



Laboratorio
Interdisciplinario de
Estadística Social



Discussion Paper N° 2020|03

**Muchas curvas, misma información:
Sobre la indeterminación del modelo SIR y su
uso en el contexto de la pandemia del COVID-19**

Jorge González & Ernesto San Martín

**Facultad de Matemáticas
Pontificia Universidad Católica de Chile
Av. Vicuña Mackenna 4860, Macul
lies.mat.uc.cl/**

Muchas curvas, misma información:

Sobre la indeterminación del modelo SIR y su uso en el contexto de la pandemia del COVID-19*

JORGE GONZÁLEZ & ERNESTO SAN MARTÍN

*Laboratorio Interdisciplinario de Estadística Social LIES &
Departamento de Estadística, Facultad de Matemáticas
Pontificia Universidad Católica de Chile*

15 de junio de 2020

*Equivocarse es humano,
perseverar en el error es diabólico.
(Canguilhem, 1981)*

1. Introducción

La pandemia del COVID-19 amenaza la salud pública. Las medidas que se toman para acotar dicha amenaza conllevan efectos colaterales negativos: las medidas de cuarentena, así como los cierres de frontera, afectan el intercambio comercial, el turismo, el desarrollo escolar. En países con baja protección social, como el nuestro, dichos efectos llegan a ser catastróficos. Ante esta situación, el hombre de Estado, los hacedores de política pública y el público en general han de preguntarse si las decisiones que se han tomado son *las mejores*, teniendo en cuenta el estado actual de la pandemia.

Una forma de responder a esta pregunta es recurriendo al paradigma de *Políticas Públicas Basadas en Evidencia*, y explorar en qué medida dichas decisiones se han tomado sobre la base de la mejor evidencia disponible (Davies y Boruch, 2001)¹. El recurso a la evidencia científica se justifica por el hecho que una determinada intervención puede ser inconsistente con las expectativas de un hacedor de política pública (Pielke Jr, 2007, Capítulo 5), produciendo incluso impactos no deseados y hasta perversos (Van Woensel, 2020, Capítulo 3). En este contexto, “la evidencia [científica] importa para la elaboración de políticas públicas” (Parkhurst, 2017) ya que su uso riguroso podría evitar potenciales daños, siendo la misma evidencia la que nos dice “lo que funciona”².

Siguiendo este paradigma, el 22 de marzo de 2020³, la Comunidad Científica (a saber, 15 sociedades

* Aunque este artículo contiene notas y apéndices técnicos al final del texto, estos pueden omitirse en una primera lectura sin comprometer la comprensión de las ideas principales.

científicas de Chile y la Academia de Ciencias) le envió una carta al Presidente de la República, al Ministro de Salud y al Ministro de Ciencia y Tecnología para decretar “Cuarentena Nacional Preventiva Obligatoria” en el país por COVID-19, una decisión que de acuerdo a la misma misiva está sustentada por la evidencia científica mundial⁴. Por lo mismo, la misiva se hizo cargo de solicitar al Gobierno el acceso a la necesaria información que permitiría producir la mejor evidencia científica para la toma de decisión⁵. El 24 de marzo, la Mesa Social COVID-19 logra cuatro acuerdos, uno de los cuales corresponde a la creación de la *Mesa de Datos* cuyo objetivo es, en palabras del Ministro de Ciencia Andrés Couve, “disponibilizar datos de la epidemia de la mejor calidad posible para la investigación científica, clínica y para obtener las proyecciones confiables que se deben hacer para anticipar y combatir esta epidemia en nuestro país a través de la implementación de medidas”⁶.

Es así como entró en escena un nuevo actor clave (Rhodes y Lancaster, 2020): los modelos matemáticos predictivos, entre los cuales se encuentran los *modelos compartimentales*. Estos se utilizan para modelar la evolución de una enfermedad infecciosa, una epidemia, en distintos compartimentos que interactúan entre sí. Esta interacción obedece a la Ley de Acción de Masas que, en términos generales, establece que la tasa de transmisión de una infección en una población es proporcional al producto del número de personas en uno de los compartimentos (el de infectadas) y al número de personas en otro compartimnto (el de susceptibles a la infección)⁷.

Uno de estos modelos es el llamado *modelo SIR*, que modela la transmisión de la enfermedad entre tres compartimentos: los susceptibles S , los infectados I y los removidos⁸ R –de ahí el acrónimo *SIR*. La interacción entre estos tres compartimentos depende de dos parámetros: uno que representa la tasa de transmisión de la enfermedad, y otro que representa la tasa de recuperación. Con el paso del tiempo, los compartimentos comienzan a interactuar entre sí de acuerdo a dichos parámetros y es así como individuos que en un principio se encontraban en el grupo de susceptibles, pasan al grupo de infectados; o aquellos que fueron infectados pasarán al grupo de removidos después de un período de tiempo. En términos formales, el modelo SIR se especifica por medio de un sistema de ecuaciones diferenciales (ver el Apéndice A.1), aunque versiones más sencillas también han sido especificadas (Pastén y Castillo-Sepúlveda, 2020).

Las recomendaciones de política pública que se desprenden de este modelo se basan en la tasa de transmisión de la enfermedad: el fenómeno de “aplanar la curva de infectados” depende de dicha tasa. Así, si esta disminuye, entonces se puede mostrar que el número predicho de infectados será menor si se lo compara con una predicción basada en una tasa de transmisión mayor; posponemos las ilustraciones para la Sección 2.1. Pese a lo prometedor e informativo que sean estas predicciones de infectados por COVID-19, el modelo SIR (y en general muchos otros modelos compartimentales) sufren de una “ambigüedad intrínseca” que los hace “nada confiables” como actor clave⁹. Esta ambigüedad se debe a lo que técnicamente se conoce como *falta de identificabilidad*. Un problema de identificabilidad ocurre cuando distintos valores de los parámetros que caracterizan a un modelo (en el caso del modelo SIR, la tasa de transmisión y la tasa de recuperación) producen la misma solución que el modelo entrega como output (en el caso del modelo SIR, la predicción del número de infectados). Cabe hacer notar que este problema ha sido estudiado ampliamente y no sólo está presente en modelos compartimentales, sino que también en variados modelos estadísticos utilizados en investigación empírica (Thurstone, 1935; Haavelmo, 1944; Manski, 1995, 2007; Gundersen y Kreider, 2008; San Martín, González, y Tuerlinckx, 2009;

San Martín, 2016).

En lo que sigue de este artículo, explicaremos brevemente el problema de identificabilidad en el contexto del modelo SIR. Mostraremos cómo medidas epidemiológicas ampliamente utilizadas para monitorear la evolución de la pandemia, tales como el número básico de reproducción, pueden obtenerse a partir de diferentes valores de los parámetros no identificados que caracterizan el modelo SIR. Ilustraremos también cómo el problema de identificabilidad afecta al monitoreo de la pandemia por medio de las curvas de infectados y removidos, ilustrando que dichas curvas pueden obtenerse a partir de distintos valores de los parámetros.

2. El modelo SIR y el problema de identificabilidad

2.1. Ilustrando cómo funciona el modelo

Como hemos mencionado anteriormente, el modelo SIR se especifica como un sistema de ecuaciones diferenciales ordinarias dado que su foco es describir cómo cambian en el tiempo las tasas de susceptibles, infectados y removidos. Estos cambios se deben al hecho que una persona susceptible de ser infectada puede transitar a un estado (compartimento) de infectado a una cierta tasa de transmisión que llamaremos β . Luego, puede transitar a un estado de removido (porque se ha recuperado, o ha fallecido) a una cierta tasa de recuperación que llamaremos γ . Fijadas las tasas de transmisión y recuperación, el modelo SIR proporciona el número de susceptibles, de infectados y de removidos en función del tiempo. Estas funciones corresponden a las soluciones del sistema de ecuaciones ya mencionado, por lo que las llamaremos *solución al modelo SIR*. Para detalles, ver el Apéndice A.1.

2.1.1. Las curvas de susceptibles, infectados y removidos

Asumamos que conocemos la tasa de transmisión y recuperación, y que sus valores son $\beta = 0,001$ y $\gamma = 0,06$. Para un caso hipotético en que el número total de individuos de una población huésped es $N = 1000$, la Figura 1 muestra la evolución de los casos susceptibles, infectados y removidos desde el inicio de la pandemia y proyectando los primeros 100 días. La interpretación de las curvas es bastante intuitiva:

- La curva de susceptibles parte en 1000 al comienzo de la pandemia y va decreciendo a medida que los individuos se van infectando o bien removiéndose con el paso del tiempo.
- La curva de infectados comienza en 1 y va creciendo hasta llegar a un peak, para luego decrecer hasta que finaliza la epidemia.
- La curva de removidos muestra el patrón esperado, comenzando en 0 cuando no hay infectados aún y creciendo hasta que todos los individuos han sido removidos.

Estas curvas corresponden a la solución al modelo SIR.

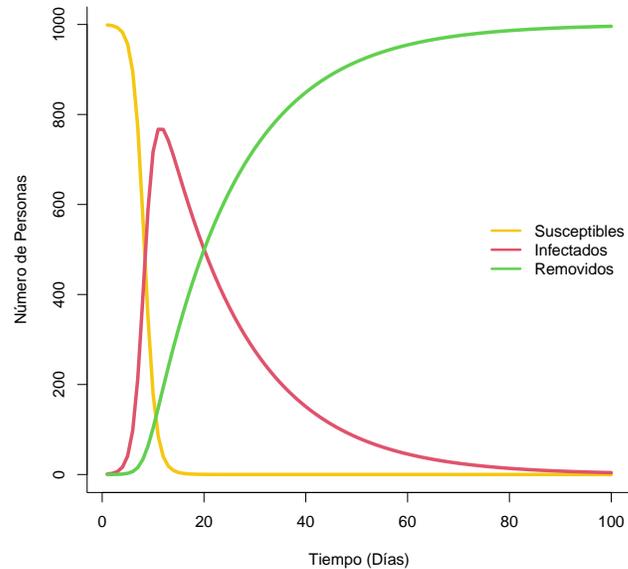


Figura 1: Evolución de la pandemia utilizando el modelo SIR para $\beta = 0,001$ y $\gamma = 0,06$

2.1.2. Aplanar la curva

De estas tres curvas, la más utilizada para recomendaciones de política pública es la curva de infectados. La misma se considera como una predicción confiable que permite no solo saber el número de infectados en un instante dado, sino además predecir el momento en el cual ocurrirá el peak de infectados. Esta información se considera relevante para “tomar decisiones basadas en evidencia” como por ejemplo organizar la logística hospitalaria en lo que dice relación a camas críticas, respiradores mecánicos y formación de personal sanitario para atender hospitalizaciones críticas.

Relacionado con esta curva, dos expresiones recurrentes que escuchamos en los medios de comunicación son las siguientes:

- Se indica que en una primera etapa de la pandemia, la curva de infectados crece de forma exponencial.
- Se habla de la necesidad de *aplanar la curva*.

Con respecto a la primera afirmación, la Figura 2 muestra el comportamiento de la infección en los primeros 8 días en donde efectivamente se observa dicha tendencia “exponencial”. En relación a la segunda afirmación, y de acuerdo al significado que se adscribe a los parámetros del modelo, este objetivo se cumpliría si se disminuye el valor de la tasa β de transmisión de la enfermedad: si se logra bajar dicha tasa, entonces la curva de infectados se aplanará. La Figura 3 ilustra este punto (se omiten las curvas para

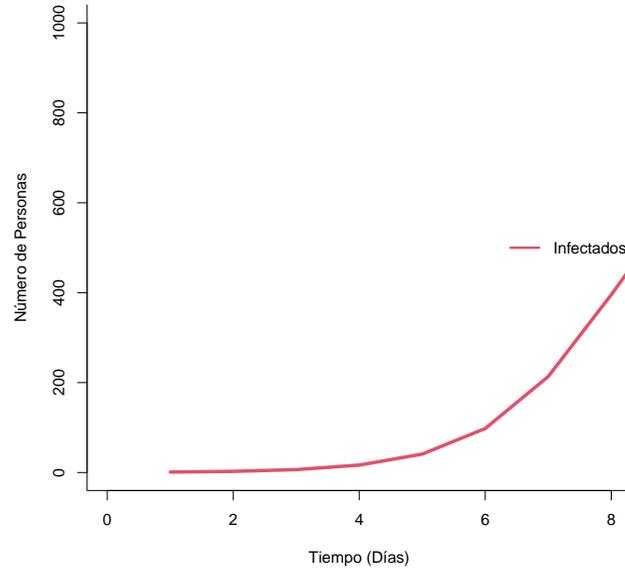


Figura 2: Evolución del número de infectados en la etapa inicial de la pandemia para $\beta = 0,001$ y $\gamma = 0,06$

S y R), en la que se mantienen fijo $N = 1000$ y $\gamma = 0,06$, y solo se disminuye el parámetro β desde un valor 0.001, a valores 0.0005 y 0.0002.

2.1.3. La evolución de una epidemia en función del número efectivo reproductivo

En la Figura 3 se puede efectivamente apreciar que si la tasa de transmisión de la enfermedad disminuye, la curva de infectados *se aplana*. Más aún, el número máximo de infectados es menor cuanto menor es la tasa de contagios β . Estas características se pueden expresar en función de dos parámetros que, por un lado, se definen en función de los parámetros del modelo SIR y, por otro lado, tienen relevancia epidemiológica: se trata del *número básico de reproducción* R_0 y del *número efectivo de reproducción* R_e . Estos se definen como

$$R_0 \doteq \frac{N\beta}{\gamma} \quad \text{y} \quad R_e \doteq R_0 \frac{S(0)}{N},$$

donde $S(0)$ es el total de susceptibles al inicio de la enfermedad (Breban, Vardavas, y Blower, 2007; Ridenhour, Kowalik, y Shay, 2014). Por otro lado, se suele suponer que la población huésped completa es susceptible al inicio de una infección, en cuyo caso $S(0) = N - 1$, $I(0) = 1$ y $R(0) = 0$ –es decir, la infección comienza con un infectado, y el resto es susceptible al contagio. En este caso, si la población huésped es “muy grande”, R_e y R_0 se consideran prácticamente iguales (Weiss, 2013).

El número efectivo de reproducción R_e corresponde al número de nuevos infectados causado por cada

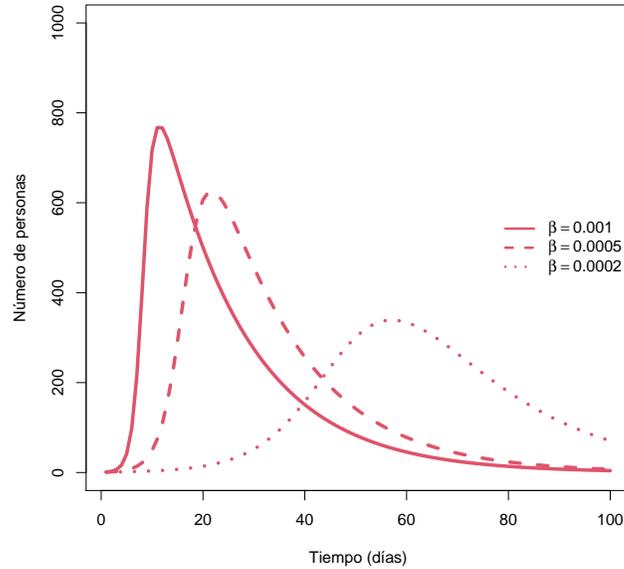


Figura 3: Efecto de la disminución de β en la curva de infectados

Tabla 1: “Aplanar” la curva en relación a R_e

γ	β	R_0	R_e	I_{\max}
0.06	0.001	16.667	16.650	771
0.06	0.0005	8.333	8.325	626
0.06	0.0002	3.333	3.339	339

individuo infectado al inicio del brote. Este parámetro se utiliza para evaluar la gravedad de la infección, así como la fuerza de las intervenciones médicas y/o conductuales necesarias para su control. Si $R_e > 1$ (lo que implica que $R_0 > 1$ pues $R_0 > R_e$), entonces el brote genera una epidemia, lo que significa que el número de infectados crece hasta alcanzar un máximo. Si, por el contrario, $R_e < 1$, el brote cesará “rápidamente”¹⁰; para detalles, ver Apéndice A.2.

Retomando el ejemplo de la Figura 3, y asumiendo que al inicio del contagio toda la población huésped es susceptible, la Tabla 1 muestra que efectivamente cuanto menor es la tasa de contagio β , el máximo de infectados es menor¹¹: el hecho que se alcance un máximo de infectados y que luego dicho número comience a decaer es una consecuencia de que $R_e > 1$.

La Figura 4 ilustra cómo la curva de infectados decrece si $R_e < 1$: manteniendo la misma tasa de recuperación ($\gamma = 0,12$), para $\beta_1 = 0,00009$, $R_e = 0,75$; para $\beta = 0,000045$, $R_e = 0,37$; y para $\beta = 0,000018$, $R_e = 0,15$.

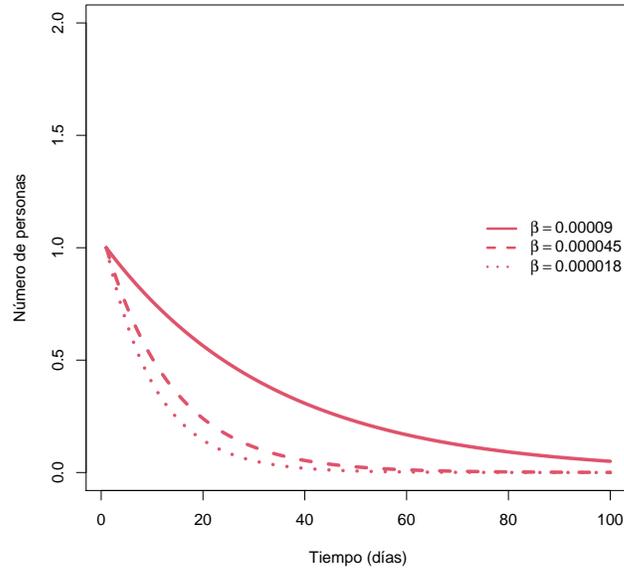


Figura 4: Curva decreciente de infectados

2.2. Consecuencias del problema de identificabilidad

Supongamos que la tasa de transmisión β y de recuperación γ , los parámetros que caracterizan al modelo SIR, no pueden ser determinados de forma única, pero aún así la información que entrega el modelo sobre la evolución de la pandemia es exactamente la misma. Esto quiere decir, por ejemplo, que para tres pares (β_1, γ_1) , (β_2, γ_2) y (β_3, γ_3) de valores distintos de los parámetros, la predicción de infectados que proporciona el modelo será exactamente la misma, o que la estimación del número de básico de reproducción sea exactamente igual. Ante esta situación ¿cuál es el valor de los parámetros sobre el cual basar de manera confiable recomendaciones de política pública? Esta ambigüedad es precisamente lo que llamamos un problema de identificabilidad. Mencionemos que este problema ha sido ampliamente estudiado (Evans, Chapman, Chappell, y Godfrey, 2002a, 2002b; Evans, White, Chapman, Godfrey, y Chappell, 2005; Chapman y Evans, 2008; Cauchemez y Ferguson, 2008; Haderler, 2011; Marinova, Marinova, Omojola, y Jackson, 2014; Tuncer y Le, 2018; Angulo y Velasco-Hernandez, 2018), aunque sus consecuencias sobre las recomendaciones de política pública no han sido suficientemente explotadas.

El problema de identificabilidad de la tasa de transmisión β y de la tasa de recuperación γ se debe al hecho que tanto el número de susceptibles en el tiempo, como el número de infectados y removidos en el tiempo (que constituyen la solución al modelo SIR), dependen de los parámetros β y γ *solo a través del cociente* β/γ ; la demostración de este resultado se puede ver en el Apéndice B.2¹². Este es el resultado principal sobre el cual se basa nuestra calificación de que el modelo SIR no es confiable para hacer recomendaciones de política pública.

Para ilustrar esta afirmación, consideremos los siguientes tres pares de parámetros del modelo SIR:

$$(\beta_1, \gamma_1) = (0,001; 0,06), \quad (\beta_2, \gamma_2) = (0,0005; 0,03), \quad (\beta_3, \gamma_3) = (0,0002; 0,012),$$

los que han sido elegidos de modo que $(\beta_2, \gamma_2) = (c_2\beta_1, c_2\gamma_1)$ con $c_2 = 0,5$ y $(\beta_3, \gamma_3) = (c_3\beta_1, c_3\gamma_1)$ con $c_3 = 0,2$. Esta elección tiene dos consecuencias importantes:

Primera consecuencia: estos tres pares de parámetros inducen el *mismo* cuociente

$$0,0167 = \frac{\beta_1}{\gamma_1} = \frac{\beta_2}{\gamma_2} = \frac{\beta_3}{\gamma_3},$$

y, por tanto, el *mismo* número básico de reproducción $R_0 = 16,67$, como el *mismo* número efectivo de reproducción $R_e = 16,68$. Si el mensaje sobre la gravedad de la epidemia se construye a partir de las tasas de transmisión de la enfermedad, entonces hay tres mensajes distintos: la curva de infectados es más plana cuando se asume que la tasa de transmisión de la enfermedad es β_3 que cuando se asume que dicha tasa es β_2 . Y, similarmente, que la curva de infectados es más plana cuando se asume que la tasa de transmisión es β_2 que cuando se asume que es β_1 . Pero ocurre que las respectivas curvas de infectados son “igual de planas”, como puede apreciarse en la Figura 5: aquí las curvas de infectados son paralelas y por tanto *no hay mejora de la epidemia si las misma se describe a partir de las tasas de transmisión y recuperación*. Es más, para los tres parámetros (β_1, γ_1) , (β_2, γ_2) y (β_3, γ_3) , el número efectivo de reproducción R_e es mayor que 1, por lo que la tasa de infectados crece alcanzando el mismo máximo para las tres curvas de la Figura 5. Así, lo que podría resultar coherente del modelo SIR con el comportamiento observado de los infectados son el número básico y el número efectivo de reproducción *pues ellos son función del cuociente β/γ* .

Segunda consecuencia: esta última afirmación nos lleva a discutir la segunda consecuencia debida a la falta de identificabilidad de las tasas de transmisión y recuperación del modelo SIR. Sabemos que para los distintos valores de los parámetros (β_1, γ_1) , (β_2, γ_2) y (β_3, γ_3) , la solución al modelo SIR es la misma. Por lo tanto, la pregunta es cómo debería especificarse el modelo SIR en función de cada uno de estos parámetros de modo que se obtenga dicha solución. La respuesta es *escalando el tiempo de acuerdo a la constante c_2 , o a la constante c_3* , según corresponda; para detalles, ver el Apéndice B.2. La Figura 6 ilustra este fenómeno:

- La Figura 6a presenta las curvas de susceptibles, infectados y recuperados cuando el modelo SIR depende de (β_1, γ_1) .
- La Figura 6b presenta las curvas de susceptibles, infectados y recuperados cuando el modelos SIR depende de (β_2, γ_2) : estas curvas son paralelas a las obtenidas en la Figura 6a. El tiempo de esta segunda solución se obtiene dividiendo por $c_2 = 0,5$ el tiempo de la primera solución. De ahí que el tiempo para esta segunda solución llegue a lo 200 días.
- La Figura 6c presenta las curvas de susceptibles, infectados y recuperados cuando el modelos SIR depende de (β_3, γ_3) : estas curvas son paralelas a las obtenidas en la Figura 6a. El tiempo de esta tercera solución se obtiene dividiendo por $c_3 = 0,2$ el tiempo de la primera solución. De ahí que el tiempo de esta tercera solución llegue a los 500 días.

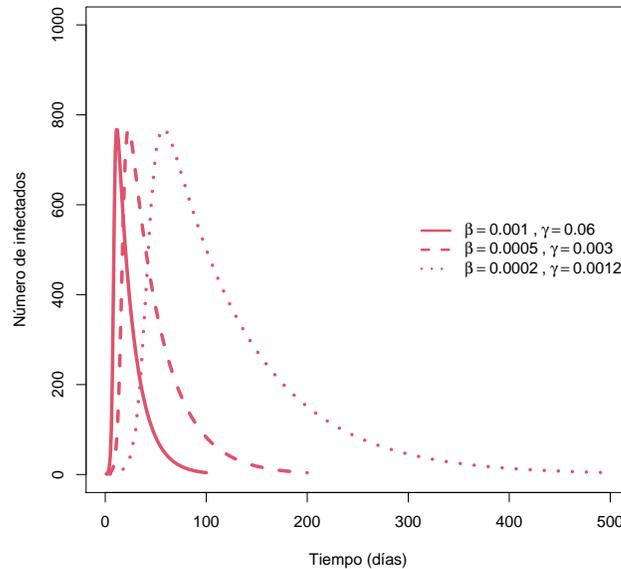
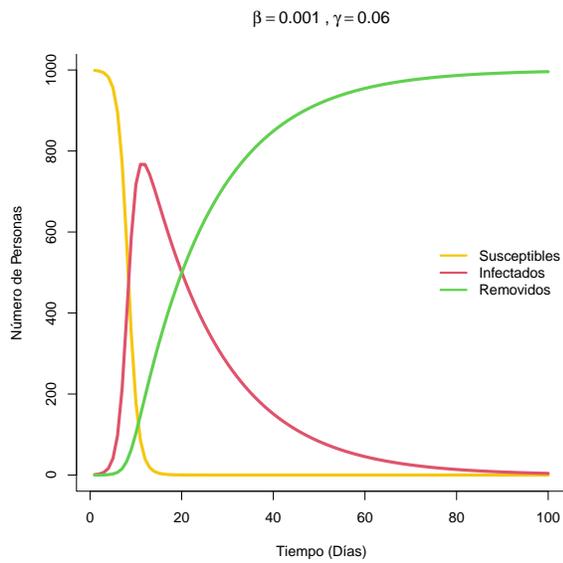


Figura 5: Curvas equivalentes de infectados para distintas tasas de transmisión y recuperación

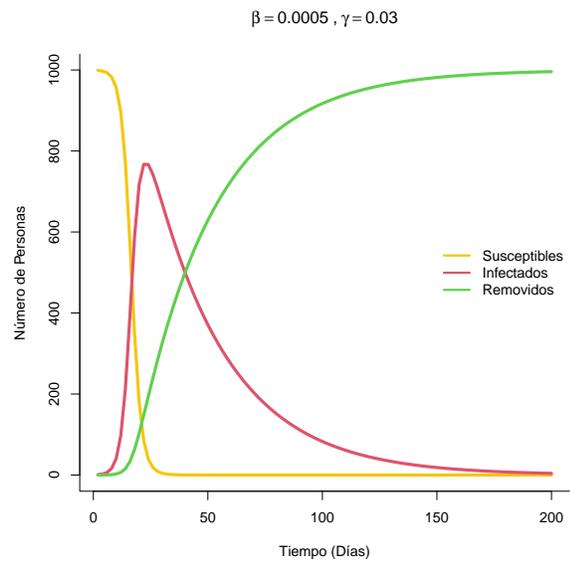
Si las recomendaciones de política pública se hacen en función de la curva de infectados, la cual es considerada como una predicción confiable, entonces las recomendaciones son totalmente ambiguas. Por ejemplo, el máximo de infectados es igual a 767 personas ¿Cuándo se alcanza este máximo? Si se considera la primera solución al modelo SIR, este máximo se alcanza a los 11 días. Pero si se considera la segunda solución al modelo SIR, este máximo se alcanza a los 22 días; y si se considera la tercera solución, este máximo se alcanza a los 55 días –todo esto, suponiendo que el origen de la epidemia está en $t_0 = 0$. El problema de identificabilidad no permite predecir de manera única cuándo se alcanzará dicho máximo. Más aún, si el modelo se usa, habrá una infinidad de días en los cuales se puede alcanzar el máximo de infectados, lo que produciría un *mensaje estresante para la población*.

2.3. El desafío de explicar los fenómenos observados

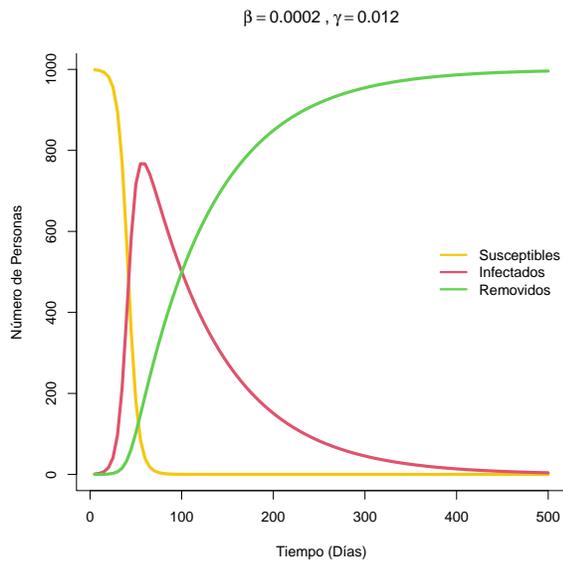
Los problemas de identificabilidad están relacionados con un aspecto fundamental de la investigación empírica: para poder extraer conclusiones de *un fenómeno observado* es necesario combinarlos con una teoría sustantiva, ya sea económica, psicológica, sociológica o, como en el caso que nos ocupa, epidemiológica. Una cosa es *asumir cómo funciona un fenómeno*, lo que se puede formalizar por medio de un modelo matemático o estadístico; otra cosa distinta es *observar cómo funciona el fenómeno*, lo cual redundaría en recolectar datos y desarrollar representaciones de los mismos. Un problema de identificabilidad significa, por tanto, que dos formulaciones diferentes de una teoría conducen al mismo “proceso generador de observaciones”, de modo que no podemos distinguir entre ellas sobre la base de las obser-



(a) Primera solución al modelo SIR



(b) Segunda solución al modelo SIR



(c) Tercera solución al modelo SIR

Figura 6: Curvas paralelas de susceptibles, infectados y removidos

vaciones, incluso cuando dichas teorías pueden no ser equivalentes en otros aspectos. Como se expresaba Haavelmo (1944), “no es un problema de pura lógica, sino un problema de *saber realmente algo* sobre los fenómenos y de hacer supuestos realistas sobre ellos” (p.29); ver también Haavelmo (1997).

El modelo SIR corresponde a una teoría que pretende representar cómo evoluciona una epidemia en un sistema cerrado. Las bases de esta teoría son, por un lado, la Ley de Acción de Masas y, por otro lado, la *hipótesis de Snow* según la cual las epidemias terminan cuando se quedan sin “combustible” (Snow, 1853), lo que significa que el brote termina cuanto menos disponibilidad de susceptibles hay; para más detalles, ver Apéndice A.3.2. Sobre estos dos supuestos generales, el modelo SIR describe la evolución de una epidemia en función de la tasa de infección y de la tasa de recuperación: se trata de dos parámetros que tienen un significado *en el modelo SIR*, de modo que si la tasa de infección decrece, el brote infeccioso disminuye, “aplanándose” la curva de infectados. Pero dado que estos parámetros no son identificados, dicho significado *en el modelo* no es suficiente para afirmar que los mismos existen *en relación al fenómeno observado*, lo que implica que dichos parámetros no representan el funcionamiento observado de la epidemia.

Las relaciones puramente empíricas, cuando son discernibles, probablemente se deban a la presencia y persistencia de una (suerte) de estructura subyacente; y, de ser así, las mismas podrían deducirse del conocimiento de esta última. Sin embargo, la dirección de esta deducción no puede invertirse –desde las relaciones empíricas hasta la estructura, como si los datos “revelasen” una estructura– excepto posiblemente con la ayuda de una teoría que especifique la forma de dicha estructura, las variables que entran en juego y cualquier otro detalle apoyado por la observación previa o la deducción de las mismas (Koopmans, 1949; Manski, 2011)¹³. Dado que los parámetros del modelo SIR no son identificados, la evolución observada de la epidemia no permite inferir que el modelo SIR subyace a las observaciones: en términos prácticos, no es posible monitorear la evolución de la pandemia usando las curvas de susceptibles, infectados y removidos que se desprenden como solución al modelo SIR. Es importante mencionar que la identificabilidad de otros modelos compartimentales, o bien aún es un problema abierto (D’Silva y Eisenberg, 2018), o algunos resultados aún son materia de debate (Roosa y Chowell, 2019).

3. Reflexiones finales

Las consecuencias de la pandemia son devastadoras. Es por ello que se requieren acciones de política pública que ayuden a controlar en algún grado tanta devastación. Hemos sido testigos de la imprudencia inicial de la clase política en Estados Unidos, Reino Unido, México, Brasil. No cabe duda que cada nación busca herramientas para que el hombre de Estado y los tomadores de decisiones de políticas públicas sean más asertivos según lo permitan las circunstancias, y lo menos imprudentes frente al mayor bien común: la vida de las personas. Es por ello que se acude a la ciencia como sostén y justificación de acciones políticas.

En este respecto, de la falta de identificabilidad del modelo SIR podemos extraer algunas lecciones relativas a la prudencia y responsabilidad de los científicos a la hora de participar en el espacio público (Abbasi, 2020):

1. Los hacedores de políticas públicas requieren oír de los científicos los supuestos que eventualmente permiten a los modelos propuestos proporcionar algún conocimiento del fenómeno observado. No basta que los modelos sean *ampliamente usados*; es necesario transparentar aquellos supuestos que no son falseables empíricamente y que todo modelo asume.

2. Extensiones y complejizaciones de modelos compartimentales deben desarrollarse con extrema cautela: debe asegurarse su identificabilidad para que puedan ser usados de manera responsable y confiable en política pública.
3. La necesaria colaboración interdisciplinaria: los análisis de identificabilidad, así como la pertinencia de los modelos matemáticos para las políticas públicas, requieren el protagonismo de las teorías sustantivas, lo que exige generar y practicar una comunicación idiomática amplia y profunda¹⁴.

Tal vez estemos frente a uno de los límites del paradigma de *Políticas Públicas Basadas en Evidencia*, por lo que uno de los roles de la ciencia en política pública sea la de ofrecer herramientas que ayuden a monitorear el impacto de las acciones políticas sobre la evolución de la pandemia –y de otros eventuales fenómenos sociales.

A. El Modelo SIR

Para evaluar los méritos de un modelo matemático es importante explicitar los supuestos sobre los cuales se construye el mismo. De esta manera, no solo resulta transparente cómo intentar generalizar o complejizar dicho modelo, sino también explicitar las justificaciones sustantivas (por ejemplo, epidemiológicas) del modelo.

A.1. Supuestos básicos y especificación del modelo SIR

Como lo hace notar, entre muchos otros, Capasso (1983), uno de los modelos más básicos que describen la evolución de una pandemia en una población huésped cerrada es el llamado *modelo SIR*, introducido por Kermack y McKendrick (1927). Los supuestos básicos de este modelo son los siguientes:

1. **División de la población huésped.** La población huésped está compuesta de tres clases de individuos:
 - La clase de los susceptibles S : el conjunto de individuos capaces de contraer la enfermedad y convertirse por tanto en infecciosos.
 - La clase de los infectados I : el conjunto de individuos capaces de transmitir la enfermedad a los susceptibles.
 - La clase de los removidos R : el conjunto de individuos que, habiendo contraído la enfermedad, han muerto o, si se han recuperado, están permanentemente inmunes, o han sido aislados, siendo así incapaces de seguir transmitiendo la enfermedad.

Como se ha mencionado en la sección introductoria, el acrónimo *SIR* se debe a este primer supuesto.

2. *Ley de acción de masas.* El proceso de infección, que corresponde al proceso de transferencia de S a I , es descrito por esta ley. Además, la transferencia de infectados I a recuperados R se considera en términos de un decaimiento exponencial.

En consecuencia, el modelo SIR se especifica de acuerdo al siguiente sistema de ecuaciones diferenciales ordinarias (que llamaremos *representación EDO del modelo SIR*):

$$\frac{d}{dt}S(t) = -\beta S(t)I(t) \quad (\text{A.1})$$

$$\frac{d}{dt}I(t) = \beta S(t)I(t) - \gamma I(t) \quad (\text{A.2})$$

$$\frac{d}{dt}R(t) = \gamma I(t), \quad (\text{A.3})$$

donde $S(0) = N_1 \geq 0$, $I(0) = N_2 \geq 0$ y $R(0) = N_3 \geq 0$ y $N_1 + N_2 + N_3 = N$. El parámetro $\beta \in (0, 1)$ corresponde a la tasa o probabilidad de transmisión de la enfermedad. Esta última interpretación se puede

justificar de la siguiente manera: supongamos que un individuo tiene κ contactos por unidad de tiempo; supongamos además que este κ es independiente del tamaño de la población. Entonces $\kappa S(t)/N$ de estos contactos se producen con individuos susceptibles. Si la fracción τ de contactos adecuados resulta en transmisión de la enfermedad, entonces cada individuo contagia $\kappa\tau S(t)/N$ individuos susceptibles por unidad de tiempo. Así, $\beta = b/N$ donde $b = \kappa\tau$. El parámetro $\gamma \in (0, 1)$ corresponde a la tasa de recuperación o, equivalentemente, a la duración $D = 1/\gamma$ de la infección.

Se puede notar que, de acuerdo a la ley de acción de masas, el número de nuevos infectados por unidad de tiempo, $\beta S(t)I(t)$, corresponde a una mezcla homogénea de las clases de infectados y susceptibles. Por otro lado, el total N de la población permanece constante, lo que se deduce del modelo SIR pues

$$\frac{d}{dt}\{S(t) + I(t) + R(t)\} = 0$$

y por tanto $S(t) + I(t) + R(t) = S(t_0) + I(t_0) + R(t_0) = N_1 + N_2 + N_3$.

A.2. Propiedades generales del modelo SIR

Dos parámetros de interés en Salud Pública son los siguientes el *número básico de reproducción* R_0 y el *número efectivo de reproducción* definidos en la Sección 2.1.3. Ahora bien, R_e puede expresarse de manera equivalente como

$$R_e = \frac{S(0)\beta}{\gamma} = \frac{S(0)b}{\gamma N} = \frac{D\kappa\tau S(0)}{N}.$$

Por tanto, R_e es el producto de la duración de la infección, el número de contactos que un infectado tiene con individuos susceptibles por unidad de tiempo y la razón de transmisión. Por lo tanto, R_e corresponde al número de nuevos infectados causado por cada individuo infectado al inicio de la epidemia. En otras palabras, R_e es una medida de la “capacidad del patógeno” (*fitness of the pathogen*).

El parámetro R_e permite interpretar en términos de Salud Pública la evolución de la epidemia. Para ello, es necesario notar en primer lugar que la función $S(t)$ es una función decreciente dado que su derivada es negativa (ver ecuación (A.1)), mientras que $R(t)$ es una función creciente dado que su derivada es positiva (ver ecuación (A.3)). Por lo tanto, $0 \leq S(t) \leq S(0) \leq N$ y $0 \leq R(0) \leq R(t) \leq N$, lo que a su vez implica que $S(\infty) = \lim_{t \rightarrow \infty} S(t)$ y $R(\infty) = \lim_{t \rightarrow \infty} R(t)$ existen. Esto permite deducir del modelo SIR las siguientes propiedades (Weiss, 2013):

1. Si $R_e \leq 1$, entonces $I(t)$ es una función que decrece a 0 a medida que t crece.
2. Si $R_e > 1$, entonces, a medida que t crece, $I(t)$ crece, alcanza su máximo y luego decrece a 0.
3. Para todo $t \geq 0$, se tiene que

$$I(t) + S(t) - \frac{\gamma}{\beta} \ln S(t) = I(0) + S(0) - \frac{\gamma}{\beta} \ln S(0). \quad (\text{A.4})$$

El número máximo de infectados I_{\max} se alcanza cuando $dI(t)/dt = 0$, lo que es equivalente, por la ecuación (A.2), a $S(t) = \frac{\gamma}{\beta}$. Usando (A.4) se deduce que

$$I_{\max} = I(0) + S(0) - \frac{\gamma}{\beta} \{1 + \ln(R_e)\}.$$

Si al inicio de una epidemia, toda la población es susceptible, entonces $S(0) = N - 1$, $I(0) = 1$ y $R(0) = 0$. En este caso,

$$\frac{I_{\max}}{N} = 1 - \frac{1}{R_0} \{1 + \ln(R_e)\} \geq 1 - \frac{1}{R_0} \{1 + \ln(R_0)\}. \quad (\text{A.5})$$

En la práctica, si N es suficientemente grande, $R_e \approx R_0$, por lo que la desigualdad anterior puede considerarse como una igualdad.

A.3. Algunas observaciones acerca de los supuestos básicos del Modelo SIR

A.3.1. Los compartimentos poblacionales

El primer supuesto del modelo SIR corresponde a una partición de la población en tres clases de individuos: los susceptibles, los infectados y los recuperados. Esta caracterización de la población es un supuesto que se puede justificar epidemiológicamente. Más aún, consideraciones de orden epidemiológico pueden contribuir a afinar dicha partición. Así, por ejemplo, la población huésped se ha dividido en cuatro clases: los susceptibles S , los expuestos pero aún no infectados E , los infectados I , y los removidos R . Esto ha dado lugar a los *modelos SEIR* y sus casos especiales (Liu, Hethcote, y Levin, 1987, en particular la Tabla 1).

Otros han añadido a la partición inicial S , I y R , la clase de asintomáticos A , la clase de los que están en cuarentena Q , la de los aislados J , obteniendo así el *modelo SAQIJR* (Brauer, 2006). Otros, siempre sobre consideraciones epidemiológicas, consideran que algunos miembros de la población infectada nunca desarrollan síntomas, sino que pasan directamente de la etapa latente a una etapa infecciosa asintomática, y luego a la etapa de removidos. Esto da lugar al *modelo SLIAR* (Arino, Brauer, van den Driessche, Watmough, y Wu, 2006) y otras variantes como la desarrollada por Arino y Portet (2020).

Es importante enfatizar que a medida que se afina la partición de la población huésped, se incorporan nuevos parámetros al modelo, lo que plantea el problema de identificación de los parámetros.

A.3.2. La ley de acción de masas en epidemiología

El modelo SIR y sus “variaciones sobre el mismo tema” se basan en la ley de acción de masas. Este principio establece que la tasa de transmisión de una infección en una población (representada por la derivada de la función $I(t)$ en la ecuación (A.2)) es proporcional al producto del número de personas infectadas I y el número de personas susceptibles a la infección S . Esta ley fue inicialmente establecida, en base a experimentos, en el campo de las reacciones cinéticas químicas, en 1864, por los noruegos Cato Maximilian Guldberg y Peter Waage; dicha ley establece que para una reacción química reversible,

en equilibrio a una temperatura constante, debe existir una relación constante entre concentraciones de reactivos y productos.

Los orígenes de la teoría epidemiológica moderna se encuentran en los trabajos de Hamer (1906), Ross (1908) y Kermack y McKendrick (1927). Como lo ha señalado (Heesterbeek, 2005), la manera en que Hamer estudia las curvas de epidemia (i.e., curvas de prevalencia/incidencia en función del tiempo) sugiere el uso de dicha ley. En el siglo XIX la evolución de una epidemia se explicaba esencialmente por medio de dos “teorías epidemiológicas”: la *hipótesis de Farr* y la *hipótesis de Snell*. En términos generales, la hipótesis de Farr establece que las epidemias terminan porque la potencia del “organismo” causal disminuye con cada individuo que atraviesa (Farr, 1866); una versión completa del texto de Farr se puede encontrar en Brownlee (1915). Por otro lado, la hipótesis de Snow asevera que las epidemias terminan porque se quedan sin “combustible”, es decir, la menor disponibilidad de susceptibles hace que el brote termine (Snow, 1853); para una reedición del texto de Snow, ver Snow (1985).

Hamer estaba convencido de la veracidad de la hipótesis de Snow. En (1906), Hamer tiene por objetivo construir una curva epidemiológica basada sobre varios supuestos de transmisión y comparar la curva resultante con los datos del sarampión de Londres. Suponiendo en particular que la disponibilidad de susceptibles es el único mecanismo en juego (la hipótesis de Snow), deduce el número de casos infectados en un período de tiempo determinado por medio de la siguiente proporcionalidad:

$$\frac{\text{Casos en el próximo intervalo de tiempo}}{\text{Casos actuales}} = \frac{\text{Susceptibles actuales}}{\text{Susceptibles en que cada uno infecta solo a uno}}.$$

Como hace notar Heesterbeek (2005), la igualdad anterior puede expresarse de la siguiente manera:

$$\text{Casos en el próximo intervalo de tiempo} \propto \text{Casos actuales} \times \text{Susceptibles actuales}.$$

Aquí, una constante de proporcionalidad es característica de la infección y de la comunidad y se toma como el inverso del tamaño de la población susceptible en estado estable cuando una infección da lugar a exactamente un nuevo caso; o, como lo dice Hamer, el inverso de los “susceptibles en que cada uno infecta solo a uno”. La igualdad anterior puede relacionarse con el modelo SIR (Heesterbeek, 2005), aunque no por ello hay que asumir que Hamer considerara explícitamente la ley de acción de masas alcanzar su objetivo.

Sin embargo, quien explícitamente introdujo esta ley en los modelos epidemiológicos fue McKendrick (1911), lo que desarrolló ampliamente en especial en Kermack y McKendrick (1927). La gran mayoría de las variaciones del modelo SIR asumen sin más esta perspectiva que, de acuerdo a Heesterbeek (2005), era una suerte de “moda” dado el enorme prestigio que dicha ley alcanzó en Química. Sin embargo, dicha ley fue cuestionada por Wilson y Worcester (1945): Wilson y Worcester enunciaba esta ley diciendo que si la tasa de nuevos infectados es C , entonces la ley de acción de masas se escribiría como

$$C = rIS, \tag{A.6}$$

donde r es una constante (Wilson y Worcester, 1945, p. 24). Wilson y Worcester explican que esta ley se basa sobre el supuesto, explícito o implícito, que los infecciosos están mezclados uniformemente entre los susceptibles en toda la población, cuestión que ya podemos rastrear en Hamer (1906), como ya lo

hemos mencionado. De acuerdo a la ley de acción de masas, “si tuviéramos una población con el doble de susceptibles e infecciosos y con la misma tasa r de mezcla, la tasa C a la que pasaría la infección no sería el doble sino cuatro veces mayor”. Para Wilson y Worcester, es poco verosímil que tal condición exista, por lo que la real utilidad de este supuesto para describir la evolución de la enfermedad se obtiene después de examinar si con una elección adecuada de la constante r , la ecuación (A.6) proporciona una curva teórica de nuevo casos que esté “de acuerdo satisfactoriamente con la curva observada de nuevos casos”. Desde una perspectiva estadística, diríamos que hay que ajustar el modelo subyacente a (A.6) a fin de estimar la constante r .

B. Falta de identificabilidad del modelo SIR

B.1. Solución exacta del modelo SIR

En esta sección desarrollamos los detalles de la falta de identificabilidad de los parámetros β y γ del modelo SIR usando un argumento de identificación global, que en cierto modo contrasta con los argumentos de identificación local desarrollados recientemente (Evans y cols., 2002a, 2002b, 2005; Chapman y Evans, 2008; Cauchemez y Ferguson, 2008; Hadeler, 2011; Marinova y cols., 2014; Tuncer y Le, 2018; Angulo y Velasco-Hernandez, 2018). Nuestro argumento¹ se basa en la solución exacta al modelo SIR desarrollada por Harko, Lobo, y Mak (2014)¹⁵.

En efecto, la solución exacta está parametrizada por una función $u(t)$ de forma que

$$S^{(\beta, \gamma)}(t) = S_0 u(t) \quad (\text{B.1})$$

$$I^{(\beta, \gamma)}(t) = \frac{\gamma}{\beta} \ln\{u(t)\} - S_0 u(t) + N \quad (\text{B.2})$$

$$R^{(\beta, \gamma)}(t) = -\frac{\gamma}{\beta} \ln\{u(t)\}, \quad (\text{B.3})$$

donde

$$u_0 := u(0) = \exp\left\{-\frac{\beta N_3}{\gamma}\right\}, \quad S_0 = N_1 \exp\left\{\frac{\beta}{\gamma} N_3\right\}$$

En las igualdades anteriores hemos explicitado la dependencia de las funciones $S(t)$, $I(t)$ y $R(t)$ con respecto a los parámetros $(\beta, \gamma) \in (0, 1)^2$. La relación entre la función $u(t)$ y el tiempo t está dada por la siguiente representación integral del tiempo:

$$t - t_0 = \int_{u_0}^{u(t)} \frac{d\xi}{\xi[-N\beta - \gamma \ln \xi + S_0 \beta \xi]}, \quad (\text{B.4})$$

donde t_0 es un tiempo inicial arbitrario que comúnmente se toma igual a 0.

Usando esta solución exacta, Harko y cols. (2014) muestran cómo, para modelos explícitos correspondientes a valores fijos de los parámetros β y γ , la solución analítica reproduce exactamente la solución

¹De acuerdo a los autores de este trabajo, dicho análisis es inédito, por lo que una versión corta en inglés ha sido enviada para que sea considerada su publicación.

numérica basada en integración numérica (Harko y cols., 2014, p.188). En términos computacionales, el procedimiento aplicado es como sigue:

1. Se fija un valor numérico $u_1 > u_0$.
2. Luego se obtiene $S_1 = S(u_1)$ y $t_1 = t(u_1)$ usando (B.4), lo que requiere integración numérica.
3. Luego se escoge $u_2 > u_1$ y se obtiene $S_2 = S(u_2)$ y $t_2 = t(u_2)$.
4. Y así sucesivamente de forma de obtener $\{(t_1, x_1), (t_2, x_2), \dots\}$.

B.2. Análisis de identificación del modelo SIR

Ahora bien, la solución exacta depende de los parámetros (β, γ) solo a través de $\frac{\gamma}{\beta}$. En efecto, sea $c \in (0, 1)$, entonces

$$\begin{aligned}
S^{(\beta, \gamma)}(t) &= N_1 \exp\left\{-\frac{\beta N_3}{\gamma} u(t)\right\} = N_1 \exp\left\{-\frac{c \beta N_3}{c \gamma} u(t)\right\} \\
&= N_1 \exp\left\{-\frac{\tilde{\beta} N_3}{\tilde{\gamma}} u(t)\right\} = S^{(\tilde{\beta}, \tilde{\gamma})}(t); \\
I^{(\beta, \gamma)}(t) &= \frac{\gamma}{\beta} \ln\{u(t)\} - S_0 u(t) + N = \frac{c \gamma}{c \beta} \ln\{u(t)\} - S_0 u(t) + N \\
&= \frac{\tilde{\gamma}}{\tilde{\beta}} \ln\{u(t)\} - S_0 u(t) + N = I^{(\tilde{\beta}, \tilde{\gamma})}(t); \\
R^{(\beta, \gamma)}(t) &= -\frac{\gamma}{\beta} \ln\{u(t)\} = -\frac{c \gamma}{c \beta} \ln\{u(t)\} \\
&= -\frac{\tilde{\gamma}}{\tilde{\beta}} \ln\{u(t)\} = R^{(\tilde{\beta}, \tilde{\gamma})}(t),
\end{aligned}$$

donde

$$\tilde{\beta} := c \beta, \quad \tilde{\gamma} := c \gamma. \tag{B.5}$$

Es decir, para $(\beta, \gamma) \neq (\tilde{\beta}, \tilde{\gamma})$, se tiene que

$$\begin{aligned}
\{S^{(\beta, \gamma)}(t) : t \geq t_0\} &= \{S^{(\tilde{\beta}, \tilde{\gamma})}(t) : t \geq t_0\}, \\
\{I^{(\beta, \gamma)}(t) : t \geq t_0\} &= \{I^{(\tilde{\beta}, \tilde{\gamma})}(t) : t \geq t_0\}, \\
\{R^{(\beta, \gamma)}(t) : t \geq t_0\} &= \{R^{(\tilde{\beta}, \tilde{\gamma})}(t) : t \geq t_0\}.
\end{aligned} \tag{B.6}$$

En otras palabras, los parámetros (β, γ) no son identificados o, equivalentemente, (β, γ) y $(\tilde{\beta}, \tilde{\gamma})$ son *observacionalmente equivalentes* (Rothenberg, 1971, Definición 1) o *indistinguibles* (Evans y cols., 2005,

Definición 1); para una “traducción” entre la terminología econométrica y la teoría de sistemas, ver Kalman (1983, en particular, Tabla 1). Más aún, las soluciones exactas nos permiten afirmar que la *parametrización identificada* está dada por $\frac{\gamma}{\beta}$, lo que significa que los *parámetros de interés* (β, γ) no son identificados; para detalles en torno a la distinción entre *parametrización identificada* y *parametrización de interés*, ver San Martín y cols. (2009).

B.3. Parametrización identificada y representación EDO del modelo SIR

Para poder ilustrar gráficamente la falta de identificabilidad de los parámetros del modelo SIR, es necesario reescribir las representación EDO (A.1)-(A.3) en términos de $(\tilde{\beta}, \tilde{\gamma})$. Dado que la representación EDO depende, por definición, del tiempo (se trata de un sistema de ecuaciones diferenciales), es necesario observar cómo debe variar el tiempo t de forma de comparar las soluciones numéricas del modelo parametrizado por (β, γ) y por $(\tilde{\beta}, \tilde{\gamma})$. Para ello, basta con usar la representación integral (B.4) de t . En efecto,

$$\begin{aligned} t - t_0 &= \int_{u_0}^u \frac{d\xi}{\xi[-N\beta - \gamma \ln(\xi) + S_0\beta\xi]} \\ &= \int_{u_0}^u \frac{d\xi}{\xi \left[\frac{-N\tilde{\beta}}{c} - \frac{\tilde{\gamma}}{c} \ln(\xi) + S_0\frac{\tilde{\beta}}{c}\xi \right]} \\ &= \int_{u_0}^u \frac{d\xi}{\frac{\xi}{c}[-N\tilde{\beta} - \tilde{\gamma} \ln(\xi) + S_0\tilde{\beta}\xi]}, \end{aligned}$$

lo que es equivalente a

$$\frac{t - t_0}{c} = \int_{u_0}^u \frac{d\xi}{\xi[-N\tilde{\beta} - \tilde{\gamma} \ln(\xi) + S_0\tilde{\beta}\xi]} \quad (\text{B.7})$$

Por lo tanto, la representación EDO del modelo SIR debe reescribirse en términos del tiempo $(t - t_0)/c$, con $c \in (0, 1)$, en el siguiente sentido: sea

$$L_{c,t_0}(t) = \frac{t - t_0}{c}, \quad t \geq t_0$$

Entonces

$$\begin{aligned} \frac{d}{dt} \left[S^{(\beta,\gamma)} \circ L_{c,t_0} \right] (t) &= \frac{d}{dt} \left\{ S^{(\beta,\gamma)} (L_{c,t_0}(t)) \right\} \\ &= \frac{d}{dt} S^{(\beta,\gamma)} (L_{c,t_0}(t)) \frac{1}{c} \end{aligned}$$

and by (A.1)

$$\frac{d}{dt} S^{(\beta,\gamma)} \left(\frac{t - t_0}{c} \right) = -\beta S^{(\beta,\gamma)} \left(\frac{t - t_0}{c} \right) I^{(\beta,\gamma)} \left(\frac{t - t_0}{c} \right).$$

Por lo tanto,

$$\begin{aligned}
\frac{d}{dt} S^{(\beta, \gamma)}(L_{c, t_0}(t)) &= -\tilde{\beta} S^{(\beta, \gamma)}(L_{c, t_0}(t)) I^{(\beta, \gamma)}(L_{c, t_0}(t)) \\
&= -\tilde{\beta} S^{(\tilde{\beta}, \tilde{\gamma})}(L_{c, t_0}(t)) I^{(\tilde{\beta}, \tilde{\gamma})}(L_{c, t_0}(t)) \text{ por (B.6)} \\
&= \frac{d}{dt} S^{(\tilde{\beta}, \tilde{\gamma})}(L_{c, t_0}(t)) \text{ por la representación EDO.}
\end{aligned}$$

De manera similar, se concluye que

$$\frac{d}{dt} I^{(\beta, \gamma)}\left(\frac{t-t_0}{c}\right) = \tilde{\beta} S^{(\tilde{\beta}, \tilde{\gamma})}\left(\frac{t-t_0}{c}\right) I^{(\tilde{\beta}, \tilde{\gamma})}\left(\frac{t-t_0}{c}\right) - \tilde{\gamma} I^{(\tilde{\beta}, \tilde{\gamma})}\left(\frac{t-t_0}{c}\right)$$

y

$$\frac{d}{dt} R^{(\beta, \gamma)}\left(\frac{t-t_0}{c}\right) = \tilde{\gamma} I^{(\tilde{\beta}, \tilde{\gamma})}\left(\frac{t-t_0}{c}\right).$$

Dado que c es arbitrario, se concluye que, a causa de la no-identificabilidad de (β, γ) , existe una representación EDO equivalente del modelo SIR:

$$\begin{aligned}
\frac{d}{dt} S^{(\beta, \gamma)}(t) &= -\tilde{\beta} S^{(\tilde{\beta}, \tilde{\gamma})}(t) I^{(\tilde{\beta}, \tilde{\gamma})}(t); \\
\frac{d}{dt} I^{(\beta, \gamma)}(t) &= \tilde{\beta} S^{(\tilde{\beta}, \tilde{\gamma})}(t) I^{(\tilde{\beta}, \tilde{\gamma}, \tilde{C}_1)}(t) - \tilde{\gamma} I^{(\tilde{\beta}, \tilde{\gamma})}(t); \\
\frac{d}{dt} R^{(\beta, \gamma)}(t) &= \tilde{\gamma} I^{(\tilde{\beta}, \tilde{\gamma})}(t).
\end{aligned}$$

Agradecimientos: Le damos las gracias al Profesor Tiberiu Harko del Department of Mathematics, University College London, por mantener discusiones en torno a Harko y cols. (2014).

Notas

¹De acuerdo a Davies y Boruch, *Las decisiones de política pública deben basarse en la mejor evidencia disponible*. En el original, la expresión reza así: *Policy decisions should be based on the better available evidence*. En inglés existen dos palabras, *politics* y *policy*, que en español suelen traducirse por un mismo término, *política*. Sin embargo, *politics* dice relación con las negociaciones y discusiones que permiten alcanzar fines deseados de acuerdo a ciertos valores políticos, mientras que *policy* es una decisión, un compromiso con un curso de acción particular; para detalles, ver Cairney (2016) y Pielke Jr (2007, Capítulos 1 y 2). Para mantener esta diferenciación, traducimos *policy* por *política pública*

²*What works* como expresión técnica

³De acuerdo al Diario *El Mostrador*, ese día se envió la misiva; ver <https://www.elmostrador.cl/destacado/2020/03/22/sociedades-cientificas-exigen-a-pinera-que-declare-cuarentena-nacional-obligatoria-para-aplanar-curva-de-propagacion-del-covid-19/>.

⁴“La evidencia científica a nivel mundial es clara, el aislamiento total de ciudades y del país, junto al distanciamiento social dentro de las ciudades, es la mejor forma de evitar los contagios y las muertes, ya que las personas no están respetando el aislamiento. Es la manera de no tener que enfrentarnos a escenarios como los que viven Estados Unidos, Italia y España por su inacción preventiva. Países vecinos como Argentina, Bolivia, Colombia, con menos casos que Chile, ya han tomado medidas drásticas de cuarentena para resguardar a la población. Si bien es claro que esta decisión tiene un gran impacto en la economía, afectando vidas humanas”; recuperado de https://docs.google.com/forms/d/161VPYjSIHG6hmMCwLOOVjTepG_Gsl4gQTh7TkDOSgeU/viewform?edit_requested=true el 4 de junio de 2020.

⁵“Asimismo, se necesita cuanto antes, el 100 % de la transparencia de los datos respecto de los casos detectados, positivos, negativos, pendientes de confirmación, cifras importantes para *tomar las decisiones oportunas para el país, que deben estar basadas en la evidencia científica*. Los firmantes, estamos comprometidos con Chile y solicitamos que *se evalúen, fundamenten y tomen las mejores decisiones para proteger la salud de la población, entendiendo que estas decisiones son dinámicas en base a la mejor información disponible, la evidencia científica y de manera transparente de cara a la ciudadanía*. ¡La vida es lo primero!”; los subrayados son nuestros. Se puede apreciar cómo el *motto* del paradigma *Políticas Públicas Basadas en Evidencia* resuena en esta petición.

⁶Los subrayados son nuestros; recuperado de <http://www.minciencia.gob.cl/noticias/ministro-de-ciencia-encabezara-mesa-de-datos-sobre-coronavirus-en-chile> el 4 de junio de 2020. Ver también la cuenta pública del Ministro de Ciencia realizada el 29 de abril de 2020 en http://www.minciencia.gob.cl/sites/default/files/200429_cuenta_publica_participativa.pdf. Es importante destacar que esta declaración está en total consonancia con la cartera que encabeza el Ministro Couve: de acuerdo a la Ley 21.105, que creó el Ministerio de Ciencia, Tecnología, Conocimiento e Innovación, el conocimiento y la innovación científica no solo proveen la evidencia oportuna para *formular* políticas públicas, sino también para sustentar su *implementación* (Art. 3). Así, la Ley 21.105 no solo le atribuye a la ciencia el rol “estándar” de elaboración de políticas pública, su seguimiento, medición y evaluación (Art. 5k), sino que además explicita el rol que tiene en el campo de la acción: el desarrollo nacional se construye a partir de la acción (implementación de políticas), por lo que es necesario fijar lineamientos generales del desarrollo científico, que en el caso chileno privilegia explícitamente la asociatividad, matizando la primacía investigación por mera curiosidad (Art. 4b).

⁷Para detalles de la Ley de Acción de Masas, su origen y críticas, ver Apéndice A.3.2.

⁸Usamos la designación *removidos* en lugar de *recuperados*, pues esta puede dar lugar a equívocos. De hecho, entre los removidos se contabilizan los recuperados y los fallecidos. Este tipo de equívocos ha sido polémico en nuestro país: la noticia de que Chile estaba contabilizando a los fallecidos por COVID-19 como “recuperados” fue publicada por medios internacionales, lo que pudo haber gatillado una precisión en el lenguaje por parte del Ministerio de Salud; para detalles, ver <https://www.t13.cl/noticia/tendencias/bbc/modelos-matematicos-de-coronavirus-muertos-recuperados-2020>.

⁹La expresión “nada confiable” ha sido usada en matemáticas para referirse a ciertas perspectivas, aún en boga, pero no por ello libres de permanente análisis crítico. Brouwer (1908) usa la expresión holandesa *onbetrouwbaarheid* –nada confiable– para criticar el uso del principio de tercer excluido, y de la lógica en general, en matemáticas por el hecho que “en la ciencia la lógica es a menudo exacta pero no se puede confiar que ella lo sea cuando se itera su aplicación de forma indefinida” y “en matemáticas no es seguro que toda la lógica sea lícita y no es seguro que se pueda resolver el problema de si toda la lógica es lícita”. Para detalles del trabajo de Brouwer, ver Van Atten, Sundholm, Bourdeau, y Van Atten (2014). Recurrimos al término *onbetrouwbaarheid* porque la crítica de Brouwer al principio de tercer excluido tiene directa aplicación en proto-política (ética), tal y como lo ha desarrollado Lorenzen (1987a) y Lorenzen (1987b).

¹⁰Es importante señalar que en este trabajo estamos usando la especificación del modelo SIR tal que $\dot{S}(t) = -\beta S(t)I(t)$, y no la especificación $\dot{S}(t) = -\frac{b}{N}S(t)I(t)$. Para ambas especificaciones las definiciones de R_0 y R_e son las mismas.

¹¹El máximo de infectados se calcula usando la expresión (A.5)

¹²La demostración que desarrollamos, que hasta donde sabemos es inédita, corresponde a lo que se llama *identificación global*, en contraste con las demostraciones que se han hecho, que se reducen a una identificación local.

¹³Este tipo de disquisiciones es de larga data. Un ejemplo ilustrativo lo encontramos en la carta que el Cardenal Bellarmino le escribe al Padre Foscarini el 12 de abril de 1615 a propósito del copernicanismo. En el tercer argumento le dice así: *Digo que si hubiera una verdadera demostración de que Sol está ubicado en el centro del mundo y la Tierra en el tercer cielo, y que el Sol no gira en torno a la Tierra, sino que la Tierra gira en torno al Sol, entonces sería necesario ir con mucho cuidado al explicar las Escrituras que parecen contrarias [...] Y si me dice que Salomón habla según las apariencias, que es que a nosotros nos parece que el Sol gira, mientras que en realidad es la Tierra la que gira, de la misma forma que a quien se aleja de la playa le parece que la playa se aleja de la nave, responderé que quien se leja de la playa sabe muy bien que le parece que la playa se aleja de él, no obstante que eso es un error y lo corrige, viendo claramente que es la nave la que se mueve y no la playa; pero en cuanto al Sol y la Tierra, no hay ningún experto que tenga necesidad de corregir el error, porque claramente experimenta que la Tierra está inmóvil y que el ojo no se engaña cuando juzga que el Sol se mueve, como también no se engaña cuando juzga que la Luna y las estrellas se mueven* (González, 1987, Apéndice II). Para desechar una u otra “impresión”, es necesario contar con alguna información adicional, es decir, el problema de identificación se podría resolver enriqueciendo la teoría. Para más discusión, ver Manski (1993).

¹⁴En matemáticas, un ejemplo paradigmático es la colaboración que el matemático N. Wiener desarrolló con el fisiólogo Arturo Rosenblueth, la cual describe en los siguientes términos: *Si un fisiólogo que no sabe de matemáticas trabaja junto con un matemático que no sabe de fisiología, el uno será incapaz de plantear su problema en términos que el otro pueda manipular, y el segundo será incapaz de poner las respuestas en cualquier forma que el primero pueda entender [puntos] Dr. Rosenblueth siempre ha insistido en que una exploración adecuada de estos espacios en blanco en el mapa de la ciencia sólo podría ser realizada por un equipo de científicos, cada uno especialista en su propio campo pero cada uno poseyendo un conocimiento profundo y entrenado en los campos de sus vecinos; todos con el hábito de trabajar juntos, de conocer las costumbres intelectuales de los demás y de reconocer el significado de la nueva sugerencia de un colega antes de que ésta haya tomado una expresión formal completa. El matemático no necesita tener la habilidad de realizar un experimento fisiológico, pero sí la de comprenderlo, criticarlo y sugerirlo. El fisiólogo no necesita ser capaz de probar un cierto teorema matemático, pero debe ser capaz de comprender su significado fisiológico y decirle al matemático lo que debe* (Wiener, 1948).

¹⁵Los autores del presente trabajo han desarrollado una demostración “más directa” a fin de encontrar dicha solución exacta, la cual está en la versión en inglés de la presente contribución.

Referencias

- Abbasi, K. (2020). *Covid-19: Questions of conscience and duty for scientific advisers*. British Medical Journal Publishing Group.
- Angulo, M., y Velasco-Hernandez, J. (2018). Robust qualitative estimation of time-varying contact rates in uncertain epidemics. *Epidemics*, 24, 98–104.

- Arino, J., Brauer, F., van den Driessche, P., Watmough, J., y Wu, J. (2006). Simple models for containment of a pandemic. *Journal of the Royal Society Interface*, 3(8), 453–457.
- Arino, J., y Portet, S. (2020). A simple model for COVID-19. *Infectious Disease Modelling*, 5, 309-315.
- Brauer, F. (2006). Some simple epidemic models. *Mathematical Biosciences & Engineering*, 3(1), 1.
- Breban, R., Vardavas, R., y Blower, S. (2007). Theory versus Data: How to Calculate R0? *PLoS One*, 2(3).
- Brouwer, L. E. J. (1908). De onbetrouwbaarheid der logische principes. *Tijdschrift voor Wijsbegeerte*, 2, 152-168.
- Brownlee, J. (1915). Historical note on Farr’s theory of the epidemic. *British Medical Journal*, 2(2850), 250.
- Cairney, P. (2016). *The Politics of Evidence-Based Policy Making*. Palgrave Macmillan.
- Canguilhem, G. (1981). *Idéologie et Rationalité dans l’Histoire des Sciences de la Vie: Nouvelles Études d’Histoire et de Philosophie des Sciences*. Vrin.
- Capasso, V. (1983). Mathematical Structures of Epidemic Systems. *Lecture Notes in Biomathematics*, 97.
- Cauchemez, S., y Ferguson, N. (2008). Likelihood-based estimation of continuous-time epidemic models from time-series data: application to measles transmission in london. *Journal of the Royal Society Interface*, 5(25), 885–897.
- Chapman, J., y Evans, N. (2008). The structural identifiability of sir type epidemic models with incomplete immunity and birth targeted vaccination. *IFAC Proceedings Volumes*, 41(2), 9075–9080.
- Davies, P., y Boruch, R. (2001). The Campbell Collaboration: Does for Public Policy what Cochrane Does for Health. *British Medical Journal Publishing Group*, 323, 294-295.
- D’Silva, J. P., y Eisenberg, M. C. (2018). Identifiability of linear compartmental models of infectious disease transmission.
- Evans, N., Chapman, M., Chappell, M., y Godfrey, K. (2002a). Identifiability of uncontrolled nonlinear rational systems. *Automatica*, 38(10), 1799–1805.
- Evans, N., Chapman, M., Chappell, M., y Godfrey, K. (2002b). The structural identifiability of a general epidemic (SIR) model with seasonal forcing. *IFAC Proceedings Volumes*, 35(1), 109–114.
- Evans, N., White, L., Chapman, M., Godfrey, K., y Chappell, M. (2005). The structural identifiability of the susceptible infected recovered model with seasonal forcing. *Mathematical Biosciences*, 194(2), 175–197.
- Farr, W. (1866). On the cattle plague. *Journal of Social Science*, March 20th.
- González, M. (1987). *Galileo Galilei. Carta a Cristina de Lorena y otros textos sobre ciencia y religión*. Alianza Editorial.
- Gundersen, C., y Kreider, B. (2008). Food Stamps and Food Insecurity: What Can Be Learned in the Presence of Nonclassical Measurement Error. *The Journal of Human Resources*, 43, 352–382.
- Haavelmo, T. (1944). The Probability Approach in Econometrics. *Econometrica*, 12, iii–115.
- Haavelmo, T. (1997). Econometrics and the Welfare State. *The American Economic Review*, 87(6), 13–15.
- Hadeler, K. (2011). Parameter identification in epidemic models. *Mathematical Biosciences*, 229(2), 185–189.

- Hamer, W. H. (1906). Epidemic disease in England –the evidence of variability and of persistency of type. *Lancet*, *ii*, 733–739.
- Harko, T., Lobo, F. S. N., y Mak, M. K. (2014). Exact analytical solutions of the Susceptible-Infected-Recovered (SIR) epidemic model and of the SIR model with equal death and birth rates. *Applied Mathematics and Computation*, *236*, 184–194.
- Heesterbeek, H. (2005). 5 - The Law of Mass-Action in Epidemiology: A Historical Perspective. En K. Cuddington y B. E. Beisner (Eds.), *Ecological paradigms lost* (p. 81 - 105). Burlington: Academic Press.
- Kalman, R.-E. (1983). Identifiability and modeling in econometrics. En P. R. Krishnaiah (Ed.), *Developments in Statistics. Volumen 4* (pp. 97–136). Elsevier.
- Kermack, W. O., y McKendrick, A. G. (1927). Contributions to the Mathematical Theory of Epidemics –I. *Proceedings of the Royal Society*, *115A*, 700–721.
- Koopmans, T. C. (1949). Identification Problems in Economic Model Construction. *Econometrica*, *17*, 125–144.
- Liu, W., Hethcote, H., y Levin, S. (1987). Dynamical Behavior of Epidemiological Models with Nonlinear Incidence Rates. *Journal of Mathematical Biology*, *25*(4), 359–380.
- Lorenzen, P. (1987a). First Lecture: Practice Based on Theory as the Object of Science. *Synthese*, *71*, 127–139.
- Lorenzen, P. (1987b). Second Lecture: Political Practice and Ethico-Political Theory. *Synthese*, *71*, 140–155.
- Manski, C. F. (1993). Identification of Endogenous Social Effects: The Reflection Problem. *The Review of Economic Studies*, *60*(3), 531–542.
- Manski, C. F. (1995). *Identification Problems in the Social Sciences*. Cambridge: Harvard University Press.
- Manski, C. F. (2007). *Identification for Prediction and Decision*. Harvard University Press.
- Manski, C. F. (2011). Policy Analysis with Incredible Certitude. *The Economic Journal*, *121*(554), F261–F289.
- Marinova, T., Marinova, R. S., Omojolaa, J., y Jackson, M. (2014). Inverse problem for coefficient identification in SIR epidemic models. *Computers and Mathematics with Applications*, *67*, 2218–2227.
- McKendrick, A. G. (1911). The chemical dynamics of serum reactions. *Proceedings of the Royal Society of London. Series B, Containing Papers of a Biological Character*, *83*(567), 493–512.
- Parkhurst, J. (2017). *The Politics of Evidence: From Evidence-Based Policy to the Good Governance of Evidence*. Routledge.
- Pastén, H., y Castillo-Sepúlveda, J. (2020). Evolución de las epidemias: la matemática de aislarse. *Extraído de <http://www.mat.uc.cl/hector.pasten/preprints/Epidemia2020.pdf>*.
- Pielke Jr, R. A. (2007). *The Honest Broker: Making Sense of Science in Policy and Politics*. Cambridge University Press.
- Rhodes, T., y Lancaster, K. (2020). Mathematical models as public troubles in COVID-19 infection control: following the numbers. *Health Sociology Review*, 1–18. doi: 10.1080/14461242.2020.1764376

- Ridenhour, B., Kowalik, J. M., y Shay, D. K. (2014). Unraveling R0: Considerations for Public Health Applications. *American Journal of Public Health*, 104(2), e32-e41.
- Roosa, K., y Chowell, G. (2019). Assessing Parameter Identifiability in Compartmental Dynamic Models Using a Computational Approach: Application to Infectious Disease Transmission Models. *Theoretical Biology and Medical Modelling*, 16(1), 1.
- Ross, R. (1908). *Report on the Prevention of Malaria in Mauritius*. Waterlow.
- Rothenberg, T. J. (1971). Identification in parametric models. *Econometrica*, 39, 577–591.
- San Martín, E. (2016). Identification of Item Response Theory Models. En W. van der Linden (Ed.), *Handbook of Item Response Theory: Models, Statistical Tools, and Applications. Volumen 2* (pp. 127–150).
- San Martín, E., González, J., y Tuerlinckx, F. (2009). Identified parameters, parameters of interest and their relationships. *Measurement: Interdisciplinary Research & Perspective*, 7, 97–105.
- Snow, J. (1853). On continuous molecular changes, more particularly in their relation to epidemic diseases. *London: Churchill*.
- Snow, J. (1885). On continuous molecular changes, more particularly in their relation to epidemic diseases. *Reviews of Infectious Diseases*, 7, 441–447.
- Thurstone, L. L. (1935). *The Vectors of Mind: Multiple-Factor Analysis for the Isolation of Primary Traits*. University of Chicago Press.
- Tuncer, N., y Le, T. (2018). Structural and practical identifiability analysis of outbreak models. *Mathematical Biosciences*, 299, 1–18.
- Van Atten, M., Sundholm, G., Bourdeau, M., y Van Atten, V. (2014). “Que les principes de la logique ne sont pas fiables”: Nouvelle traduction française annotée et commentée de l’article de 1908 de L.E.J. Brouwer. *Revue d’Histoire des Sciences*, 67(2), 257–281.
- Van Woensel, L. (2020). *A Bias Radar for Responsible Policy-Making: Foresight-Based Scientific Advice*. Palgrave Macmillan, Springer Nature.
- Weiss, H. H. (2013). The SIR Model and the Foundations of Public Health. *Materials Mathematics*, 0001–17.
- Wiener, N. (1948). *Cybernetics or Control and Communication in the Animal and the Machine*. The M.I.T. Press.
- Wilson, E. B., y Worcester, J. (1945). The law of mass action in epidemiology. *Proceedings of the National Academy of Sciences of the United States of America*, 31(1), 24.