bootstrap a la regresión para observar las diferencias con la regresión original.

DESARROLLO DEL TEMA REPORTE DE RÉFERI

La mortalidad por COVID-19 se encuentra negativamente asociada con el número de pruebas y la efectividad del gobierno

Autores: Li-Lin Liang, Ching-Hung Tseng, Hsiu J. Ho, Chun-Ying Wu

RESUMEN

En esta investigación se estudia que factores, a nivel de país, determinan la mortalidad por COVID-19. Para identificar esto, los autores construyen una base de datos con información de incidencia de Covid-19 además de datos económicos y políticos para 190 países. Con estos datos estiman un modelo lineal multivariado para mostrar que número de pruebas y la efectividad del gobierno se encuentran negativamente relacionadas con la mortalidad. Finalmente, para validar su estudio comparan los valores de mortalidad obtenidos en el estudio con los valores de mortalidad recopilados en bases de datos.

COMENTARIOS PRINCIPALES

En esta sección presento los comentarios que fundamentan mi decisión final respecto al paper.

1. Los investigadores estiman su regresión principal asumiendo que los errores estándar son homocedásticos. Los datos, sin embargo, corresponden a información de países, que son unidades de observación que tienden a estar agrupadas (en este caso la agrupación obvia ocurre a nivel de continente). Por este motivo, es de esperar que los errores estándar estén correlacionados a nivel de continentes, por lo que su estimación debería realizarse a través de clusters.

El problema con esta estrategia es que existen clusters (Oceanía, por ejemplo) con muy pocas observaciones. Por lo tanto, lo ideal sería utilizar Bootstrap para calcular los errores estándar. Si esta estructura no es tomada en cuenta (como es el caso del paper que estoy discutiendo), entonces los errores estándar estarán sub-estimados, exagerando así la significancia de los resultados.

Esto lo podemos observar en la Tabla 1. La regresión (1) tiene los valores de la investigación original, la regresión (2) contiene valores recreados con información actual y la regresión (3) corrige los errores estándar estimados en la regresión (2) usando bootstrap. Al comparar la regresión (2) y (3) podemos observar como la significancia disminuye, hasta desaparecer por completo para la variable *número de pruebas por cada 100 personas*. Aparte de esto los errores estándar aumentan para todas las variables.

2. Existe falta de robustez en los resultados. Al construir regresiones basadas en controles más apropiados como las que se encuentran en la Tabla 2, solo se mantuvo la significancia para la *razón de positividad*, número de personas mayores a 65 años, razón de casos críticos y número de camas de hospital. Sin embargo, esta última solo tiene significancia en la regresión (7). Esta regresión mide la mortalidad por COVID-19 rezagada 15 días antes del pico de contagios. Además, solo la regresión (6) tiene valores significantes al aplicar bootstrap.
3. Las variables *número de pruebas por cada 100 personas* y *casos totales por cada 1000 personas* están relacionadas. Esto se debe a que, si existe mayor número de pruebas realizadas, mayor número de casos serán reportados. Por lo que, para corregir este problema es recomendable usar la *razón de positividad* de las pruebas realizadas.

4. Dos de las variables utilizadas en la regresión podrían estar arrojando resultados diferentes a las interpretaciones reportadas en la investigación, debido a que pueden estar siendo afectadas por un sesgo de atenuación, causado por errores de medición en las variables independientes, lo que implica que la pendiente es menor:

Primero, la variable enfermedades comunicables mide el porcentaje de muertes por enfermedades transmisibles, condiciones maternas, prenatales y mala nutrición. Dentro de la investigación es usada para medir patrones de contagio. Sin embargo, este porcentaje está construido por la mayor cantidad de información que pueda obtener. Por lo que, si en un país no existen reportes en enfermedades transmisibles, pero si tiene información en las otras causas de muerte no se observara directamente los patrones de contagio.

Segundo, la variable *infraestructura de transporte*, de la misma forma, posee un error de implementación dentro del modelo. Esta variable mide carreteras, aeropuertos, capacidad de importación, entre otras cosas. En el modelo está siendo usada para medir movilidad dentro de un país, cuando representa más bien una medición de conectividad entre países. Esto se debe a que los datos muestran que tan fácil es comerciar productos entre países debido a la calidad de infraestructura de transporte.

COMENTARIOS SECUNDARIOS

En esta sección explico ciertas observaciones adicionales.

1. Las fuentes del sitio web *World Odometer* son muy buenas pero la información es remplazada con datos más actuales. En otras palabras, la información utilizada por los investigadores ya no está disponible, porque fue substituida por

información actualizada. Por ejemplo, los datos acerca de las Islas Marshall que fueron utilizadas durante la elaboración de la investigación original, ya no existen en la base actual. Sería recomendable usar bases de datos con las que se puedan recrear los resultados, como las del sitio web *Our world in data*.

2. En la investigación existen dos tipos de variables: las que son relacionadas al COVID-19 y las que tienen datos acerca de los países. Las explicaciones acerca de los datos usados para COVID-19 se encuentran bien realizadas. Sin embargo, las explicaciones sobre los datos relacionados a los países poseen ambigüedades. Esto dificulta comprobar la validez de los resultados debido a que no es posible recrear con facilidad la base utilizada.
3. Por todas estas razones la investigación debería ser retirada de la plataforma para analizar si debiera mantenerse en publicación o no. Debido a que tiene problemas con los resultados actuales, lo que dificulta el estudio de la COVID-19. Además de que los investigadores pueden beneficiarse de un análisis más profundo de su modelo.

Tabla 1: Regresión comparación investigación contra recreación

Número de regresión	(1)	(2)
VARIABLES	Mortalidad por cada 100 personas	Mortalidad por cada 100 personas con Bootstrap
Casos totales por cada 1000 personas	0.001*** (.0002)	0.001*** (.0004)
Número de pruebas por cada 100 personas	-0.0002** (.00009)	-0.0002 (.0001)
Razón casos críticos sobre casos totales	0.02 (.07)	0.02 (.4)
Efectividad del gobierno	-0.0007** (.0003)	-0.0007** (.0003)
Porcentaje de la población sobre 65 años	0.1*** (.03)	0.1*** (.04)
Número de camas de Hospital	-0.002*** (.0007)	-0.002*** (.0007)
Enfermedades comunicables	0.002 (.01)	0.002 (.008)
Infraestructura de transporte	0.001*** (.0004)	0.001** (.0004)
Constante	-0.03*** (.01)	-0.03** (.01)
Observaciones	125	125
R-cuadrada	0.46	0.46
Número de repeticiones con Bootstrap		1000
Errores estándar en paréntesis		
*** p<0.01, ** p<0.05, * p<0.1		
R.R significa Riesgo Relativo		

Tabla 1 Regresión comparación investigación contra recreación

Número de regresión	Tabla 2: Regresiones más robustas							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
VARIABLES	Mortalidad 10 días del caso 100	Mortalidad 10 días del caso 100 con Bootstrap	Mortalidad 10 días del caso 100 rezagada 15 días	Mortalidad 10 días del caso 100 rezagada 15 días con Bootstrap	Mortalidad en el tope de casos nuevos	Mortalidad en el tope de caso nuevos con Bootstrap	Mortalidad en el pico rezagada 15 días	Mortalidad en el pico rezagada 15 días con Bootstrap
Razón de positividad	0.001* (.0006)	0.001 (.0008)						
Razón de positividad rezagada 15 días			0.01*** (.004)	0.01 (.009)				
Razón de positividad en el tope de casos nuevos					0.001*** (.0003)	0.001*** (.0003)		
Razón de positividad en el tope de casos nuevos rezagado 15 días							-0.0005 (.001)	-0.0005 (.002)
Razón casos críticos sobre casos totales	-0.03 (.02)	-0.03 (.02)	-0.1* (.05)	-0.1 (.1)	-0.02 (.02)	-0.02 (.02)	0.04* (.02)	0.04 (.04)
Efectividad del gobierno	0.00001 (.00001)	0.00001 (.00001)	0.00001 (.00002)	0.00001 (.00003)	-0.00001 (.00001)	-0.00001 (.00001)	-0.00001 (.00001)	-0.00001 (.00001)
Porcentaje de la población sobre 65 años	-0.0006 (.0009)	-0.0006 (.001)	-0.001 (.002)	-0.001 (.004)	0.002** (.001)	0.002** (.001)	0.001 (.0009)	0.001 (.001)
Número de camas de Hospital	0.00001 (.00002)	0.00001 (.00002)	0.00002 (.00003)	0.00002 (.00006)	-0.00001 (.00002)	-0.00001 (.00001)	-0.00002* (.00001)	-0.00002 (.00003)
Enfermedades comunicables	-0.00006 (.0003)	-0.00006 (.0001)	-0.0004 (.0006)	-0.0004 (.0008)	-0.0001 (.0003)	-0.0001 (.0002)	-0.00005 (.0002)	-0.00005 (.0003)
Infraestructura de transporte	-0.00001 (.00001)	-0.00001 (.00001)	-0.00001 (.00003)	-0.00001 (.00004)	0.00001 (.00001)	0.00001 (.00002)	0.000007 (.00001)	0.000007 (.00002)
Rigor gubernamental	0.0002 (.0002)	0.0002 (.0002)						
Rigor gubernamental rezagada 15 días			-0.0002 (.0003)	-0.0002 (.0004)				
Rigor gubernamental en el tope de casos nuevos					0.0001 (.0001)	0.0001 (.0001)		
Rigor gubernamental en el tope de casos nuevos rezagado 15 días							-0.000002 (.000001)	-0.000002 (.000003)
Constante	0.0002 (.0004)	0.0002 (.0003)	0.0006 (.001)	0.0006 (.001)	-0.0006 (.0004)	-0.0006 (.0007)	-0.00002 (.0004)	-0.00002 (.0007)
Observaciones	61	61	23	23	70	70	22	22
R-cuadrada	0.136	0.136	0.563	0.563	0.304	0.304	0.415	0.415
Número de repeticiones con Bootstrap		1000		1000		1000		1000
Errores estándar en paréntesis								
*** p<0.01, ** p<0.05, * p<0.1								

Tabla 2 Regresiones más robustas

CONCLUSIONES

En este reporte se pretendió revisar la investigación de Liang, Tseng, Ho & Wu en la que se concluyó que el número de pruebas realizadas y la efectividad del gobierno poseen una relación negativa con la mortalidad. Cuando se comparó las respuestas del modelo antes de realizar bootstrap se encontró que los valores reportados para el modelo original cambian. Además, se encontró fue que la medición de los errores estándar mejoró al usar bootstrap. Por otra parte, cuando se construyó regresiones con variables más detalladas la robustez del modelo se perdió, esto ocurre al estandarizar la variable dependiente al medir tasa de mortalidad después del día 100 del primer caso reportado y en el número máximo de casos. Aparte de esto se cambió la forma en que se mide número de pruebas y casos totales, para evitar que exista relación entre estas variables. Finalmente, se observó que existen errores de medición en dos variables dentro de la regresión. Como último comentario, recomiendo utilizar con cautela, porqué existe un gran número de errores en el modelo que evita observar claramente lo que se pretende medir.

REFERENCIAS BIBLIOGRÁFICAS

- Liang, L., Tseng, H., Ho., H & Wu, C. (2020). Covid-19 mortality is negatively associated with test number and government effectiveness. *Scientific Reports*, 10(12567). DOI: 10.1038/s41598-020-68862-x
- Our World in Data. (2020). COVID Data Base. Actualizado a la Fecha. Recuperada el 14 de Octubre de 2020, de: <https://ourworldindata.org/search?q=covid>
- Pischke, S. (2007). Lecture Notes on Measurement Error. Obtenido el 21 de diciembre de 2020 de http://econ.lse.ac.uk/staff/spischke/ec524/Merr_new.pdf
- Stas, K. (2013). Is bootstrapping standard errors and confidence intervals appropriate in regressions where homoscedasticity assumption is violated?. *StackExchange*. Recuperado el 3 de Noviembre de 2020 de: <https://stats.stackexchange.com/questions/56870/is-bootstrapping-standard-errors-and-confidence-intervals-appropriate-in-regress>.
- Stata Corp.(S.f). *bootstrap*. Obtenido el 16 de noviembre de 2020 de <https://www.stata.com/manuals13/rbootstrap.pdf>
- Wooldridge, J. (Eds). (2010). Breve introducción al Bootstrapping. *Wooldridge Introducción a la Econometria Un Enfoque Moderno*. (245-246). Michigan: Cengage Learning.
- Wooldridge, J. (Eds). (2010). Propiedades de MCO bajo error de medición. *Wooldridge Introducción a la Econometria Un Enfoque Moderno*. (315-322). Michigan: Cengage Learning.
- World Bank. (2020). *World Bank Data*. Actualizado a la Fecha. Recuperada el 10 de Octubre de 2020, de: <https://www.worldbank.org>
- World Odometer. (2020). *Coronavirus Updates*. Actualizado a la Fecha. Recuperada el 10 de Octubre de 2020, de: <https://www.worldometers.info>