

Brechas de Datos y Política de Respuesta al Nuevo Coronavirus



James H. Stock

Universidad de Harvard, marzo de 2020

Harvard Environmental Economics Program

Traducción: Enrique A. Bour

Resumen

En esta nota se expone el modelo epidemiológico básico de contagio Susceptible-Infectado-Recuperado (SIR), con un público objetivo de economistas que buscan un marco para comprender los efectos de las políticas de distanciamiento y contención sobre la evolución del contagio y las interacciones con la economía. Un parámetro clave, la tasa de asintomáticos (fracción de los infectados que no son sometidos a examen en el marco de las actuales pautas), no está bien estimado en la literatura porque las pruebas para el coronavirus se han enfocado en enfermos y vulnerables, sin embargo, ésta podría ser estimada mediante un muestreo aleatorio de la población. En este modelo simple, las diferentes políticas que dan lugar a la misma tasa de transmisión β tienen el mismo resultado en términos de salud, pero pueden tener costos económicos muy diferentes. Por lo tanto, una forma de enmarcar la economía de la política de confinamiento sería buscando las políticas más eficientes para lograr un determinado β , y luego determinando el rumbo del β que compensa el costo económico en comparación con el costo del exceso de vidas perdidas por sobrecargar el sistema de atención sanitaria.

Palabras clave: Modelo epidemiológico, macroeconomía, confinamiento

Autor agradece a Andy Atkeson, Oleg Itshoki, Erin Lake y Arash Nekoei por sus útiles comentarios.

James Harold Stock (n. 1955) es un economista estadounidense y profesor de economía en la Universidad de Harvard. Es co-autor de *Introduction to Econometrics*, un reconocido libro de texto para estudiantes universitarios, y co-editor de *Brookings Papers on Economic Activity*. Stock ocupó la presidencia del Departamento de Economía de Harvard de 2007 a 2009 y fue miembro del Consejo de Asesores Económicos del Presidente Obama de 2013 a 2014

En esta nota se expone el modelo epidemiológico básico de contagio Susceptible-Infectado-Recuperado (SIR), con un público objetivo de economistas que buscan un marco para comprender los efectos de las políticas de distanciamiento y contención de la evolución del contagio y sus interacciones con la economía.¹ El modelo está calibrado con los datos más recientes. Su naturaleza simple hace abstracción de muchas cuestiones importantes y su resultado no tiene por objeto sustituir las estimaciones de estudios epidemiológicos más sofisticados.

Esta nota establece cuatro puntos principales. Primero, el efecto del distanciamiento social y los cierres de negocios en la dinámica de la epidemia entra en el modelo a través de un único parámetro, la tasa de transmisión de casos β . Para una tasa de transmisión de casos dada, la cuestión de diseño de políticas es cómo lograr esa tasa de transmisión de casos minimizando el costo económico. Una segunda pregunta de índole económica es cuál es la trayectoria óptima del caso para β , comparando el costo económico de ese camino con los costos en defunciones. El modelo sirve para enfocar la atención sobre estas cuestiones de importancia central en la interacción entre política sanitaria y política económica.

En segundo lugar, los parámetros del modelo no están bien estimados en la literatura sobre el coronavirus debido a la falta de datos disponibles. Los datos sobre prevalencia, por ejemplo, se obtienen de las tasas positivas de pruebas de detección del coronavirus, pero hasta ahora las pruebas se han racionado y se han administrado casi en su totalidad a una población seleccionada, los que están en riesgo y muestran síntomas. Así pues, la fracción de pruebas que resultan positivas no permite estimar la tasa de infección de la población.

En tercer lugar, utilizando el principio de Bayes, es posible reexpresar el modelo en términos de β y la tasa de asintomáticos, que es la fracción de infectados que muestran síntomas lo suficientemente leves, o ninguno, como para que no sean sometidos a tests bajo las normas actuales de pruebas. La virtud de reexpresar el modelo de

¹ Después de escribir esta nota, dos documentos excelentes y muy relacionados entre sí han aparecido como documentos de trabajo del NBER, Atkeson (2020) y Eichenbaum, Rebelo y Trabandt (2020). Atkeson (2020) opera a través del modelo SIR y proporciona simulaciones para una versión calibrada bajo diferentes escenarios de aislamiento. Eichenbaum, Rebelo y Trabandt (2020) fusionan el modelo SIR con un modelo macro de agente representativo para hacer un seguimiento de los resultados macroeconómicos. Con relación a esos documentos, la contribución aquí consiste en ilustrar la dependencia de los resultados económicos de algunos parámetros que están muy mal determinados en la literatura actual, y mostrar cómo el modelo puede ser calibrado usando datos de una muestra seleccionada según las pautas de pruebas históricas.

esta manera es que hace uso de la tasa de pruebas positivas, sobre las que hay buenos datos. La tasa de asintomáticos de COVID-19 no está identificada en nuestro modelo y las estimaciones puntuales recientes en la literatura epidemiológica oscilan entre 0,18 y 0,86 (con un rango más amplio si se incorpora la incertidumbre del muestreo). Sin embargo, la tasa de asintomáticos podría estimarse con precisión y rápidamente mediante el análisis de una muestra aleatoria de la población total.

En cuarto lugar, la respuesta política y sus consecuencias económicas dependen del valor de la tasa de asintomáticos. Ilustramos esto con tres líneas de política para β que representan varios niveles de cierre y medidas de seguimiento: dos líneas de "aplanamiento de la curva" y una línea de contención del virus. Estas diferentes vías tendrían consecuencias económicas muy diferentes, aunque no están modeladas. Los resultados sanitarios de las medidas para "aplanar la curva" dependen en gran medida de la tasa de asintomáticos.

Un modelo epidemiológico calibrado sencillo

Bajo los supuestos simplificadores de que la población está homogéneamente mezclada, que los asintomáticos son tan infecciosos como los sintomáticos (tal vez esto no sea cierto, véase Li et. al. (2020)), que en la población todos son igualmente susceptibles al contagio, y que aquellos que han sido infectados previamente ya no son susceptibles, la tasa de infección sigue el llamado modelo SIR (véase, por ejemplo, Allen (2017)). El modelo SIR simple que se utiliza aquí, hace abstracción de la mortalidad. La versión de tiempo discreto del modelo SIR en la escala de tiempo semanal es:

$$(1) \quad \Delta S_t = -\beta I_{t-1} (S_{t-1} / N)$$

$$(2) \quad \Delta R_t = \gamma I_{t-1}$$

$$(3) \quad \Delta I_t = \beta I_{t-1} (S_{t-1} / N) - \gamma I_{t-1}$$

donde S_t es el número de susceptibles, I_t es el número de infectados, R_t es el número de recuperados, y N es la población total (constante). Suponiendo que todos en la población son inicialmente susceptibles, entonces $N = S_t + I_t + R_t$. El coeficiente β es la tasa de transmisión y γ es la tasa de recuperación.

La ecuación (1) dice que el número de nuevos infectados es el número de infectados anteriores multiplicado por la tasa de transmisión por la fracción de la población que es susceptible; el número de susceptibles se reduce uno por uno con el número de nuevos infectados. La Ecuación (2) dice que el número de recuperados aumenta con el número de recuperados en el período actual. La Ecuación (3) dice que el número actual de infecciones aumenta con el número de nuevas infecciones, menos el número de recuperaciones, que se desprende de la identidad $N = S_t + I_t + R_t$.

Un resumen común de la transmisión de la enfermedad es el número básico de reproducción, R_0 , que es $R_0 = \beta/\gamma$. R_0 es el número total de casos producidos por el contagio de un solo caso, cuando toda la población es susceptible y β y γ se encuentran en sus valores de ausencia de políticas.

En este modelo, la política opera manipulando los valores de los parámetros. Los valores de base de β y γ pueden considerarse valores no políticos. La autocuarentena, el distanciamiento social y el cierre de escuelas y empresas actúan para reducir la tasa de transmisión β . Las intervenciones sanitarias, como el tratamiento médico o las drogas (si estuvieran disponibles) podrían servir para aumentar la tasa de recuperación γ . Las políticas que disminuyen β y/o aumentan γ sirven para reducir la R_0 .

Valores de los parámetros e implicancias observables

El modelo tiene dos parámetros desconocidos, β y γ . En el caso del coronavirus, sorprendentemente existen pocos datos para estimar β y γ porque las pruebas han sido limitadas y las que se han hecho se han dirigido en gran medida a los enfermos, especialmente a los enfermos que son más vulnerables o que podrían beneficiarse más de la hospitalización. Es decir, se han hecho pruebas en gran medida de los sintomáticos. Esas guías de pruebas pasan por alto los casos que se denominan asintomáticos o no detectados, que son conceptos diferentes aunque los trato aquí como sinónimos. Un caso puede pasar desapercibido porque la persona no tiene síntomas, porque tiene síntomas suficientemente leves (síntomas de resfriado o alergia) como para no haber informado del caso, o porque informó de sus síntomas a un profesional médico pero no cumplía con las pautas estrictas para recibir la prueba.

La estimación del modelo en un sentido convencional, es decir, el ajuste (1) - (3) utilizando datos de series de tiempo, no es posible porque no hay datos sobre I_t y R_t con los que se pueda ajustar el modelo. La obtención de estimaciones de I_t exige pruebas aleatorias continuas de la población, lo que no ha sucedido. De manera similar, las estimaciones de R_t podrían deducirse de I_t (dado γ), o alternativamente podrían obtenerse mediante un muestreo aleatorio continuo de pruebas de anticuerpos séricos en respuesta al coronavirus, sin embargo tales pruebas no están aun ampliamente disponibles y no han sido desplegadas en muestras aleatorias en Estados Unidos.

La ausencia de pruebas aleatorias de la población plantea un problema adicional. Aunque la tasa de recuperación γ de los enfermos graves puede calcularse a partir de los datos de aquellos cuya progresión de la enfermedad se ha rastreado, no es estimada entre los asintomáticos. La tasa de recuperación difiere plausiblemente entre enfermos sintomáticos y asintomáticos, lo que complica la estimación directa de γ a partir de datos de casos sanitarios.

Aunque no hay datos sobre I_t y R_t , hay datos ampliamente disponibles sobre los resultados de las pruebas (por ejemplo, Roser, Ritchie y Ortiz-Ospina (2020)). Dado

que las pruebas en Estados Unidos se han enfocado en gran medida en los sintomáticos (dejando de lado los pequeños grupos asintomáticos no representativos como los jugadores de la NBA), es plausible que la tasa de pruebas positivas estime la tasa de infección entre los sintomáticos. Utilizando el algoritmo de Bayes el modelo puede ser aumentado aprovechando datos de las series temporales sobre la tasa de pruebas positivas.

La tasa de pruebas positivas puede utilizarse para calibrar el modelo SIR de la siguiente manera. Dividiendo ambos miembros de (1) - (3) por N se expresan todas las cantidades como tasas de población o, a nivel individual, probabilidades. Bajo el supuesto simplificador de que sólo se someten a prueba los sintomáticos, podemos utilizar el principio de Bayes para expresar la tasa de pruebas positivas en términos de la tasa de los sintomáticos (la fracción de las infecciones que son sintomáticas):

$$(4) \quad \Pr [I_t \mid \text{Sintomático}_t] = \frac{\Pr [\text{Sintomático}_t \mid I_t] \Pr [I_t]}{\Pr [\text{Sintomático}_t]} = (1 - \pi_a) \frac{\Pr [I_t]}{\Pr [\text{Sintomático}_t]}$$

donde I_t y S_t se refieren a los infectados y susceptibles como arriba y donde $\pi_a = \Pr [\text{Asintomático}_t \mid I_t] = 1 - \Pr [\text{Sintomático}_t \mid I_t]$ es la tasa de asintomáticos (la tasa de infección no detectada).

La fracción de la población que es sintomática (el denominador en (4)) es,

$$(5) \quad \Pr [\text{Sintomático}_t] = \Pr [\text{Sintomático}_t \mid I_t] \Pr [I_t] + \Pr [\text{Sintomático}_t \mid S_t] \Pr [S_t] + \Pr [\text{Sintomático}_t \mid R_t] \Pr [R_t]$$

$$= (1 - \pi_a) \Pr [I_t] + s_o (\Pr [S_t] + \Pr [R_t]),$$

donde s_o es la tasa de referencia de síntomas entre susceptibles y recuperados (resfriados normales y alergias).

Suponiendo que las pruebas hayan sido aleatorias entre los sintomáticos, la fracción de pruebas que son positivas se estima $\Pr [I_t \mid \text{Sintomático}_t]$. El sistema expandido (1) - (5) tiene cinco ecuaciones y cuatro parámetros: β , γ , π_a , y s_o .

No exploro aquí la estimación del modelo utilizando datos de series de tiempo sobre la tasa de pruebas positivas, aunque eso sería posible. En su lugar, ilustraré su uso y la importancia política del parámetro clave π_a , la tasa de los asintomáticos en una simulación calibrada.

Calibración y simulación del modelo

Paso ahora a una ilustración numérica del modelo y las intervenciones políticas. Para γ , supongo que la vida media de una infección es de 6 días ($\gamma = 0,55$). He esta-

blecido $s_0 = 0,02$. Para la semana del 21 de marzo de 2020, la tasa de pruebas positivas en los Estados Unidos fue de aproximadamente 10% (Roser, Ritchie y Ortiz-Ospina (2020)). Como condiciones iniciales, asumo que había 50 casos (desconocidos) en los EE.UU. en la semana que terminó el 4 de enero, 2020. Incluso con estas calibraciones, el modelo (1) - (5) está subidentificado en un parámetro. Por lo tanto, fijo la tasa asintomática π_a en un valor predeterminado y resolvemos en β de tal manera que la tasa de pruebas positivas sea 10% en la semana que termina el 21 de marzo de 2020.

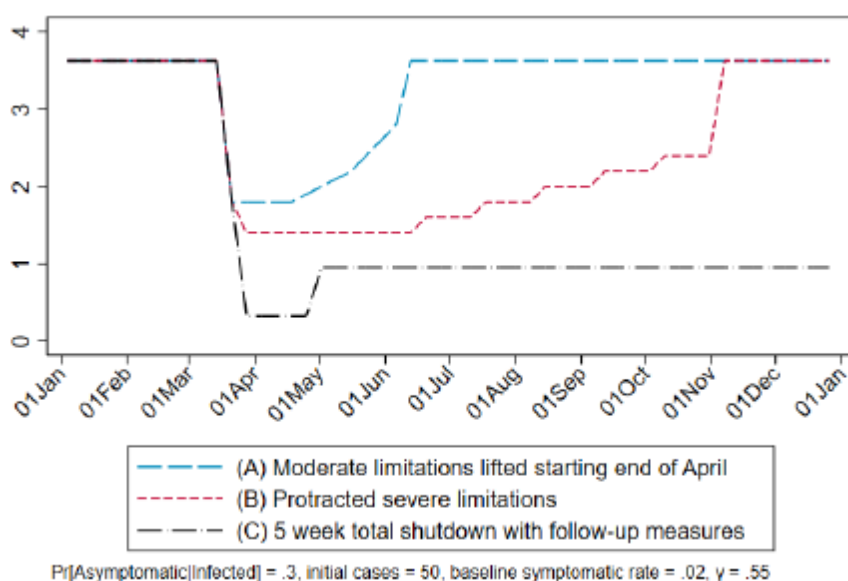
La limitada evidencia disponible sobre la tasa de asintomáticos π_a ha sido revisada en Qui (20 de marzo de 2020). La estimación de Mizumoto y otros (12 de marzo de 2020) de la tasa de asintomáticos sugirió que podría ser tan baja como 18%, sin embargo ese estudio utilizó datos de *Diamond Princess*, que están muy sesgados hacia turistas de edad avanzada y por lo tanto es poco probable que sea representativa. Otras estimaciones son más elevadas, incluido el 31% (Nishiura et al (13 de febrero de 2020)) para 565 ciudadanos japoneses evacuados de Wuhan, sin embargo, se definió como asintomático el hecho de no mostrar ningún síntoma, y no simplemente el hecho de no cumplir con las pautas de pruebas de los EE.UU. Los primeros datos de Islandia sugieren una tasa asintomática de aproximadamente la mitad, sin embargo, en ellos se utiliza una definición estrecha de asintomático; una definición menos estricta, más cercana a las pautas de pruebas de los EE.UU., sugiere una tasa asintomática mucho más elevada.² Li y otros (16 de marzo de 2020) estiman una tasa mucho más alta de 86% en casos no detectados para China. Ninguno de estos estudios es para muestras aleatorias representativas en Estados Unidos. Sobre la base de esta limitada evidencia, adoptamos dos valores de la tasa de asintomáticos, 0,30 (por ejemplo, utilizado en Pueyo (19 de marzo de 2020) y 0,86.

Políticas. La política de cierre opera a través de β . Consideramos tres casos de distanciamiento social/cierre económico. Ninguno de ellos está optimizado y los valores numéricos no deben ser tomados literalmente. Se trata más bien de ilustrar la sensibilidad de los resultados a la tasa de asintomáticos π_a , de manera equivalente, ilustrar cuán diferentes deben ser las trayectorias de β (π_a , en forma semejante, $R_0 = \beta/\gamma$) para lograr una determinada tasa de infección bajo diferentes valores de la tasa de asintomáticos.

² Islandia inició un programa de pruebas voluntarias a gran escala en la segunda semana de marzo. Un informe del 15 de marzo sobre los resultados iniciales dice "De las muestras que se han tomado hasta ahora, 700 han sido probadas. Kári [CEO de deCODE Genetics, la compañía que realiza las pruebas] dice que cerca de la mitad de los que dieron positivo no han mostrado ningún síntoma, y la otra mitad muestra síntomas de tener un resfriado regular". (The Reykjavik Grapevine, 15 de marzo de 2020 en <https://grapevine.is/news/2020/03/15/first-results-of-general-population-screening-about-1-of-icelanders-with-coronavirus/>, véase también <https://www.government.is/news/article/2020/03/15/Large-scale-testing-of-generalpopulation-in-Iceland-underway/>.) Bajo las actuales pautas de prueba de EE.UU., ambas categorías de casos estarían en gran parte no probadas y por lo tanto no serían detectadas.

Las tres líneas de cierre de R_0 se muestran en la Figura 1.³ Las líneas (A) y (B) representan dos estrategias diferentes de "aplanamiento de la curva", en las que la tasa de infección se gestiona de forma que no sobrecargue el sistema sanitario, pero que al final la población consiga la inmunidad de manada. En cambio, la opción (C) tiene por objeto suprimir el virus hasta que se desarrolle una vacuna. Específicamente, la ruta (A) plantea cierres moderados que continúan hasta fines de abril, que luego se levantan lentamente, con levantamiento completo en la última semana de junio. El método (B) plantea cierres más severos que continúan durante tres meses, que luego se levantan lentamente durante los cinco meses siguientes. El sendero (C) muestra lo que es en efecto un cierre total que dura tres meses, con un seguimiento posterior, pruebas continuas, distanciamiento social continuo y seguimiento de los contactos con cuarentena. En el sendero (C), R_0 es 0,32 durante 5 semanas, un valor tomado de la estimación de Wang y otros (6 de marzo de 2020) de R_0 en Wuhan después de que se impusieran medidas de cierre total (estiman un R_0 previo a la política de 3,86).

Figura 1. Tres senderos de R_0 inducidos por la política



En las figuras 2 a 4 se muestran los resultados epidemiológicos de las tres rutas para las políticas (A), (B) y (C) respectivamente. En cada figura, el panel izquierdo muestra las tasas de los actualmente infectados y sintomáticos (en la notación anterior, $Pr [Sintomático_t, I_t]$) y de los siempre infectados (actualmente infectados + recuperados) para el valor bajo de la tasa de asintomáticos (0,3), y el panel derecho muestra estas vías para el valor más alto (0,86).

Los resultados de los dos escenarios de "aplanamiento de la curva" dependen en gran medida de la tasa de asintomáticos (desconocida). La política menos restrictiva (A) aplanar la curva de manera efectiva si la tasa asintomática es alta, y la inmunidad

³ Los valores sin política de R_0 en estas cifras son calculados con $\pi_a = 0,3$.

de manada se logra a mediados del verano. Si la tasa de asintomáticos es baja, la política menos restrictiva da lugar a tasas muy elevadas de sintomatología infectada que abrumarían al sistema de salud. La política de limitaciones severas prolongadas (B) aplanan la curva conforme a la tasa asintomática más baja, aunque la tasa sintomática infectada alcanza un 5% a fines del verano. Aunque no modelizo los costos económicos, es probable que los costos económicos de las vías (A) y (B) sean muy diferentes, y que (B) dé lugar a costos elevados y a una recesión más profunda y prolongada.

Figura 2. Trayectoria de política (A) con tasas de asintomáticos baja (izquierda) y alta (derecha)

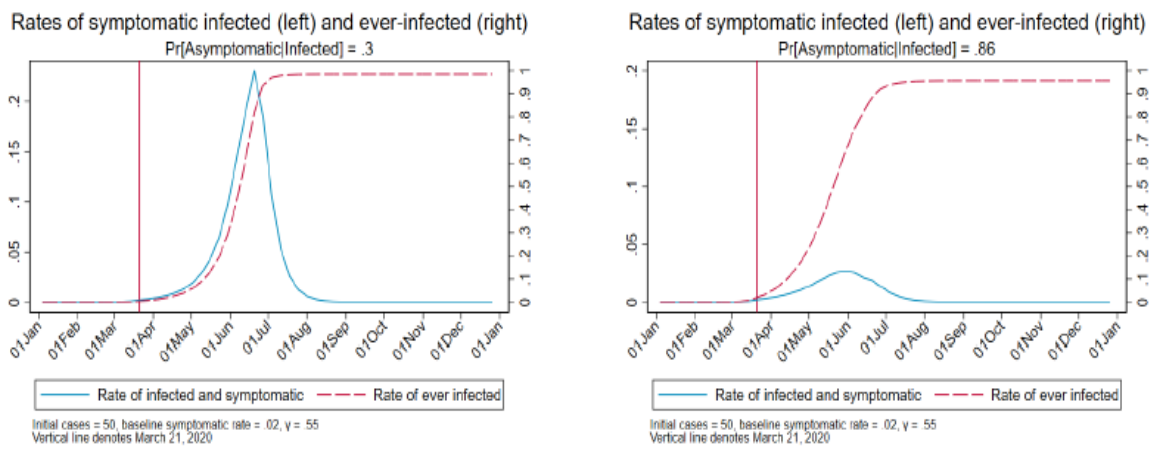
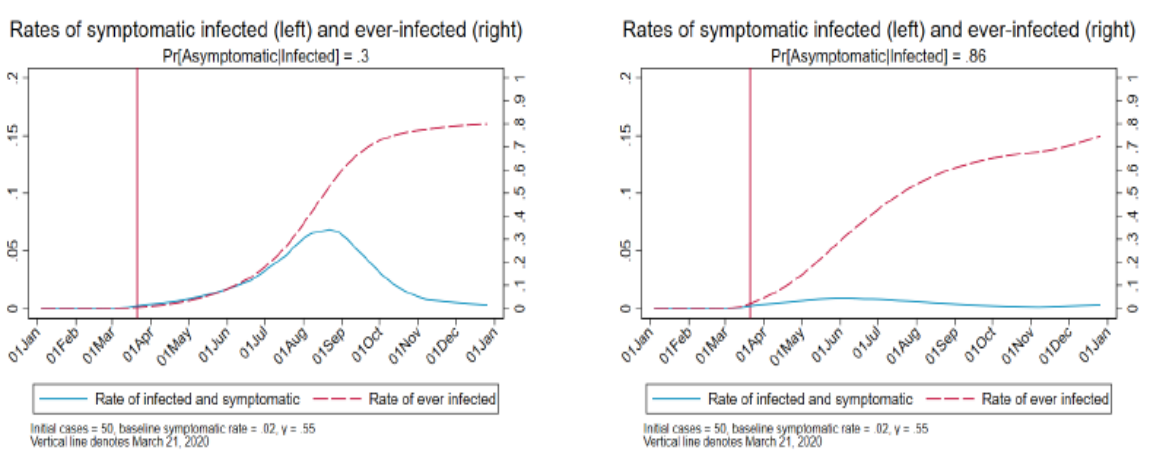


Figura 3. Trayectoria de política (B) con tasas de asintomáticos baja (izquierda) y alta (derecha)

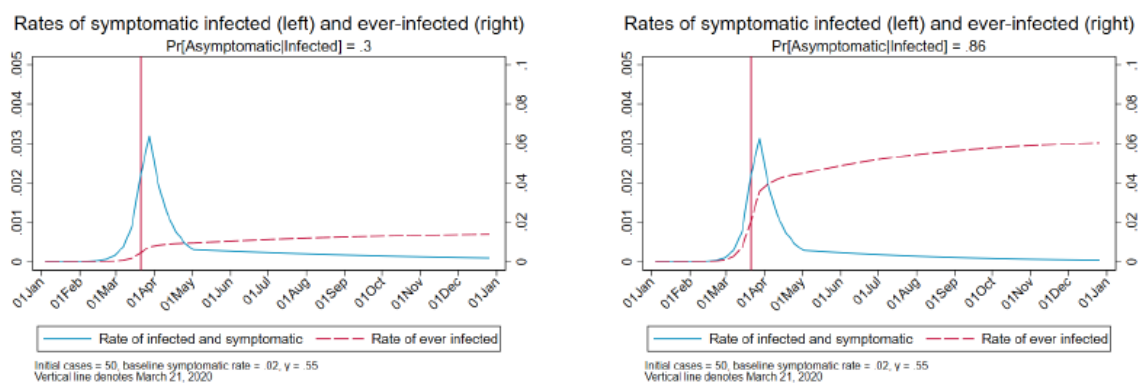


El panel izquierdo de la figura 2 y el panel derecho de la figura 3 representan dos hipótesis de alto costo, la primera costosa en muertes, la segunda costosa en resultados económicos; para evitar cualquiera de las dos es necesario conocer la tasa de asintomáticos y las tasas de infecciones y recuperaciones.

El camino alternativo (C), esencialmente un cierre total de 5 semanas de interacción cara a cara como se hizo en Wuhan, tendría los costos económicos inmediatos más severos. Bajo los parámetros de aquí, esto es suficiente para suprimir el virus. Si se mantiene la R_0 por debajo de 1 hasta que se desarrolle una vacuna, entonces la tasa total de infectados en la población sigue siendo baja.

Figura 4. Trayectoria de política (C) con tasas de asintomáticos baja (izquierda) y alta (derecha)

Nota: Los valores del eje difieren de las Figuras 2 y 3.



Resumen

Los resultados de las políticas dependen fundamentalmente de un parámetro clave desconocido, la fracción de infectados que son asintomáticos, y de las tasas actuales de infectados y recuperados en la población. Las pruebas sobre la tasa de asintomáticos son escasas, pero podrían estimarse fácilmente mediante ensayos randomizados.

Desde la perspectiva del diseño de política económica, este modelo simplificado tiene la virtud de resumir el efecto epidemiológico de las políticas de cierre en un único parámetro, el parámetro de contagio β . En este modelo simple, las diferentes políticas que den lugar al mismo β tendrán los mismos resultados en materia de salud. Sin embargo, políticas distintas pueden tener costos económicos muy diferentes. Por lo tanto, una forma de enmarcar la economía de las políticas de cierre es buscar las políticas más eficientes para lograr un β dado, y luego resolver el camino óptimo de β que equilibra el costo económico con el costo del exceso de vidas perdidas al abrumar el sistema de atención sanitaria.

- Allen, L.J.S. (2017). “A primer on stochastic epidemic models: Formulation, numerical simulation, and analysis.” *Infectious Disease Modeling* 2(2), 128-142.
- Atkeson, A. (2020), “What will be the economic impact of COVID-19 in the US? Rough estimates of disease scenarios,” NBER Working Paper 26867, March 2020.
- Eichenbaum, M., S. Rebelo, and M. Trabant (2020), “The macroeconomics of epidemics,” NBER Working Paper 26882, March 2020.
- Li, R. et. al. (2020), “Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV2),” *Science*, published online March 16, 2020 DOI: 10.1126/science.abb3221 at <https://science.sciencemag.org/content/early/2020/03/13/science.abb3221>.
- Mizumoto, K. et. al., “Estimating the asymptomatic proportion of coronavirus disease 2019 (COVID-19) cases on board the Diamond Princess cruise ship, Yokohama, Japan, 2020,” *Eurosurveillance* 25(10), March 12, 2020 at <https://www.eurosurveillance.org/content/10.2807/1560-7917.ES.2020.25.10.2000180>.
- Nishiura, H. et. al. (2020), “Estimation of the asymptomatic ratio of novel coronavirus infections (COVID-19),” forthcoming, *International Journal of Infectious Disease*, posted online February 13, 2020 at <https://doi.org/10.1016/j.ijid.2020.03.020>.
- Pueyo, T. (2020), “Coronavirus: The Hammer and the Dance,” Medium March 19, 2020 at <https://medium.com/@tomaspueyo/coronavirus-the-hammer-and-the-dance-be9337092b56> Traducción “Coronavirus: El Martillo y la Danza” http://ebour.com.ar/index.php?option=com_weblinks&task=view&id=20363&catid=166

⁴ A las referencias indicadas por el autor, agrego un capítulo que puede consultarse para rever el teorema de Bayes utilizado en el texto: *Elementos de Teoría Estadística*, que he traducido a partir de un curso para abogados de David Kaye y David Freedman: http://ebour.com.ar/ensayos_meyde2/Elementos%20de%20Teoria%20Estadistica.pdf