



Efecto de factores socio-económicos y condiciones de salud en el contagio de COVID-19 en los estados de México

Effects of socio-economic factors and health conditions on COVID-19 contagion in the states of Mexico

Pablo Mejía Reyes*, Annel Hurtado Jaramillo, Liliana Rendón Rojas

Centro de Investigación en Ciencias Económicas. Facultad de Economía. Universidad Autónoma del Estado de México, México

*Recibido el 28 de octubre de 2020; aceptado el 18 de octubre de 2020
Disponible en Internet el: 28 de octubre de 2020*

Resumen

El objetivo de este documento es analizar el impacto de múltiples factores demográficos, sociales, de salud y económicos en la magnitud e intensidad del contagio de SARS-CoV-2 en los estados mexicanos. Para ello se desarrolla un análisis de límites extremos (extreme-bounds analysis) en modelos de regresión de corte transversal, que pueden incluir efectos espaciales. Los resultados sugieren que una mayor densidad de población (que dificulta el distanciamiento social), el padecimiento de obesidad y/o enfermedades crónico-degenerativas (diabetes e hipertensión) y el no respeto a las disposiciones sanitarias han favorecido el contagio de COVID-19. Las condiciones sociales de la población y las características económicas de los estados no resultaron relevantes. Las implicaciones de política pública que se derivan de este resultado son directas.

Código JEL: C21, I18, J11, K00, R15

Palabras clave: COVID-19; Densidad de población; Condiciones de salud; Regulación sanitaria; Econometría espacial

Abstract

The aim of this paper is to analyze the effect of a large battery of demographic, social, health and economic factors on the magnitude and intensity of SARS-CoV-2 contagion in the Mexican states. To do so, an extreme-bounds analysis in cross-section econometric models, with possible spatial dependence, is carried out. Our findings suggest that a greater population density (that impedes social distancing), the suffering of obesity and/or chronic degenerative diseases (diabetes and hypertension) and the lack of respect for health regulations have favored the spread of COVID-19. Social conditions of population and economic characteristics seem to be not relevant. The public policy implications from our results are straightforward.

JEL code: C21, I18, J11, K00, R15

Keywords: COVID-19; Population density; Health conditions; Health regulations; Spatial econometrics

* Autor para correspondencia

Correo electrónico pmejiare@uaemex.mx (P. Mejía Reyes).

La revisión por pares es responsabilidad de la Universidad Nacional Autónoma de México.

<http://dx.doi.org/10.22201/fca.24488410e.2020.3127>

0186- 1042/© 2019 Universidad Nacional Autónoma de México, Facultad de Contaduría y Administración. Este es un artículo Open Access bajo la licencia CC BY-NC-SA (<https://creativecommons.org/licenses/by-nc-sa/4.0/>)

Introducción

La enfermedad denominada COVID-19 (del inglés, *Coronavirus Disease-2019*) es causada por el nuevo Coronavirus del Síndrome Respiratorio Agudo Severo Tipo 2 (SARS-CoV-2), el cual se ubica taxonómicamente en la familia *Coronaviridae* (Gorbalenya *et al.*, 2020). Esta familia está compuesta por diferentes virus que causan enfermedades en humanos y animales que van desde el resfriado común hasta enfermedades respiratorias graves. El COVID-19 es una enfermedad transmisible que se extiende a gran velocidad (Gupta *et al.*, 2020) pues se estima que una persona infectada propagaría la enfermedad entre 2.2 y 3.5 personas, en promedio (Díaz-Castrillón y Toro-Montoya, 2020; Callaway *et al.*, 2020).

Los primeros casos de COVID-19 se confirmaron en diciembre de 2019 en la ciudad de Wuhan, provincia de Hubei de China, como un brote de neumonía de causa desconocida.¹ El virus se propagó rápidamente hacia el resto de China y otros países del mundo. A mediados de enero se identificaron pacientes en otros países de Asia (Indonesia, Japón Tailandia y Corea del Sur) y desde la última semana de enero en varios de Europa (Francia, Alemania, España, Reino Unido e Italia). Al igual que en los casos anteriores, el virus llegó a América con viajeros internacionales. Al parecer, el primer caso se confirmó en Estados Unidos (EE.UU.) el 19 de enero de 2020 en el estado de Washington (Holshue *et al.*, 2020). Ante la acelerada dispersión del virus en varios países de prácticamente todos los continentes del mundo, la Organización Mundial de la Salud (OMS) declaró el 30 de enero al brote de SARS-CoV-2 como una emergencia de salud de importancia mundial (OMS, 2020a; ECDC, 2020). A su vez, el nuevo coronavirus apareció en América Latina casi un mes después: el 26 de febrero se confirmó el primer caso en Sao Paulo, Brasil, mientras que en México apareció el 28 de febrero en la Ciudad de México (PAHO/WHO, 2020). En ambos casos se trató de personas que habían estado recientemente en Italia. Ante este panorama, la OMS declaró oficialmente la existencia de una pandemia de COVID-19 el 11 de marzo de 2020 (OMS, 2020a).

Dado que el COVID-19 es una nueva enfermedad no existe vacuna ni tratamiento comprobado para su cura o control efectivo. Por ello, la inmensa mayoría de países del mundo han tomado como medidas generales de prevención el distanciamiento físico y la práctica de una higiene profunda (Lakshmi-Priyadarsini y Suresh, 2020), lo que ha requerido que las personas reduzcan sus actividades al aire libre, la concentración de personas y los desplazamientos geográficos. Como consecuencia, en prácticamente todo el mundo se han suspendido actividades escolares y se han reducido o cancelado significativamente actividades turísticas, recreativas y de entretenimiento, entre otras. En diferente grado y con efectividad variante, se ha impuesto el confinamiento domiciliario y el uso de cubre-bocas.

No obstante, la virulencia del COVID-19, el incumplimiento de las medidas sanitarias y otros factores sociales, sanitarios y demográficos han llevado a elevados niveles de contagio en el mundo, especialmente cuando no se han tomado las medidas adecuadas. Así, por ejemplo, hasta el 31 de julio, el número de contagios alcanzó 4 388 566 en EE.UU., 3 300 000 en los países europeos y 770 412 en África. Sólo en países donde se tomaron medidas estrictas de distanciamiento social los contagios se controlaron de manera relativamente rápida, como en China, donde hasta la misma fecha se habían encontrado 87 956 casos (OMS, 2020b).

En México, a 178 días de que se identificara el primer caso, se han tenido 454 322 personas contagiadas y 46 688 muertes. La información desagregada, por su lado, muestra que los contagios se han distribuido de manera muy heterogénea entre los estados del país. En particular, hasta el 31 de julio, los tres estados que han tenido un mayor número de contagios han sido la Ciudad de México, el Estado de México y Tabasco con 76 804, 55 726 y 22 961, respectivamente. Por el contrario, los estados con el menor número de casos han sido Colima (1 985), Zacatecas (2 905) y

¹ La identificación de los primeros pacientes en China y en otros países fue muy incierta, dado que se trataba de un virus desconocido que solamente fue identificado hasta el 12 de enero en China. Véase OMS (2020a) para una secuencia de los eventos desde que se presentó el primer caso.

Nayarit (3 306). Desde luego que, entre otros factores, esto puede deberse al tamaño de la población de los mismos estados (efecto de escala), por lo que una medida más apropiada de su intensidad es la tasa de prevalencia. De acuerdo a ésta, los estados más afectados hasta el 31 de julio son Tabasco, que presentó 8.9 casos por cada 1 000 habitantes, Ciudad de México, 8.5, y Sonora 6.0. En contraste, Chiapas solamente tuvo 1.0, Chihuahua 1.5 y Jalisco 1.6 (Gobierno de México, 2020).

Sin duda el mundo se enfrenta a una crisis de salud sin precedentes causada por el nuevo virus SARS-CoV-2. Por ser una variante diferente, mucho se ha especulado sobre los principales factores que propician su contagio. No obstante, relativamente poco se ha investigado sobre la relevancia que diversos factores pueden haber tenido en su propagación, a lo que también ha contribuido que se trata de un proceso en curso. Por ello, el objetivo de este documento es analizar el impacto de factores sociales, demográficos y económicos y condiciones prevalecientes de salud en los niveles acumulados de contagio y en la tasa de prevalencia del COVID-19 en los estados mexicanos. Para ello, se realiza un análisis de límites extremos en diferentes modelos econométricos con el fin de identificar dentro de una amplia gama a las variables que pueden explicarlos. Los resultados indican que, en general, la densidad de población, el acatamiento de las recomendaciones sanitarias y las enfermedades preexistentes condicionan su magnitud e intensidad.

El resto del documento consta de cinco secciones. En la primera se presenta brevemente la evolución de los contagios de SARS-CoV-2 a nivel internacional hasta su llegada a México para así analizar los contagios a nivel estatal. En la segunda sección se revisa la literatura relevante sobre el tema. En la tercera se especifica el modelo econométrico y se define metodología que se usan para identificar los factores que pueden explicar el nivel acumulado y la tasa de prevalencia del COVID-19. En la sección cuatro se presentan y discuten los principales resultados y, por último, se establecen las conclusiones.

Contagio de COVID-19 en México y en el mundo

A finales de 2019 se inicia la peor crisis sanitaria que ha experimentado el mundo en los últimos casi cien años derivada del contagio del novel virus denominado SARS-CoV-2. De acuerdo con el *South China Morning Post*, el primer enfermó se presentó en China el 17 de noviembre de 2019 y a partir de entonces los casos aumentaron rápidamente a lo largo y ancho de ese país. El 27 de diciembre, Zhang Jixian, médico del Hospital Provincial de Medicina China y Occidental Integrada de Hubei, informó que la enfermedad había sido causada por un nuevo coronavirus (Ma, 2020).

Según la Organización Mundial de la Salud (OMS, 2020a), el día 9 de enero del 2020 se presentó la primera muerte a causa del virus y el 13 del mismo mes se registró un caso en Tailandia, el primero fuera de China. En América, EE.UU. reportó el primer paciente el 19 de enero. Francia, por su lado, informó sobre sus dos primeros enfermos el 24 de enero, con lo que se iniciaba la epidemia en Europa. Un día después, el 25 de enero, Australia da a conocer que tiene 4 casos en el país, con lo que la enfermedad llega a Oceanía (Reuters, 2020).

El 11 de febrero la OMS (2020a) anuncia que el nombre oficial de la enfermedad sería “Enfermedad por Coronavirus”, abreviada como “COVID-19”, por sus siglas en inglés. En los siguientes días se confirman los primeros casos en Egipto (primero de África), seguido por Israel, y después en Brasil (el 26 de febrero), convirtiéndose así este último en el primer país de Latinoamérica con portadores del nuevo virus.

Para detener los contagios, los gobiernos del mundo han adoptado una amplia gama de medidas. Ante la ausencia de procedimientos médicos probados o de una vacuna para eliminar el virus, se optó por imponer restricciones a la movilidad de las personas, tanto en su entorno inmediato (confinamiento domiciliario) como en su desplazamiento geográfico (viajes nacionales e internacionales). Así, por ejemplo, el 22 de enero Corea del Sur anunció la suspensión temporal de la entrada de turistas en todo su territorio; el gobierno italiano, por su parte, determinó el aislamiento total del

país, además de imponer medidas más estrictas, como la prohibición de viajes.² España decretó el Estado de Alarma, por lo que se confinó a la población, en tanto que los trabajadores no imprescindibles dejaron de laborar o lo hicieron desde sus hogares, además de que los colegios, institutos y universidades cerraron (Carrión *et al.*, 2020). En América, Perú y Chile pasan a estado de emergencia y cierran sus fronteras el 15 de marzo (OMS, 2020a).

A pesar de estas medidas, y otras que se tomaron en los diferentes países, el número de contagios aumentó en todo el mundo, aunque a diferentes ritmos, lo que llevó a la OMS a declarar el COVID-19 como pandemia global el 11 de marzo (OMS, 2020a). El rápido crecimiento de los contagios ha llevado a un elevado número de casos acumulados en varios países. Por ejemplo, hasta el 31 de julio, se ha tenido un total de 4 495 014 en EE.UU.; 288 522 en España; 186 573 en Francia; 247 158 en Italia; 2 610 102 en Brasil, 407 492 en Perú, 353 536 en Chile y 454 322 en México. Al parecer, los países donde hubo un menor número fue donde se adoptaron medidas de control rápidas y estrictas, como China y Corea del Sur, donde se tuvieron hasta la misma fecha un total de 87 489 y 14 305 personas afectadas, respectivamente (ECDC, 2020).

En el caso de México, el primer paciente con COVID-19, importado de Italia, se dio a conocer el 28 de febrero de 2020, lo que llevó al gobierno federal a anunciar el 29 del mismo mes el inicio de la Fase 1 de la enfermedad, con casos importados. La expansión exponencial del virus en las primeras semanas llevó a que las autoridades mexicanas implementaran la “Jornada Nacional de Sana Distancia” a partir del 23 de marzo (y que duraría hasta el 19 de abril). Entre las medidas adoptadas se encontraban el confinamiento voluntario, la suspensión de actividades no esenciales, el desarrollo de actividades educativas y laborales desde casa en la medida de lo posible, la suspensión de eventos masivos (más de 5 000 personas) y el cuidado especial de los adultos mayores de 60 años. El mismo día, la OMS señala que en México ya había transmisión local de la enfermedad, por lo que se transitaba a la Fase 2 de la pandemia. Por ello, el 30 de marzo, con 1 094 casos confirmados y 28 defunciones, el gobierno federal decretó la emergencia sanitaria y se determinó el resguardo domiciliario de personas mayores de 60 años, así como de aquellas con padecimientos de diabetes, hipertensión arterial e insuficiencia cardíaca (Gobierno de México, 2020).

El 20 de abril, después de 857 decesos y 9 501 casos confirmados, se anuncia el inicio de la Fase 3. Entre las medidas aprobadas destacan la extensión de la Jornada Nacional de Sana Distancia hasta el 30 de mayo y la ampliación de la suspensión de actividades no esenciales. Al finalizar la jornada de sana distancia, se reportaron 87 517 casos confirmados y 9 779 defunciones (Gobierno de México, 2020). Entonces, a partir de principios de junio, México, al igual que otros países, se preparó para la reapertura de la economía transitando a la llamada fase de la “Nueva Normalidad”.³ Con la puesta en marcha de un sistema basada en la idea del “semáforo”,⁴ la administración federal dio por concluida la contingencia sanitaria decretada a finales de marzo y transfirió a los gobiernos estatales las atribuciones para que cada territorio decidiera cómo llevar a cabo la reapertura (Gobierno de México, 2020).

² Italia impuso el bloqueo más draconiano fuera de China continental, al restringir la movilidad de más de 10 millones de personas en la parte norte del país. Los lugares afectados por el cierre incluyen el centro financiero de Milán y el destino turístico de Venecia, se impusieron medidas como el cierre de escuelas y universidades y la cancelación de todos los eventos públicos, eventos deportivos y ceremonias civiles y religiosas (Carteni *et al.*, 2020). Los italianos permanecieron en cuarentena hasta el 3 de mayo.

³ Por entonces, la economía mexicana se encontraba sumida en la recesión más profunda de las últimas nueve décadas. De acuerdo a cifras del Instituto Nacional de Estadística y Geografía, INEGI (2020), la producción había disminuido en 17.3 a tasa anual, la inversión había caído 31.7% y se habían perdido 12.8 millones empleos formales e informales.

⁴ La idea del Semáforo de riesgo epidemiológico para transitar hacia una nueva normalidad es un sistema de monitoreo para la regulación del uso del espacio público de acuerdo con el riesgo de contagio de COVID-19. Este semáforo es estatal y consta de cuatro colores: Rojo, se permiten únicamente actividades económicas esenciales. Naranja, además se permite que las empresas de actividades económicas no esenciales trabajen con el 30% del personal, siempre tomando en cuenta las medidas de cuidado máximo para las personas con mayor riesgo de presentar un cuadro grave de COVID-19, se abren los espacios públicos abiertos con un aforo (cantidad de personas) reducido. Amarillo, todas las actividades laborales están permitidas, cuidando a las personas con mayor riesgo. El espacio público abierto se abre de forma regular, y los espacios públicos cerrados se abren con aforo reducido. Verde, se permiten todas las actividades, incluidas las escolares (Gobierno de México, 2020).

La estrategia del gobierno mexicano no ha sido exitosa desde el punto de vista del número de contagios diarios, que siguieron aumentando a ritmos crecientes hasta los últimos días de julio; sólo recientemente se ha presentado un estancamiento por encima de 50 casos diarios por millón de habitantes, cifra muy elevada en comparación con las que tienen otros países por las mismas fechas. Más aun, al 31 de julio, México tenía 454 322 personas contagiadas en total, lo que correspondía a una tasa de prevalencia⁵ de 3 800.9 personas por millón de habitantes (Gobierno de México, 2020).

A nivel de estados la situación ha sido muy heterogénea debido a sus propias características socioeconómicas y a las reacciones de la población ante las medidas sanitarias, así como a las políticas específicas que cada estado ha adoptado.⁶ En particular, la figura 1 muestra los contagios acumulados de los estados que presentan los niveles más altos. Se puede observar un crecimiento exponencial en la mayoría de ellos al menos durante los primeros 100 días; de hecho, en casos como Tabasco, Veracruz, Guanajuato, Puebla y Nuevo León se mantiene la misma tendencia. A su vez, la Ciudad de México y el Estado de México han tenido un crecimiento lineal en las últimas semanas, mientras que en Sonora y Baja California, y un poco en Tamaulipas, el ritmo de crecimiento parece empezar a perder fuerza. En cualquier caso, los niveles de contagio han sido muy elevados.

La figura 2, a su vez, presenta la evolución de contagios por cada 1 000 habitantes en los estados donde se han alcanzado los mayores niveles; nótese que éstos son diferentes a los que aparecen en la figura 1, aunque algunos se repiten. También en este caso se presentan diferentes patrones. La tendencia hasta finales de julio sigue siendo creciente en Coahuila, Yucatán, Tabasco y Baja California Sur, mientras que en la Ciudad de México, Quintana Roo y Campeche parece haber un crecimiento lineal estable. Solamente en Tamaulipas, Sonora y Sinaloa se aprecia una desaceleración del ritmo de contagios.

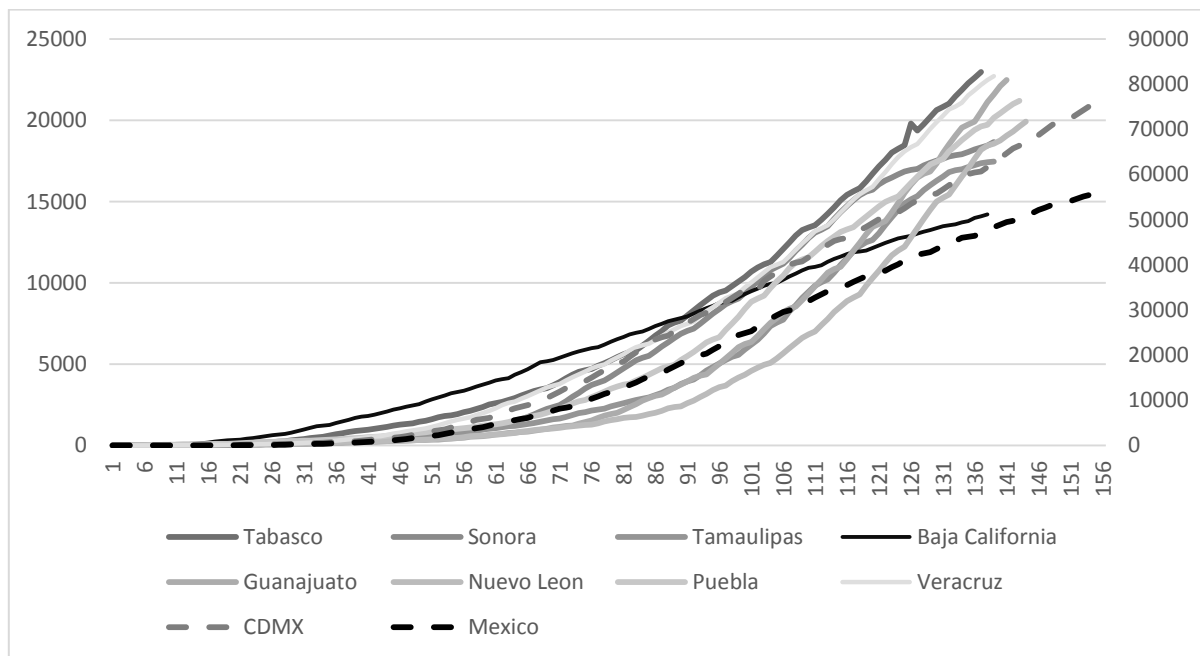


Figura 1. Contagios acumulados de COVID-19 en los estados de México con mayores niveles (Número de personas). Los datos de la Ciudad de México y el Estado de México se miden en el eje de la derecha.
 Fuente: elaboración propia con base en información de <https://coronavirus.gob.mx/datos/#COMNac>.

⁵ La tasa de prevalencia pondera los contagios por un cierto número de habitantes, por lo que puede verse como una medida de intensidad de la pandemia.
⁶ Los estados adoptaron también medidas muy heterogéneas. Un resumen está disponible a solicitud expresa.

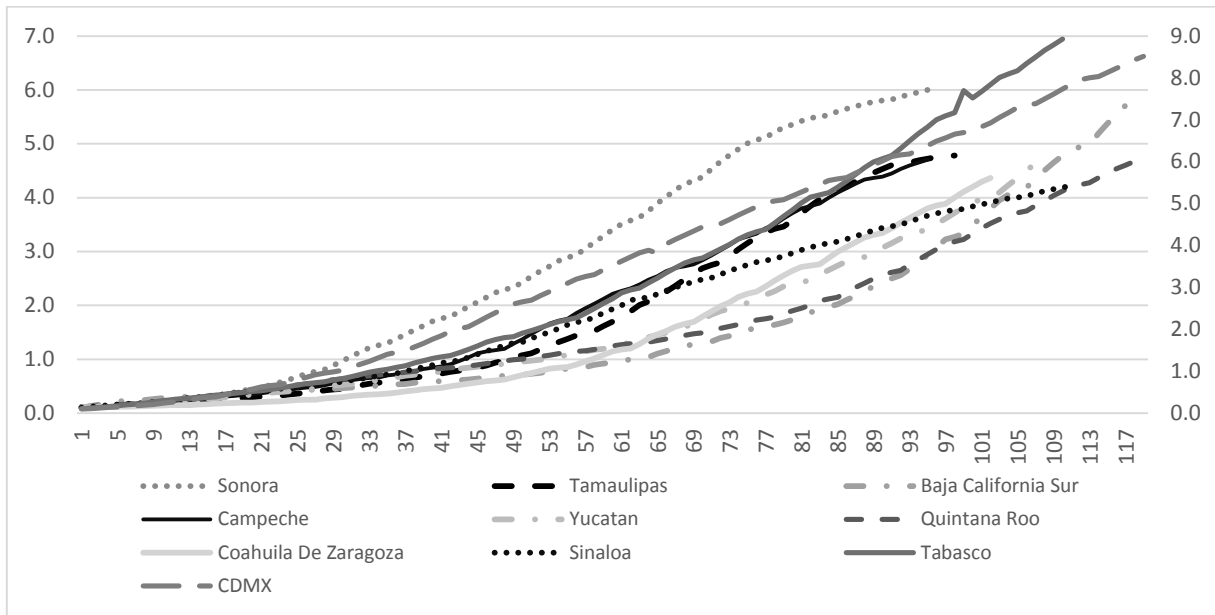


Figura 2. Tasa de prevalencia de COVID-19 en los estados con mayores niveles (Contagios por cada 1 000 habitantes). En el eje horizontal se mide el número de días desde que presentó un promedio 0.1 casos por 1 000 habitantes. Los casos de Ciudad de México y Tabasco se miden en el eje derecho.
 Fuente: elaboración propia con base en información de <https://coronavirus.gob.mx/datos/#COMNac>

La distribución espacial del número acumulado de contagios entre los estados del país aparece en el mapa de la figura 3. Se aprecia que hasta el 31 de julio de 2020, los estados con el mayor número de contagios eran la Ciudad de México (76 804), seguida por el Estado de México (55 726), Tabasco (22 961), Veracruz (22 703), Guanajuato (22 479), Puebla (21 192), Nuevo León (19 910) y Sonora (18 657), varios de ellos con niveles de población elevados. Por el contrario, estados caracterizados por menores volúmenes de población, grandes extensiones territoriales o ambientes predominantemente rurales han tenido un menor número de contagios, como Chihuahua (5 913), Durango (4 365), Zacatecas (2 905), Baja California Sur (4 692), Chiapas (5 481) o Campeche (4 706).

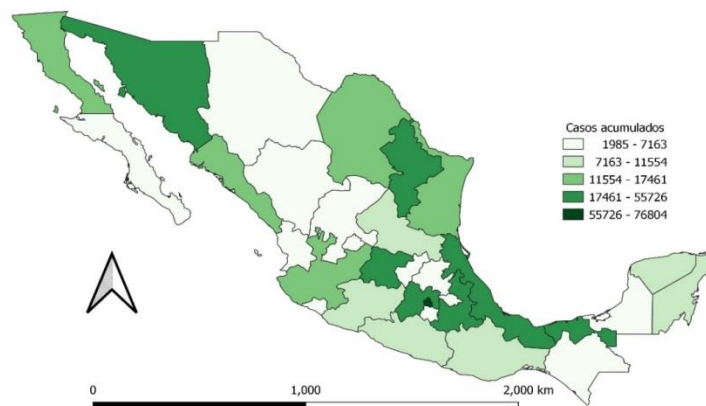


Figura 3. Contagios acumulados de COVID-19 en los estados de México (Personas).
 Fuente: elaboración propia con base en información de <https://coronavirus.gob.mx/datos/#COMNac>.

En el mapa de la figura 4, por otro lado, se presenta la tasa de prevalencia del COVID-19 (número de contagios por cada 1 000 habitantes) en los estados del país. Aunque hay diferencias con respecto a la figura 3, algunos patrones se repiten. En particular, Tabasco (8.9) y Ciudad de México (8.5) presentan los niveles más altos, seguidos de Sonora (6.4), Baja California Sur (5.8), Tamaulipas (4.7), Campeche (4.7), Quintana Roo (4.6) y Yucatán (4.6). En el otro extremo se ubican los estados de Chihuahua (1.5), Zacatecas (1.7), Jalisco (1.6), Querétaro (1.6) y Chiapas (1.0) con el menor número de contagio por cada 1 000 habitantes.

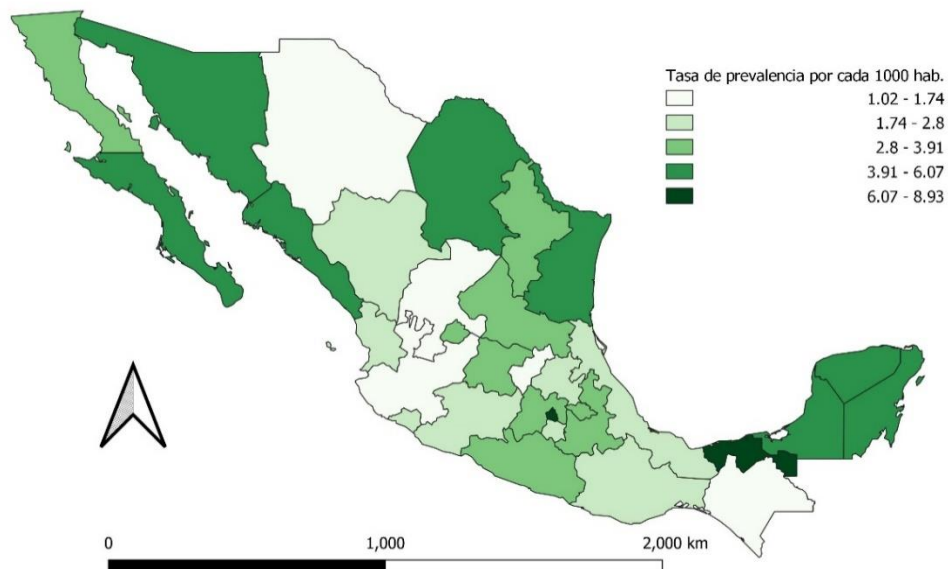


Figura 4. Tasa de prevalencia de COVID-19 en los estados de México (Contagios por cada 1 000 habitantes).
Fuente: elaboración propia con base en información de <https://coronavirus.gob.mx/datos/#COMNac>.

Como se puede ver en los mapas de las figuras 3 y 4, los patrones espaciales de la magnitud y tasa de contagios de COVID-19 cambian de manera importante. Como era de esperarse, los primeros se presentan en mayor medida en los estados con mayor población simplemente por un factor de escala, aunque no siempre. El panorama cambia cuando esos contagios se miden como proporción de la población estatal, lo que da una medida de la intensidad del contagio. En cualquier caso, lo que nos interesa dilucidar en las siguientes secciones es cuáles factores pueden explicar estas diferencias en magnitud e intensidad del COVID-19 entre los estados de México, para lo cual se utiliza un modelo econométrico.

Breve revisión de literatura

El COVID-19 es una enfermedad muy virulenta que sorprendió a los científicos y a los responsables de las políticas públicas de los diferentes países del mundo, especialmente durante los primeros meses. Poco a poco, sin embargo, se ha ido aprendiendo más sobre su naturaleza y las medidas que deben tomarse para controlar o reducir el ritmo de su contagio. En ese sentido, aun cuando desde un punto de vista clínico son evidentes las medidas a seguir, resulta menos claro el

diagnóstico con respecto a los factores que pueden explicar la magnitud e intensidad del contagio entre países y entre regiones del mismo país. Ante la gravedad de la situación, y todavía con el fenómeno en curso, han aparecido diferentes estudios que tratan de explicarlo para regiones de China (Li *et al.*, 2020; Xie *et al.*, 2020), EE.UU. (Papageorge *et al.*, 2020; Gerritse, 2020; Mollalo *et al.*, 2020; Yanga *et al.*, 2020) o Europa (Onder *et al.*, 2020; Carteni *et al.*, 2020), entre otros.

En primer lugar, Lakshmi-Priyadarsini y Suresh (2020) emplean la metodología TISM (Total Interpretive Structural Modelling)⁷ para identificar los factores que influyen en las características epidemiológicas del COVID-19 en todo el mundo. Entre sus principales resultados están que el distanciamiento social y la conciencia comunitaria, la edad, la temperatura del aire, el flujo de aire y la ventilación, la densidad de población y la humedad constituyen factores clave o de conducción en este modelo. Además, encuentran que el comportamiento del huésped y la cantidad de contactos y prácticas de higiene personal son los factores de enlace.

A nivel de países, Xie *et al.* (2020) realizan un análisis exploratorio de datos espaciales para analizar las diferencias espaciales y temporales de la dispersión de la pandemia en China. Además, con el método geo-detector, que permite analizar la heterogeneidad estratificada espacial, identifican los factores que influyen en la propagación del COVID-19. Encuentran que la afluencia de población de Wuhan y la intensidad de la conexión económica regional fueron los principales factores que influyeron en la propagación de la epidemia, junto con la distribución de la población, la accesibilidad al transporte, la temperatura promedio y las instalaciones médicas. En tanto, Yanga *et al.* (2020) evalúan la prevalencia de comorbilidades en los pacientes infectados con SARS-CoV-2 para lo cual realizan un meta-análisis de siete estudios realizados en hospitales de China. Utilizan razones de probabilidad, modelos de efectos aleatorios e intervalos de confianza (IC) del 95% para llegar a sus conclusiones, en las que destacan que los padecimientos subyacentes, incluidas enfermedades cardiovasculares y del sistema respiratorio e hipertensión, pueden ser factores de riesgo para pacientes graves en comparación con pacientes no graves.

Carteni *et al.* (2020), por su parte, analizan el efecto de los hábitos de movilidad en la propagación del coronavirus en Italia a través de un modelo de regresión lineal múltiple. La estimación vincula el número de casos diarios con variables socioeconómicas, ambientales, de atención médica y hábitos de movilidad. Los resultados de la estimación mostraron que los hábitos de movilidad constituyen la variable principal que explica el número de infecciones por COVID-19. También fueron significativas las variables ambientales, el número de pruebas por día y la proximidad al primer brote italiano, especialmente en la etapa inicial de la infección. Onder *et al.* (2020), por su parte, buscan explicar las elevadas tasas de mortalidad por COVID-19 en las regiones italianas en comparación con las de otros países. Observan que había una alta proporción de pacientes mayores con infección confirmada por COVID-19, por lo que este factor podría explicar en parte las diferencias, y que las muertes se presentaban principalmente entre pacientes masculinos mayores con multi-morbilidad.

Para el caso de España, Carrión *et al.* (2020) estudian el origen de las diferencias espaciales en la propagación del COVID en las Áreas Básicas de Salud (ABS) de Cataluña, empleando diferentes modelos de regresión múltiple y regresiones cuantílicas. Concluyen que la transmisión del virus se da rápidamente en áreas que concentran volúmenes elevados de población y, aunque no hay relación concluyente, sugieren que el nivel socioeconómico bajo también podría tener un efecto positivo en la propagación del virus.

En el caso de EE.UU., Mollalo *et al.* (2020) utilizan cinco diferentes modelos –mínimos cuadrados ordinarios (OLS), modelo de rezago espacial (SLM), modelo de error espacial (SEM), regresión ponderada geográficamente (GWR) y GWR multiescala (MGWR)– para explicar las variaciones de la tasa de incidencia del COVID-19 a nivel de condado.

⁷ Esta metodología sirve para analizar la interrelación entre factores que influyen en cierto fenómeno, estableciendo relaciones directas o transitivas entre factores.

Entre las variables explicatorias consideran 35 factores socioeconómicos, ambientales, topográficos, de comportamiento y demográficos. Entre sus resultados destacan que los factores explicativos más importantes se encuentran la desigualdad de ingresos, el ingreso familiar promedio, la proporción de mujeres negras y la proporción de enfermeras practicantes. A su vez, Gerritse (2020) encuentra, mediante la estimación de ecuaciones de regresión epidemiológica, que la transmisión y la incidencia de infecciones por COVID-19 difieren notablemente entre las áreas de EE. UU. y que la densidad de población se asocia con mayores velocidades de transmisión, en particular al inicio de los brotes. Por su lado, Dietz y Santos-Burgoa (2020) analizan el impacto desproporcionado de la gripe H1N1 y ahora COVID-19 en pacientes con obesidad y obesidad grave. Afirman que ésta se asocia con una disminución del volumen de reserva espiratoria, de la capacidad funcional y de la distensión del sistema respiratorio, lo que agrava las infecciones provocadas por SARS-CoV-2.

Para el caso de México, Padilla-Santamaría *et al.* (2020) realizan un estudio transversal retrospectivo analítico, donde analizan el número oficial de casos confirmados y defunciones por COVID-19 en México hasta el 9 de mayo de 2020. Ellos afirman que los hombres se contagian con mayor frecuencia que las mujeres y que la letalidad también es más alta en el género masculino, además de que tanto las muertes como los contagios están relacionados con la densidad poblacional.

Esta revisión de literatura muestra que existen relativamente pocos análisis, especialmente en el caso de México, que buscan comprender la naturaleza del contagio de COVID-19, posiblemente porque es un fenómeno en curso del que todavía se desconocen muchos aspectos. No obstante, estos estudios constituyen avances significativos para identificar algunos de los factores socioeconómicos principales para explicar su magnitud e intensidad. Con base en los resultados de estos trabajos, en la siguiente sección se formula un modelo econométrico para explicar los contagios en el caso de los estados mexicanos.

Especificación del modelo econométrico

Las variables que se han utilizado en los estudios revisados en la sección anterior para explicar el contagio de SARS-CoV-2 (c) se pueden organizar en diferentes grupos asociados a factores demográficos (X_1), sociales (X_2), condiciones de salud (X_3) y económicos (X_4). Para analizar el efecto de estas variables sobre el número de casos acumulados (magnitud) y la tasa de prevalencia (intensidad) del contagio del SARS-CoV-2 en los estados de México se plantea el siguiente modelo lineal general:

$$c = X_1b_1 + X_2b_2 + X_3b_3 + X_4b_4 + u \quad (1)$$

donde c es la variable a explicar y corresponde a un vector de orden $(n \times 1)$ (con $n = 32$ estados) que contiene el número de casos acumulados de contagio (CA) o la tasa de prevalencia (TC); X_i (para $i = 1, 2, 3, 4$) denota una matriz de orden $(n \times k_i)$ de variables explicatorias, con parámetros asociados contenidos en el vector b_i de orden $(k_i \times 1)$, mientras que u representa el vector de perturbaciones de orden $(n \times 1)$ que sigue un proceso ruido blanco, es decir $u \sim iid N(0, \sigma^2 I)$.

En particular, X_1 contiene factores demográficos, como la densidad de población (DP) y el porcentaje de población masculina (TH).⁸ La primera se asocia con mayores velocidades de transmisión, ya que entre más personas se concentran en áreas más reducidas es mayor la cercanía física, lo que eleva la posibilidad de contagio (Kumar, 2015; Padilla-Santamaría *et al.*, 2020; Hoehl *et al.*, 2020; Gilbert *et al.*, 2020; ECDC, 2020; Carrión *et al.*, 2020). Las estadísticas, por otra parte, han mostrado una mayor incidencia de COVID-19 en hombres que en mujeres (Onder *et al.*,

⁸ Las definiciones y fuentes de las variables se presentan en la tabla A1 del anexo.

2020; Padilla-Santamaría *et al.*, 2020) en una proporción de 52.6% a 47.4% en el caso de México (Gobierno de México, 2020), por lo que se espera que haya más contagios en estados con un porcentaje más alto de hombres. De esta forma, la matriz X_1 estaría integrada por los vectores correspondientes a estas variables, $X_1 = [DP TH]$.⁹

Por otra parte, X_2 recolecta variables que hemos denominado sociales y que se refieren principalmente a la actitud de las personas frente a las recomendaciones sanitarias. De hecho, hay numerosos reportes en los medios nacionales e internacionales que atribuyen el aumento de los contagios a la falta de respeto al distanciamiento social (e. g. Flournoy y Morell, 2020). Un problema con esta variable es que no es directamente observable, por lo que se usan diferentes indicadores *proxy* que tratan de medirla. En primer lugar, se usa el Índice de Cultura de la Legalidad (*ICL*), el cual combina variables que dan elementos para conocer si las personas encuestadas declaran tener respeto, conocimiento y cumplimiento de las leyes y prácticas cívicas (véase MUCD, 2014). En segundo lugar, se utiliza un conjunto de variables que miden en general la percepción o el respeto al estado derecho *ex post*. Aquí se cuentan variables como percepción de la seguridad (*PINS*), percepción de corrupción estatal (*PCORE*), costos del delito (*CDEL*), delitos denunciados (*DEN*) e incidencia delictiva (*INDEL*). La matriz $X_2 = [ICL PINS PCORE CDEL DEN INDEL]$.¹⁰

En la matriz X_3 se consideran variables asociadas a enfermedades preexistentes, ya que hay evidencia que sugiere que el sistema inmunológico de personas con obesidad y enfermedades crónico-degenerativas pueden tener complicaciones respiratorias o cardiovasculares, lo que las haría más propensas a contraer el SARS-CoV-2 (Onder *et al.*, 2020; Gupta *et al.*, 2020; Dietz y Santos-Burgoa, 2020; Yanga *et al.*, 2020). Para captar estos efectos se consideran variables como la proporción de la población total con diabetes (*TDIA*), hipertensión (*THIP*) y obesidad (*TOB*), así como la que presenta multi-morbilidad asociada a hipertensión y obesidad (*THIPOB*) y diabetes, hipertensión y obesidad (*THIPOBDIA*).¹¹

Por último, los contagios de COVID-19 pueden estar sujetos a factores económicos que impidan a las personas confinarse en sus domicilios o realizar trabajo desde casa, como puede ser la situación de trabajadores de sector informal o de aquellos con empleos precarios. Más aun, la población en condiciones de pobreza o marginación puede ser más vulnerable al contagio al no contar con los medios financieros, las condiciones físicas de salud o higiene apropiada para enfrentar el virus (Flournoy y Morell, 2020). En esta categoría se tomaron en cuenta también otras variables, pero solamente la brecha laboral *BL*, definida como $BL = (\text{Desocupados} + \text{Subocupados} + \text{Disponibles}) / (\text{Población económicamente Activa} + \text{Disponibles})$ ¹² resultó significativa en la explicación; sus valores se recogen en la matriz X_4 .¹³

Por otro lado, el tránsito de personas de un lugar a otro, principalmente, puede hacer que el contagio se disperse entre países (como de hecho ocurrió) y regiones del mismo país. Para modelar esa posibilidad, el modelo (1) se extiende para capturar la interacción espacial de los contagios de COVID-19 entre los estados de México. Entonces, agrupando a todas las variables explicatorias y a los parámetros correspondientes en la matriz $X = [X_1 X_2 X_3 X_4]$ y el vector $b = [b_1 b_2 b_3 b_4]$, el modelo (1) sería:

⁹ Adicionalmente se consideraron el número de personas por hogar y el número de habitaciones por persona, así como el índice de movilidad del Observatorio COVID-19 (2020), pero ninguna de ellas tuvo efectos estadísticamente significativos sobre *CA* o *TC*.

¹⁰ Las variables que no resultaron significativas en esta categoría fueron el índice de marginación, el porcentaje de población en situación de pobreza, el índice de rezago social y el coeficiente de Gini.

¹¹ También se tomaron en cuenta medidas alternativas como el número de médicos en instituciones públicas y el índice de salud, basado en la esperanza de vida al nacer (PNUD, 2015). Sin embargo, ninguna tuvo efectos estadísticamente significativos.

¹² La brecha laboral busca reflejar la incidencia del subempleo (personas que trabajan a tiempo parcial y quieren un empleo a tiempo completo) y el alcance del desempleo oculto (personas que no están buscando activamente, pero que se reincorporarían a la fuerza laboral si el mercado lo permitiera) (Blanchflower y Levin, 2015).

¹³ También se consideraron la tasa de informalidad laboral, la tasa de condiciones críticas de ocupación y la tasa de desempleo extendido, pero ninguno presentó efectos significativos.

$$\begin{aligned}c &= \rho Wc + Xb + u \\u &= \lambda Wu + \varepsilon \\ \varepsilon &\sim N(\mathbf{0}, \sigma^2 I)\end{aligned}\tag{2}$$

donde W es la matriz de pesos espaciales estandarizada por renglón, con entradas iguales a 1 si los estados son vecinos y 0 de otro modo (contigüidad tipo reina). Por lo tanto, Wc es el término autorregresivo del modelo, de manera que ρ es el parámetro que mide la dependencia espacial sustantiva entre cada estado y sus vecinos; Wu y λ se definen de manera análoga en relación con la parte no explicada por las variables sugeridas por la teoría; ε es un vector de orden $(n \times 1)$ que sigue un proceso ruido blanco. A partir de este modelo general se pueden definir diferentes alternativas: si $\lambda = 0$ se tiene el modelo de rezago espacial, en tanto que si $\rho = 0$ se obtiene un modelo de error espacial; en el caso de que ambos coeficientes difieran de 0 se tendría el modelo general denominado *SARMA* (*Spatial autorregressive and moving average*) por Anselin y Bera (1998).

La elección del modelo espacial particular a estimar se basa en los estadísticos *LM* (multiplicadores de Lagrange) usados ampliamente. En sus dos versiones, LM_ρ ó LM_λ , evalúan la hipótesis nula de no dependencia espacial sustantiva y residual, respectivamente. Si la hipótesis nula se rechaza con alguno de ellos, se estima el modelo correspondiente, pero si ambos estadísticos de prueba son significativos se opta por la mejor especificación; por ejemplo si $LM_\rho > LM_\lambda$, se estima el modelo de rezago espacial, y viceversa. Adicionalmente, y bajo la misma lógica, se utilizaron las versiones robustas de los mismos estadísticos, los cuales son válidos aun en presencia de errores de especificación. En general, si la hipótesis nula no se rechaza se opta por estimar modelos convencionales de corte transversal.

Factores explicativos del contagio de SARSCoV-2

El análisis de los determinantes de los contagios de COVID-19 de los estados de México se basa en la especificación general del modelo y las técnicas econométricas descritas en la sección anterior. En particular, la especificación de los modelos estimados consta de dos etapas. Primero, se seleccionan las variables que tienen efectos *robustos* sobre el número acumulado de contagios (magnitud) o sobre la tasa de prevalencia (intensidad) desarrollando un análisis de límites extremos (EBA, por sus siglas del inglés *extreme-bounds analysis*) propuesto por Leamer (1983) y popularizado por Levine y Renelt (1992). Segundo, ante la posibilidad de que los contagios hayan transitado de un estado a otro se busca evidencia de dependencia espacial y se estiman los correspondientes modelos econométricos.

El EBA consiste en estimar *relaciones robustas* entre la variable explicada y las que sugiere la literatura relevante. Se parte de un modelo pequeño que se va ampliando si las variables que se van agregando son estadísticamente significativas, del signo esperado y no afectan a las estimaciones de las variables previamente incluidas. Este enfoque podría verse como uno de “lo particular a lo general”, con claras ventajas cuando se tienen pocos grados de libertad, como es este caso (32 estados), y una amplia variedad de posibles variables explicatorias.

Entonces, el objetivo es ir definiendo el orden de entrada de las variables explicatorias al modelo. Para ello se parte de una regresión bivariada que incluye como variable explicatoria solamente la que se espera que tenga una relación robusta con los contagios (con base en la teoría o en evidencia empírica de otros estudios). En el caso de la variable dependiente tasa de prevalencia (*TC*), la variable explicatoria robusta resultó ser la densidad poblacional (*DP*), mientras que para los contagios acumulados (*CA*) fue una del grupo de variables sociales que miden el respeto a las normas;

específicamente *ICL*, *CDEL*, *PCORE*, *DEN* o *INDEL*, lo que nos lleva a varios modelos iniciales.¹⁴ En las tablas 1 y 2 se resaltan estas variables en negritas.

Una vez identificada esa primera variable se amplía cada modelo resultante (varios) con una segunda de los diferentes grupos, con los requerimientos mencionados sobre su significancia estadística, signo y no afectación de las estimaciones de la primera variable. En esta segunda etapa resultaron robustos los efectos de variables del conjunto de condiciones de salud y la densidad de población (*DP*), entre otras.¹⁵

Se procedió de esta manera para conformar modelos en los que finalmente se incorporaron hasta tres variables explicatorias; las que se agregaron posteriormente no mostraron efectos estadísticamente significativos. De esta forma, se tienen catorce modelos lineales que contienen variables con efectos robustos sobre el contagio. A partir de ahí, en una segunda etapa, se hace el análisis para determinar si es necesario incorporar efectos espaciales. Los estadísticos de prueba correspondientes al *I* de Moran (Moran, 1948) y a los multiplicadores de Lagrange, *LM* (Burrige, 1980; Anselin, 1988a, 1988b) se presentan en la tabla A2 del anexo. Los modelos en los que no es necesario incorporar efectos espaciales se muestran en la tabla 1 y los modelos espaciales en la 2.¹⁶ En ambas tablas se reportan también las pruebas de especificación de normalidad y homoscedasticidad; en la segunda se agrega la prueba de dependencia espacial remanente.¹⁷

En general, el *I* de Moran sólo es estadísticamente significativo en el modelo 7 (con un valor *p* igual a 0.050). No obstante, los estadísticos *LM* son significativos en siete modelos: sugieren la estimación de un modelo de rezago espacial en los casos 1, 2, 3, 4, 8 y 9 y de error espacial en el 13. En el resto no se encuentra evidencia de dependencia espacial, por lo que se estiman por MCO (véase tabla A2 de anexo).

En las regresiones presentadas en la tabla 1 se observa que ninguna muestra problema de normalidad o heteroscedasticidad, pues los valores *p* correspondientes sugieren no rechazar las hipótesis nulas al 5% de significancia. El poder explicativo de los modelos, a su vez, es relativamente alto, con coeficientes de determinación por encima de 0.7 en un caso y de alrededor de 0.9 en el resto.

Destaca el hecho de que todas incluyen un efecto positivo de la densidad de población (*DP*) sobre la magnitud de los contagios (modelos 5, 6, 7, 10, 11 y 12) y la tasa de prevalencia (modelo 14). Este último, además incluye a la otra variable demográfica considerada *TH*, la cual tiene un coeficiente positivo y significativo estadísticamente, por lo que se puede sostener que a mayor proporción de hombres en un estado mayor será la tasa de contagios.¹⁸ La tercera variable que explica la tasa de prevalencia es la proporción de la población estatal que padece de hipertensión, lo que comprueba la hipótesis de que padecimientos previos que debilitan el sistema inmunológico favorecen la adquisición del virus.

Es interesante observar que la primera variable (en negritas) con efectos robustos sobre los contagios acumulados es del grupo de indicadores que buscan medir el respeto a las normas. La idea de incorporar estas variables es que sirvan como *proxy* del acatamiento de la población de las disposiciones sanitarias. En ese sentido, el Costo del delito (*CDEL*) resulta significativo y con efectos positivos sobre el nivel de contagio en los modelos 5, 6 y 7, al igual que la proporción de población con hipertensión (*THIP*), obesidad (*TOB*) o ambas (*THIPOB*). A su vez, la variable Delitos denunciados (*DEN*) tiene efectos similares en el modelo 10, junto con la multi-morbilidad asociada a obesidad, hipertensión y diabetes (*THIPOBDIA*). Por último, la Incidencia delictiva (*INDEL*) se combina con el padecimiento de

¹⁴ Se realizó también un análisis amplio partiendo de prácticamente todas las variables mencionadas en la sección anterior como variable inicial en los modelos bivariados.

¹⁵ En las etapas subsiguientes se introduce la variable que resultó más significativa hasta la menos significativa en modelos bivariados previamente estimados.

¹⁶ Son las pruebas de heteroscedasticidad (Breusch y Pagan, 1980) y normalidad (Jarque y Bera, 1987).

¹⁷ Prueba de razón de verosimilitud que contrasta si existe algún remanente de autocorrelación espacial residual.

¹⁸ Además, se ha observado que las muertes por COVID-19 en México se han dado principalmente entre pacientes varones mayores que también tienen múltiples comorbilidades (Padilla-Santamaría *et al.* (2020)).

diabetes (*TDIA*) o de obesidad (*TOB*) para explicar la acumulación de contagios de COVID-19 en los modelos 11 y 12, respectivamente.

Tabla 1

Factores que influyen en la magnitud y tasa de prevalencia de contagios de COVID-19 en México

Variables	Variable dependiente: CA						Variable dependiente: TC (14)*
	(5)	(6)	(7)	(10)	(11)	(12)	
Intercepto	-9100.3 (0.093)	-5087.1 (0.135)	-11495.2 (0.064)	-1140.26 (0.697)	-11015.6 (0.081)	-8744.36 (0.211)	-47.000 (0.003)
<i>DP</i>	7.214 (0.000)	7.381 (0.000)	7.424 (0.000)	6.567 (0.000)	6.818 (0.000)	7.684 (0.000)	0.001 (0.001)
<i>TH</i>							0.971 (0.003)
<i>CDEL</i>	0.882 (0.000)	0.883 (0.000)	0.895 (0.000)				
<i>DEN</i>				0.069 (0.000)			
<i>INDEL</i>					0.289 (0.000)	0.280 (0.000)	
<i>TDIA</i>					1436.417 (0.015)		
<i>THIP</i>	709.173 (0.014)						0.131 (0.043)
<i>TOB</i>			412.129 (0.012)			342.442 (0.060)	
<i>THIPOB</i>		835.286 (0.005)					
<i>THIPOBDIA</i>				1104.767 (0.033)			
R^2	0.910	0.916	0.911	0.898	0.891	0.881	0.719
Normalidad	1.212 (0.546)	0.839 (0.657)	2.024 (0.363)	0.576 (0.749)	0.297 (0.862)	1.553 (0.460)	1.906 (0.385)
Heteroscedasticidad	1.360 (0.715)	3.732 (0.292)	1.374 (0.712)	4.411 (0.220)	2.222 (0.528)	2.408 (0.492)	2.106 (0.834)

Los coeficientes en negritas, cursivas y normales corresponden a las variables identificadas en primero, segundo y tercer lugar en los diferentes modelos. Los valores *p* se presentan entre paréntesis. * El modelo se estima con dos variables binarias que valen 1 en Chihuahua y Tabasco, y 0 en los demás estados, respectivamente, con el objetivo de captar valores anómalos en los residuos. Los coeficientes estimados son 5.140 y -2.583, con valores *p* de 0.000 y 0.029, respectivamente. Fuente. Elaboración propia.

Sobre los factores que explican los contagios, nuevamente destacan las variables sociales que buscan medir el respeto a las normas sanitarias. En los modelos 1 y 2 se encuentra un efecto robusto del Índice de cultura de la legalidad (*ICL*), construido por variables que evalúan la forma en la que las personas encuestadas declaran el respeto, conocimiento y cumplimiento de las leyes y prácticas cívicas: los coeficientes estimados son negativos y estadísticamente significativos, sugiriendo que una mayor cultura de la legalidad lleva a un menor nivel de contagios, y viceversa. Interesantemente, en estos modelos también entran variables que buscan medir la misma conducta de las personas ante las normas: entre mayor es la incidencia delictiva (*INDEL*) se tienen más contagios estatales. Curiosamente, la percepción de seguridad de la población (*PINS*) tiene un signo contrario al esperado, ya que la lógica sugeriría que entre mayores valores tome esta variable más respeto se tiene por las normas, en general, y por las sanitarias, en particular, por lo que debería haber menos contagios. Por último, al igual que en los modelos lineales, el padecimiento de enfermedades crónico-degenerativas, en este caso diabetes (*TDIA*), ha contribuido a que haya más casos de COVID-19.

Por otra parte, los modelos de regresión espacial aparecen en la tabla 2. De acuerdo a las implicaciones de los estadísticos *LM*, en primer lugar se muestran los modelos de rezago espacial y finalmente los de error espacial. En general los coeficientes de autocorrelación espacial sustantiva (ρ), en los modelos de los contagios acumulados, y residual (λ), en

los de la tasa de prevalencia, son significativos estadísticamente al menos al 5% (sólo en el modelo 4, ρ es significativa al 10% y en el 8 no lo es). Llama la atención, sin embargo, que todos ellos sean negativos, lo que sugiere la inexistencia de conglomerados en los que estados con altos niveles de contagio sean vecinos de otros con similares patrones, y viceversa. Los mapas de las figuras 3 y 4 sugieren que efectivamente ese es el caso: hay un corredor central en el que se ubican estados con bajos niveles de contagio que colindan con otros de la costa del Océano Pacífico y del centro de México que han sido severamente afectados por la pandemia. En suma, esta evidencia sugiere la inexistencia de transmisión interestatal de los contagios.¹⁹

Tabla 2
 Factores que influyen en la magnitud y tasa de prevalencia de contagios de COVID-19 en México (Modelos de regresión espacial)

Variables	Rezago espacial. Variable dependiente: CA						Error espacial Variable dependiente: TC
	(1)*	(2)**	(3)	(4)	(8)	(9)	(13)
Intercepto	26546.4 (0.019)	41785.6 (0.001)	-10680 (0.058)	-1987.14 (0.479)	-5757.87 (0.333)	-11877.4 (0.041)	-5.699 (0.000)
<i>DP</i>			<i>7.453</i> (0.000)	<i>7.723</i> (0.000)	<i>6.348</i> (0.000)	<i>6.706</i> (0.000)	0.001 (0.000)
<i>ICL</i>	- 92758.4 (0.001)	-122759 (0.000)					
<i>CDEL</i>			0.898 (0.000)	0.873 (0.000)			
<i>PCORE</i>					0.009 (0.000)		
<i>DEN</i>						0.071 (0.000)	
<i>INDEL</i>	0.295 (0.000)						
<i>PINS</i>		0.009 (0.010)					
<i>TDIA</i>	<i>1838.04</i> (0.000)	<i>1777.99</i> (0.000)	149892 (0.000)		1191.08 (0.014)	162722 (0.000)	
<i>TOB</i>							0.200 (0.000)
<i>THIPOBDIA</i>				107859 (0.012)			
<i>BL</i>							7.873 (0.025)
ρ	-0.293 (0.008)	-0.356 (0.007)	-0.275 (0.009)	-0.209 (0.078)	-0.206 (0.108)	-0.232 (0.035)	
λ							-0.437 (0.039)
R^2	0.925	0.906	0.930	0.919	0.903	0.922	0.637
Normalidad	0.183 (0.913)	5.157 (0.076)	0.186 (0.911)	0.518 (0.771)	0.333 (0.285)	0.106 (0.948)	0.281 (0.869)
Heteroscedasticidad	0.713 (0.950)	2.159 (0.827)	2.677 (0.444)	1.596 (0.660)	1.817 (0.611)	4.859 (0.182)	***
Dependencia espacial	5.841 (0.016)	6.663 (0.010)	6.490 (0.011)	3.413 (0.065)	2.875 (0.090)	4.349 (0.037)	

Los coeficientes en negritas, cursivas y normales corresponden a las variables identificadas en primero, segundo y tercer lugar en los diferentes modelos. Los valores p se presentan entre paréntesis. * El modelo se estima con una variable

¹⁹ Nótese que las pruebas de especificación sugieren la existencia de autocorrelación remanente. Para tratar de modelarla se estimaron modelos SARMA, que incorporan tanto autocorrelación sustantiva como residual; no obstante, los coeficientes autorregresivos adicionales no fueron significativos. Por ello, se mantuvo la especificación sugerida por los estadísticos LM.

binaria que valen 1 en CDMX y 0 en los demás estados; el valor de su coeficiente es 40497.9 y el valor p respectivo es igual a 0.000. ** El modelo se estima con dos variables binarias que valen 1 en CDMX y Estado de México y 0 en los demás estados; los valores de los coeficientes estimados son 60825.3 y 37200.9, con valores p ambos de 0.000. *** Se usó la prueba propuesta por Kelejian y Prucha (2010) para la obtención de estimadores robustos con heteroscedasticidad y autocorrelación.

Fuente. Elaboración propia.

Al igual que en los modelos lineales, la densidad de población (DP) contribuye positivamente a los contagios acumulados (modelos 3, 4, 8 y 9) y a la tasa de prevalencia (modelo 13). Similarmente, las variables asociadas al respeto a la norma resultan significativas en la explicación de los contagios acumulados, tales como el costo del delito ($CDEL$) en los modelos 3 y 4; percepción de corrupción ($PCORE$) en el modelo 8, y delitos denunciados (DEN). Como era anticipado, estas variables tienen efectos positivos y significativos estadísticamente. La tasa de diabetes ($TDIA$), por su parte, entra en los modelos 3, 8 y 9, mientras que la multi-morbilidad de diabetes, hipertensión y obesidad ($THIPOBDIA$) lo hace en el modelo 4.

La tasa de prevalencia (modelo 13) es explicada, además de DP , por la tasa de obesidad (TOB) y por la única variable económica que entra en nuestras especificaciones, la brecha laboral BL . Esta última tiene un efecto positivo y significativo, sugiriendo que entre mayor es la proporción de población que no cuenta con un puesto de trabajo le es más difícil mantener la sana distancia por tener que buscar un medio de subsistencia, lo que la expone más al contagio.

En general, la evidencia reportada en este documento es consistente con algunos de los hallazgos de la temprana literatura internacional sobre el tema. En particular, coincide con los reportes sobre los efectos positivos de la densidad de población sobre los contagios, ya que entre más personas se concentra en espacios menores es mayor la cercanía física, lo que eleva la posibilidad de contagio (Kumar, 2015; Hoehl *et al.*, 2020; Gilbert *et al.*, 2020, y Padilla-Santamaría *et al.*, 2020). Por lo tanto, los lugares densamente poblados son más propensos a una propagación viral más rápida (Gerritse, 2020; Lakshmi-Priyadarsini y Suresh, 2020). Similarmente, apoya los planteamientos de Yanga *et al.* (2020) y de Dietz y Santos-Burgoa (2020), quienes encuentran que las enfermedades subyacentes, incluyendo hipertensión, diabetes, enfermedades del sistema respiratorio y enfermedades cardiovasculares, y la obesidad, respectivamente, pueden ser factores de riesgo para pacientes graves en comparación con pacientes no graves. Finalmente, da soporte empírico a la conjetura de que la falta de seguimiento de las normas sanitarias dictadas por la autoridad provoca una mayor magnitud e intensidad del contagio de COVID-19.

Conclusiones

En este documento se analiza el impacto de factores demográficos, sociales, sanitarios y económicos en los niveles acumulados de contagio y en la tasa de prevalencia del COVID-19 en los estados de México mediante un análisis de límites extremos (*extreme-bounds analysis*) desarrollado en modelos de corte transversal en los que se determina también la existencia de efectos espaciales. Aun cuando son varias las especificaciones de los modelos a las que se llega, hay una serie de resultados consistentes con los hallazgos en la literatura internacional que vale la pena resaltar.

En primer lugar, las variables usadas como *proxy* para medir el respeto a las normas, en general, y, de manera indirecta, a las disposiciones sanitarias dictadas por las autoridades para contener el contagio del coronavirus tienen efectos significativos sobre la magnitud e intensidad de COVID-19 en los estados. En particular, el no cumplimiento de las normas sanitarias parece ser un factor importante en la propagación del virus. En segundo lugar, el contagio de SARS-CoV-2 ha sido más frecuente en estados donde es mayor la proporción de población que padece obesidad y/o enfermedades crónico-degenerativas debido al deterioro del sistema inmunológico que éstas generan. En tercer lugar, la mayor cercanía física entre las personas en ciudades altamente pobladas ha sido uno de los factores centrales en el contagio. Por último, es importante mencionar que el poder explicativo de los modelos es relativamente elevado, lo que

sugiere que estos factores y la combinación de ellos pueden explicar tanto la magnitud como la intensidad de los contagios en los estados de México.

Las principales implicaciones de política que se pueden derivar directamente de nuestros resultados se relacionan con la necesidad de proteger a la población con obesidad y/o con padecimientos de hipertensión y diabetes ya que, como se ha dicho, constituyen un grupo altamente vulnerable; mantener y reforzar las medidas sanitarias, especialmente el distanciamiento físico, el uso de cubre-bocas y la higiene personal, en particular en las localidades densamente pobladas.

Llama la atención, por otro lado, que algunas variables consideradas importantes por los diseñadores de política y la opinión pública no hayan resultado relevantes. Por ejemplo, las políticas públicas dictadas por los estados o la movilidad física, así como diferentes indicadores de las condiciones de vida de la población (pobreza, empleo precario y desigualdad) no resultaron significativas en nuestros modelos. Tampoco variables económicas como la proporción de actividades altamente sensibles al confinamiento (turismo y entretenimiento) o el tamaño del sector informal parecen ser importantes en este proceso. Más aun, de acuerdo a los resultados de los modelos espaciales estimados, tampoco hay evidencia de transmisión inter-estatal del coronavirus.

La pandemia de COVID-19 es un proceso complejo, que sigue su curso y que tiene múltiples aristas que se deben analizar en el futuro cercano para comprender mejor su naturaleza y sus implicaciones. La velocidad y tiempos de contagio y los fallecimientos son algunos aspectos que podrían analizarse con métodos similares a los aquí usados. El análisis, no obstante, se puede extender para buscar posibles relaciones no lineales entre las variables explicatorias, en ámbitos más desagregados, como ciudades o colonias. Posiblemente entonces resulten relevantes algunas variables intuitivamente importantes.

Referencias

- Anselin, L. (1988a). *Spatial Econometrics: Methods and Models*. Kluwer Academic Publishers, Dordrecht, The Netherlands.
- Anselin, L. (1988b). A test for spatial autocorrelation in seemingly unrelated regressions. *Economics Letters*, 28 (4), 335–341. [https://doi.org/10.1016/0165-1765\(88\)90009-2](https://doi.org/10.1016/0165-1765(88)90009-2)
- Anselin, L. & Bera, A. K. (1998). Spatial Dependence in Linear Regression Models with an Introduction to Spatial Econometrics. In A. Ullah y D. Giles (eds.), *Handbook of Applied Economic Statistics*. (pp. 237-289). New York, NY. <https://doi.org/10.1201/9781482269901-36>
- Blanchflower, D. & Levin, A. (2015). Labor market slack and monetary policy. *NBER Working Paper*, No. 21094. Cambridge, MA. <https://doi.org/10.3386/w21094>
- Breusch, T. & Pagan, A. (1980). The Lagrange multiplier and its applications to model specification in econometrics. *Review of Economics Studies*. 47 (1), 239-253. <https://doi.org/10.2307/2297111>
- Burridge, P. (1980). On the Cliff-Ord test for spatial correlation. *Journal of the Royal Statistical Society*, 42 (1), 107–108. <https://doi.org/10.1111/j.2517-6161.1980.tb01108.x>
- Callaway, E., Cyranoski, D., Mallapaty, S., Stoye, E. & Tollefson, J. (2020). The coronavirus pandemic in five powerful charts. *Nature*, 579, 482–483. <http://doi.org/10.1038/d41586-020-00758-2>
- Carrión, J. Ll., Di Paolo, A., García, A., López-Bazo, E., López-Tamayo, J., Matano, A., Moreno, R., Pons, E., Ramos, R., Royuela, V. & Suriñach, J. (2020). La propagación del COVID y su distribución geográfica en áreas pequeñas. *Un análisis para Cataluña (I) Densidad y nivel socioeconómico. La Riqueza de las Naciones. Asociación Española de Ciencia Regional*. (Disponible en: <https://aecr.org/es/la-propagacion-del-covid-y-su-distribucion-geografica-en-areas-pequenas-un-analisis-para-cataluna-i-densidad-y-nivel-socioeconomico/>) (Consultado: 25/08/2020).
- Carteni, A., Di Francesco, L. & Martino, M. (2020). How mobility habits influenced the spread of the COVID-19 pandemic: Results from the Italian case study. *Science of the Total Environment*, 741, 1-9. <https://doi.org/10.1016/j.scitotenv.2020.140489>
- Díaz-Castrillón, F. J. & Toro-Montoya, A. I. (2020). SARS-CoV-2/COVID-19: el virus, la enfermedad y la pandemia. *Medicina & laboratorio*, 24 (3), 183-205. <https://doi.org/10.36384/01232576.268>
- Dietz, W. & Santos-Burgoa C. (2020). Obesity and its implications for COVID-19 Mortality. *Obesity*, 28 (6). <https://doi.org/10.1002/oby.22818>

- ECDC (2020). Data on the geographic distribution of COVID-19 cases worldwide. European Centre for Disease Prevention and Control. (Disponible en: <https://www.ecdc.europa.eu/en/publications-data/download-todays-data-geographic-distribution-covid-19-cases-worldwide>) (Consultado: 03/08/2020).
- Flournoy, M. & Morell, M. (2020, April 19). The 6 factors that determine coronavirus containment or devastation. *The Washington Post*. (Disponible en: <https://www.washingtonpost.com/opinions/2020/04/19/6-factors-that-determine-coronavirus-containment-or-devastation/>) (Consultado: 26/08/2020).
- Gerritse, M. (2020). Cities and COVID-19 infections: Population density, transmission speeds and sheltering responses. *Covid Economics vetted at real-time papers*, Centre for Economic Policy Research (CEPR), 37, 1-26.
- Gilbert, M., Pullano, G. & Pinotti, F. (2020). Preparedness and vulnerability of African Countries against importations of COVID-19: a modelling study. *THE LANCET Regional Health*, 395 (10227), 871-877. [http://doi.org/10.1016/S0140-6736\(20\)30411-6](http://doi.org/10.1016/S0140-6736(20)30411-6)
- Gobierno de México (2020). *Covid-19 México*. (Disponible en: <https://coronavirus.gob.mx/datos/#COMNac>) (Consultado: 18/08/2020).
- Gorbalenya, A. E., Barker, S.C., Baric, R.S., de Groot, R.J., Drosten, C., Gulyaeva, A.A., Haagmans, B.L., Lauber, C., Leontovich, A. M., Neuman, B. W., Penzar, D., Perlman, S., Poon, L. L. M., Samborskiy, D. V., Sidorov, I. A., Sola, I. & Ziebuhr, J. (2020). The species Severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2. *Nature Microbiology*, 5, 536-544. <https://doi.org/10.1038/s41564-020-0695-z>.
- Gupta, R., Ghosh, A., Singh, A. & Misra, A. (2020). Clinical considerations for patients with diabetes in times of COVID-19 epidemic. *Diabetes & Metabolic Syndrome: Clinical Research & Reviews*, 14 (3), 211-212. <https://doi.org/10.1016/j.dsx.2020.03.002>
- Hoehl, S., Rabenau, H., Berger, A., Kortenbusch, M., Cinatl, J. & Bojka, S. (2020). Evidence of SARS-CoV-2 Infection in Returning Travelers from Wuhan, China. *The New England Journal of Medicine*, 382, 1278-1280. <http://doi.org/10.1056/NEJMc2001899>
- Holshue, M.L., DeBolt, C., Lindquist, S., Lofy, K.H., Wiesman, J., Bruce, H. & Díaz, G. (2020). First case of 2019 novel coronavirus in the United States. *New England Journal of Medicine*, 382, 929-936. <https://doi.org/10.1056/NEJMoa2001191>
- INEGI (2020). Banco de Información Económica. (Disponible en: <https://www.inegi.org.mx/sistemas/bie/>) (Consultado: 15/08/2020).
- Jarque, C. M. & Bera, A. K. (1987). A test for normality of observations and regression residuals. *International Statistical Review*, 55 (2), 163-172. <https://doi.org/10.2307/1403192>.
- Kelejian, H. H. & Prucha, I. R. (2010). Specification and estimation of spatial autoregressive models with autoregressive and heteroskedastic disturbances. *Journal of Econometrics*, 157 (1), 53-67. <https://doi.org/10.1016/j.jeconom.2009.10.025>
- Kumar, M. (2015). A study of population distribution. *International Journal of Latest Technology, Engineering Management and Applied Science*, 4 (3), 24-28.
- Lakshmi-Priyadarsini, S. & Suresh, M. (2020). Factors influencing the epidemiological characteristics of pandemic COVID 19: A TISM approach. *International Journal of Healthcare Management*, 13 (2), 89-98. <http://doi.org/10.1080/20479700.2020.1755804>
- Leamer, E. E. (1983). Let's take the con out of econometrics. *The American Economic Review*, 73 (1), 31-43. (Disponible en: <https://www.jstor.org/stable/1803924>) (Consultado: 17/08/2020).
- Levine, R. & Renelt, D. (1992). A sensitivity analysis of cross-country growth regressions. *The American Economic Review*, 82 (4), 942-963. (Disponible en: <https://www.jstor.org/stable/2117352>) (Consultado: 17/08/2020).
- Li, B., Yang, J. Zhao, F., Zhi, L., Wang, X., Liu, L., Bi, Z. & Zhao, Y. (2020). Prevalence and impact of cardiovascular metabolic diseases on COVID-19 in China. *Clinical Research in Cardiology*, 109 (5), 531-538. <http://doi.org/10.1007/s00392-020-01626-9>
- Ma, J. (2020). Coronavirus: China's first confirmed Covid-19 case traced back to november 17. *South China Morning Post*. (Disponible en: <https://www.scmp.com/news/china/society/article/3074991/coronavirus-chinas-first-confirmed-covid-19-case-traced-back>) (Consultado: 15/08/2020).
- Mollalo, A., Vahedi, B., & Rivera, K. M. (2020). GIS-based spatial modeling of COVID-19 incidence rate in the continental United States. *Science of the Total Environment*, 728, 1-8. <https://doi.org/10.1016/j.scitotenv.2020.138884>
- Moran, P. A. P. (1948). The Interpretation of Statistical Maps. *Journal of the Royal Statistical Society B (Methodological)* 10 (2), 243-251. <https://doi.org/10.1111/j.2517-6161.1948.tb00012.x>
- MUCD (2014). Primer Índice de Cultura de la Legalidad en México. (Disponible en: <https://www.mucd.org.mx>) (Consultado: 10/08/2020).
- Observatorio COVID-19 (2020). Observatorio para la Contención del COVID-19 en América Latina. (Disponible en: www.ObsContenCovid.INFO) (Consultado:08/08/2020).
- OMS (2020a). Actualizaciones continuas sobre la enfermedad por coronavirus (COVID-19). (Disponible en: <https://www.who.int/emergencias/diseases/novel-coronavirus-2019/events-as-they-happen>) (Consultado: 15/08/2020).

- OMS (2020b). Tablero de la enfermedad por coronavirus (COVID-19). (Disponible en: <https://covid19.who.int/table>) (Consultado: 15/08/2020).
- Onder, G., Rezza, G. & Brusaferro, S (2020). Case Fatality Rate and Characteristics of Patients Dying in Relation to COVID-19 in Italy. *American Medical Association*, 323 (18), 1775-1776. <http://doi.org/10.1001/jama.2020.4683>
- Padilla-Santamaría, F., Maya-Franco, L. & Ferman-Cano, F. (2020). COVID-19 en México: Panorama Epidemiológico. *Revista Cadena de Cerebros*, 4 (1), 31-42. <https://doi.org/10.5281/zenodo.3926806>
- PAHO/WHO (2020). Novel coronavirus (COVID-19). Epidemiological Update. Pan American Health Organization. (Disponible en: <https://www.paho.org/en/documents/epidemiological-update-novel-coronavirus-covid-19-28-february-2020>) (Consultado: 30/08/2020).
- Papageorge, N., Zahn, M. V., Belot, M., van den Broek-Altenburg, E., Choi, S., Jamison, J. C. & Tripodi, E. (2020). Socio-demographic factors associated with self-protecting behavior during the Covid-19 pandemic. *NBER Working Paper No. 27378*. <https://doi.org/10.3386/w27378>
- PNUD (2015). Índice de Desarrollo Humano para las entidades federativas, México 2015. (Disponible es: <https://www.mx.undp.org/content/mexico/es/home/library/poverty/indice-de-desarrollo-humano-para-las-entidades-federativas--mexi.html>) (Consultado: 14/08/2020).
- Reuters (2020). Coronavirus. Francia confirma dos casos, los primeros en Europa. *El Universal*. (Disponible en: <https://www.eluniversal.com.mx/mundo/coronavirus-francia-confirma-dos-casos-los-primero-en-europa>) (Consultado: 10/08/2020).
- Xie, Z., Qin, Y., Li, Y., Shen, W., Zheng, Z. & Liu, S. (2020). Spatial and temporal differentiation of COVID-19 epidemic spread in mainland China and its influencing factors. *Science of the Total Environment*, 744, 1-9. <https://doi.org/10.1016/j.scitotenv.2020.140929>
- Yanga, J., Zheng, Y., Gou, X., Pu, K., Chen, Z., Guo, Q., Ji, R., Wang, H., Wang, Y. & Zhou, Y. (2020). Prevalence of comorbidities and its effects in patients infected with SARS-CoV-2: a systematic review and meta-analysis. *International of Journal of Infectious Diseases*, 94, 91-95. <https://doi.org/10.1016/j.ijid.2020.03.017>

Anexo

Tabla A1
 Definición de variables

Variables	Nomenclatura	Unidades	Fecha	Fuente de los datos
Casos acumulados de casos COVID-19 por estado	CA	Personas	31 de julio 2020	https://coronavirus.gob.mx/datos/#COMNac
Tasa de prevalencia*	TC	Tasa	31 de julio 2020	https://coronavirus.gob.mx/datos/#COMNac
Densidad de población	DP	Hab/km ²	2020	CONAPO e INEGI
% de población masculina	TH	Porcentaje	2015	CONAPO (estimaciones 2020)
Índice de Cultura de la Legalidad**	ICL	Índice	2014	MUCD https://www.mucd.org.mx
Costos del delito	CDEL	Pesos por persona	2019	
Delitos denunciados	DEN	Delitos	2019	
Percepción de seguridad	PINS	Porcentaje	2019	INEGI Encuesta Nacional de Victimización y Percepción sobre Seguridad
Incidencia de corrupción***	INCOR		2019	
Percepción de corrupción estatal	PCORE	Porcentaje	2019	
Incidencia delictiva ⁺	INDEL	Delitos /1000 habitantes	2020	Secretaría de Gobernación
Porcentaje de población con obesidad	TOB	Porcentaje	2018	
Porcentaje de población con diagnóstico previo de hipertensión	THIP	Porcentaje	2018	
Porcentaje de población con diagnóstico previo de diabetes	TDIA	Porcentaje	2018	Encuesta Nacional de Salud y Nutrición (ENSANUT) 2018; Secretaría de Salud, el Instituto Nacional de Salud Pública (INSP); y, el Instituto Nacional de Estadística y Geografía (INEGI)
Porcentaje de población con diagnóstico de hipertensión y obesidad	THIPOB	Porcentaje	2018	
Porcentaje de población con diagnóstico de hipertensión, obesidad y diabetes	THIPOBDIA	Porcentaje	2018	
Brecha laboral [#]	BL		2020.I	INEGI https://www.inegi.org.mx/sistemas/Infoenoe/Default

* $TC = \frac{CA}{Población} \times 1000$. ** Proporciona elementos para conocer si las personas encuestadas declaran tener respeto, conocimiento y cumplimiento de las leyes y prácticas cívicas. *** Incidencia de corrupción por cada 100 mil habitantes. + Delitos del fuero común por cada mil habitantes. # Es el porcentaje de la población económicamente activa potencial.
 Fuente: Elaboración propia.

Tabla A2
 Pruebas de diagnóstico para dependencia espacial

Pruebas	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
I de Moran	-0.655 (0.512)	-1.051 (0.293)	0.275 (0.783)	-0.300 (0.764)	-0.064 (0.949)	0.519 (0.603)	1.956 (0.050)	-0.464 (0.642)	-0.545 (0.586)	-0.135 (0.892)	-0.015 (0.988)	1.391 (0.164)	-1.436 (0.151)	1.146 (0.252)
LM rezago	4.460 (0.034)	5.651 (0.017)	5.165 (0.023)	3.381 (0.066)	2.668 (0.102)	1.766 (0.184)	1.009 (0.315)	2.880 (0.090)	3.749 (0.053)	1.940 (0.164)	2.304 (0.129)	0.219 (0.640)	2.630 (0.105)	0.014 (0.907)
LM error	0.787 (0.374)	1.486 (0.223)	0.000 (1.000)	0.306 (0.580)	0.115 (0.734)	0.012 (0.912)	1.895 (0.169)	0.463 (0.496)	0.566 (0.452)	0.153 (0.695)	0.065 (0.798)	0.763 (0.382)	3.212 (0.073)	0.365 (0.545)
LM rezago robusta	3.699 (0.054)	4.218 (0.040)	6.458 (0.011)	3.251 (0.071)	2.869 (0.090)	2.561 (0.110)	3.852 (0.050)	2.489 (0.115)	3.211 (0.073)	1.938 (0.164)	2.638 (0.104)	1.517 (0.218)	0.106 (0.745)	0.327 (0.567)
LM error robusta	0.027 (0.869)	0.053 (0.818)	1.293 (0.255)	0.176 (0.675)	0.317 (0.573)	0.808 (0.369)	4.738 (0.030)	0.073 (0.788)	0.028 (0.868)	0.151 (0.697)	0.400 (0.527)	2.061 (0.151)	0.687 (0.407)	0.678 (0.410)
LM SARMA	4.487 (0.106)	5.704 (0.057)	6.458 (0.040)	3.557 (0.169)	2.985 (0.225)	2.573 (0.276)	5.747 (0.057)	2.952 (0.229)	3.777 (0.151)	2.091 (0.351)	2.704 (0.259)	2.280 (0.320)	3.317 (0.190)	0.692 (0.708)

Fuente: Elaboración propia.