

# Modelos logísticos de crecimiento aplicados a la COVID-19 en España

Víctor Martín Barroso

## Nº 5

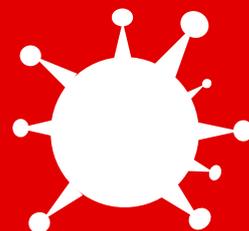
27-04-2020

Documentos de trabajo COVID-19

Centro de Estudios Economía de Madrid



Universidad  
Rey Juan Carlos



# Modelos logísticos de crecimiento aplicados a la COVID-19 en España

Víctor Martín Barroso

Los primeros casos de neumonía atípica causados por un nuevo tipo de coronavirus, fueron confirmados en la ciudad de Wuhan (provincia de Hubei, China), el 31 de diciembre de 2019 (WHO, 2020a). Esta enfermedad infecciosa, denominada como COVID-19, se ha expandido rápidamente a un gran número de países, obligando a la Organización Mundial de la Salud a declararla como pandemia mundial durante la segunda semana del mes de marzo (WHO, 2020b). A 22 de abril de 2020, el número total de casos confirmados a nivel mundial superaba los 2,5 millones, de los cuales prácticamente la mitad corresponden a la Unión Europea (WHO, 2002,c). España, precedida tan solo por los EEUU, ocupa un segundo puesto con más de 200 mil casos confirmados.

Dentro de España, la prevalencia de la COVID-19 varía de forma considerable entre comunidades autónomas. Así, la Comunidad de Madrid (60.847 confirmados, 22 de abril) y Cataluña (44.892, 22 de abril) son las comunidades con un mayor número de casos confirmados. Entre ambas comunidades acumulan prácticamente la mitad de los casos confirmados en el conjunto de España. A estas le siguen Castilla La-Mancha (17.557 confirmados, 22 de abril) y Castilla y León (17.402 confirmados, 22 de abril), las cuales acumulan el 8.3% y el 8.2% del total de casos respectivamente. En el resto de comunidades, con la salvedad de País Vasco (13.436 confirmados, 6,3%), Andalucía (11.845 confirmados, 5,6%) y Comunidad Valenciana (10.692 confirmados, 5%), el número de casos confirmados, en cada una de ellas, representa menos del 5% del conjunto total de España.

El objetivo del presente documento es modelizar la evolución de la COVID-19 en España tanto a nivel agregado como por comunidades autónomas, haciendo uso de los denominados modelos logísticos de crecimiento. Estos modelos son ampliamente utilizados para la modelización de procesos de crecimiento acumulativo con forma de S, como es el caso de la propagación de enfermedades infecciosas en epidemiología.

En particular, se han considerado dos modelos distintos: el modelo de Verhulst (1845) y el modelo de Gompertz (1825). Si denotamos con  $C(t)$  el número total de casos acumulados en un momento del tiempo  $t$ , los modelos considerados vienen dados por las siguientes dos ecuaciones,

$$C(t) = \frac{K}{1 + e^{-\gamma(t-\tau)}} \quad (1)$$

$$C(t) = Ke^{-\gamma(t-\tau)} \quad (2)$$

en donde  $K$  es el número máximo de casos acumulados o tamaño final de la epidemia de un brote con una sola oleada,  $\gamma$  es la tasa intrínseca de crecimiento per cápita de la población infectada, y  $\tau$  es el número de unidades de tiempo que transcurren desde el inicio del brote hasta que se alcanza el punto de inflexión. De este modo, el parámetro  $\tau$  indica el punto en donde el número total de casos acumulados deja de crecer de forma exponencial, coincidiendo con el denominado punto álgido o cresta de la curva.

Los datos utilizados en el análisis son datos diarios de casos confirmados acumulados para el conjunto de España y las distintas comunidades autónomas<sup>1</sup>. Los datos han sido obtenidos del Instituto de Salud Carlos III (Ministerio de Sanidad)<sup>2</sup>. Los modelos han sido estimados mediante Mínimos Cuadrados No Lineales (MCNL), utilizando como periodo muestral en cada caso, el periodo comprendido entre el primer dato disponible y el 22 de abril de 2020. Los resultados de la estimación para España se presentan en el cuadro 1, mientras que los resultados de la estimación del modelo de Verhulst y del modelo de Gompertz para las comunidades autónomas se presentan en el cuadro 2 y 3 respectivamente. En cada caso se presenta la fecha de inicio del periodo muestral considerado, el valor estimado de los parámetros  $K$ ,  $\gamma$  y  $\tau$  y sus respectivas desviaciones típicas entre paréntesis robustas a problemas de heterocedasticidad, el número de observaciones ( $N$ ), el valor del coeficiente de determinación ( $R^2$ ), el valor del criterio de información de Akaike (AIC) y la fecha estimada en la que se produce el punto de inflexión o cresta de la curva.

<sup>1</sup> No se han tenido en cuenta las ciudades autónomas de Ceuta y Melilla por el bajo número de casos confirmados que presentan.

<sup>2</sup> Datos disponibles en: [https://covid19.isciii.es/resources/serie\\_historica\\_acumulados.csv](https://covid19.isciii.es/resources/serie_historica_acumulados.csv)

**Cuadro 1.** Resultados de la estimación por MCNL de los modelos logísticos de crecimiento para España

MODELO (fecha inicial)	$K$ (d.t.)	$\gamma$ (d.t.)	$\tau$ (d.t.)	N	$R^2$	AIC	Punto de inflexion
<b>Verhulst</b> 20/02/2020	207443.3*** (3975.709)	0.1567*** (0.0052)	41.8420*** (0.3580)	63	0.9982	1241	02/04/2020
<b>Gomperzt</b> 20/02/2020	236106.9*** (3623.913)	0.0851*** (0.0020)	39.2434*** (0.3569)	63	0.9996	1143	30/03/2020

Notas: significatividad estadística al 1% \*\*\*, 5% \*\*, 10% \*. Desviaciones típicas de los parámetros estimados robustos a problemas de heterocedasticidad entre paréntesis.

Todos los modelos han sido estimados mediante el comando “nl” de stata.

En el caso del modelo de Verhulst, el tamaño final de la epidemia se sitúa levemente por encima de los 207 mil casos, siendo el límite superior del intervalo de confianza al 95% algo superior a los 215 mil casos acumulados. La tasa de crecimiento es de un 15,7% y el punto de inflexión se produce 42 días después del 20 de febrero, es decir, el 2 de abril.

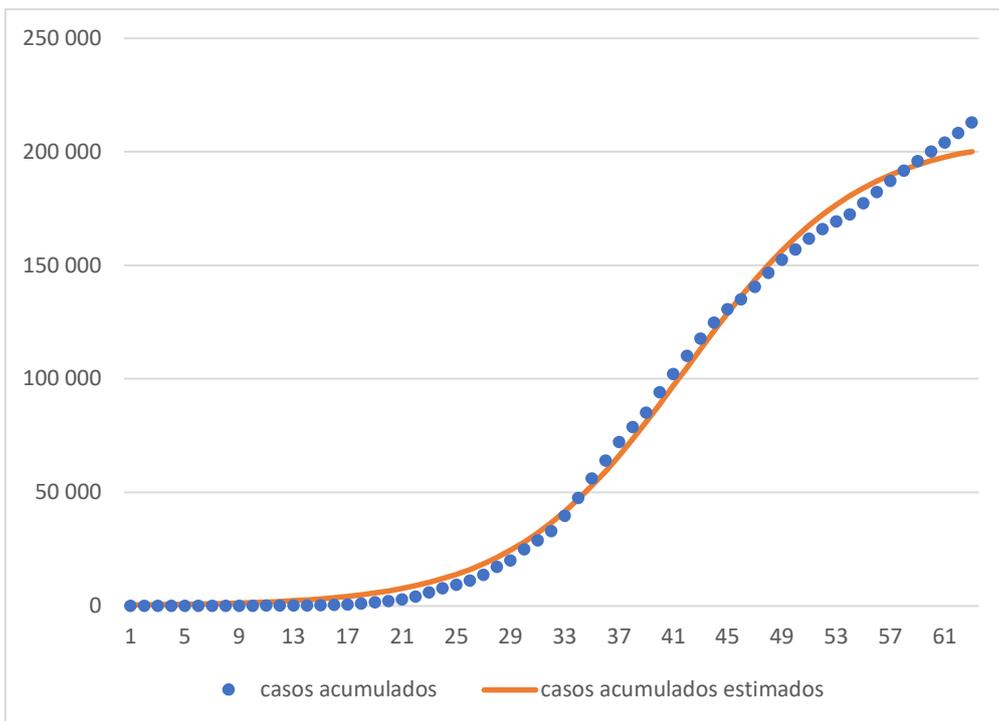
En el caso del modelo de Gomperzt, el tamaño final estimado de la epidemia es notablemente superior, con un total de 236 mil casos acumulados y en donde el límite superior del intervalo de confianza al 95% de confianza sitúa el tamaño final en 243 mil casos. La tasa intrínseca de crecimiento es de un 8,5%, prácticamente la mitad de la estimada mediante el modelo de Verhulst. El punto de inflexión o cresta de la curva se alcanzó 39 días después del 20 de febrero, es decir, a principios de marzo.

El grado de ajuste de ambos modelos es muy elevado, como se desprende del valor del coeficiente de determinación, si bien los valores obtenidos del criterio de información de Akaike sugieren que el modelo de Gomperzt presenta un ajuste mayor.

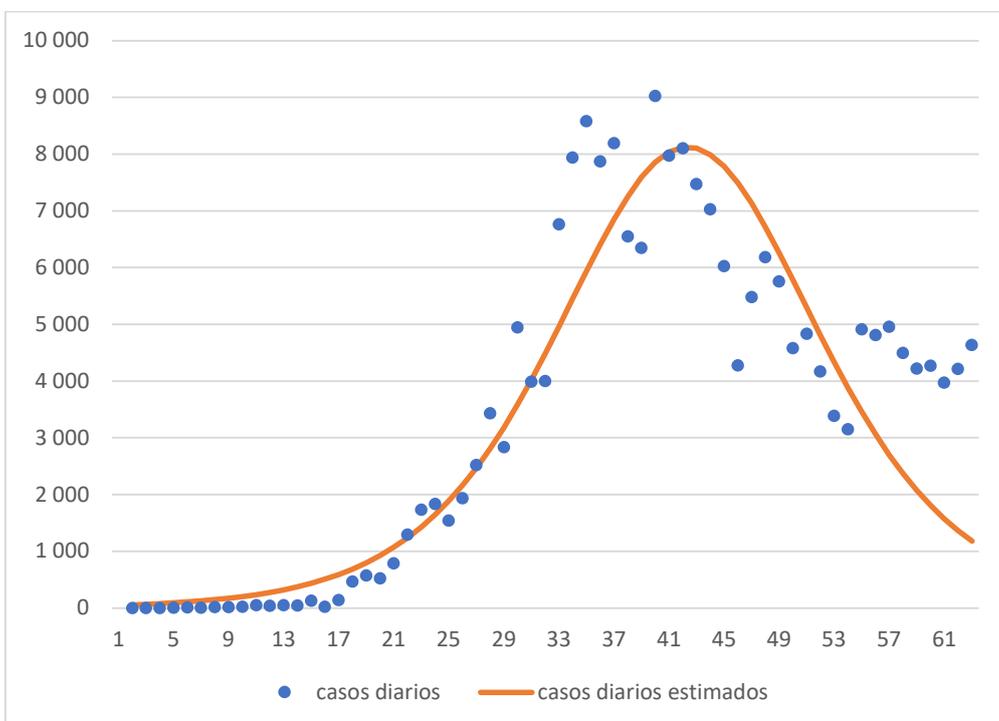
La figura 1 y la figura 2 permite una aproximación visual al grado de ajuste del modelo de Verhulst y del modelo de Gomperzt respectivamente. En cada uno de los dos modelos se representa el valor estimado y el valor real del número total de casos acumulados y del número de casos confirmados diarios durante el periodo de análisis. Como puede apreciarse, si bien ambos modelos presentan un grado de ajuste relativamente alto, posiblemente el tamaño final de la epidemia se esté infraestimando en ambos casos.

**Figura 1.** Modelo de Verhulst aplicado a España

**Casos acumulados**



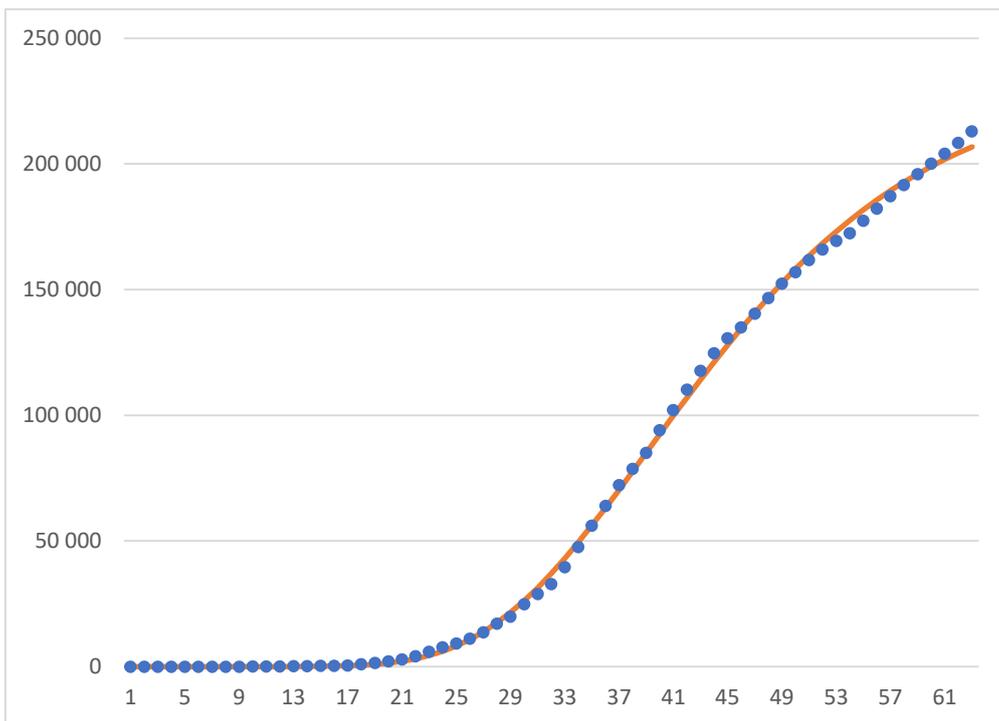
**Casos diarios**



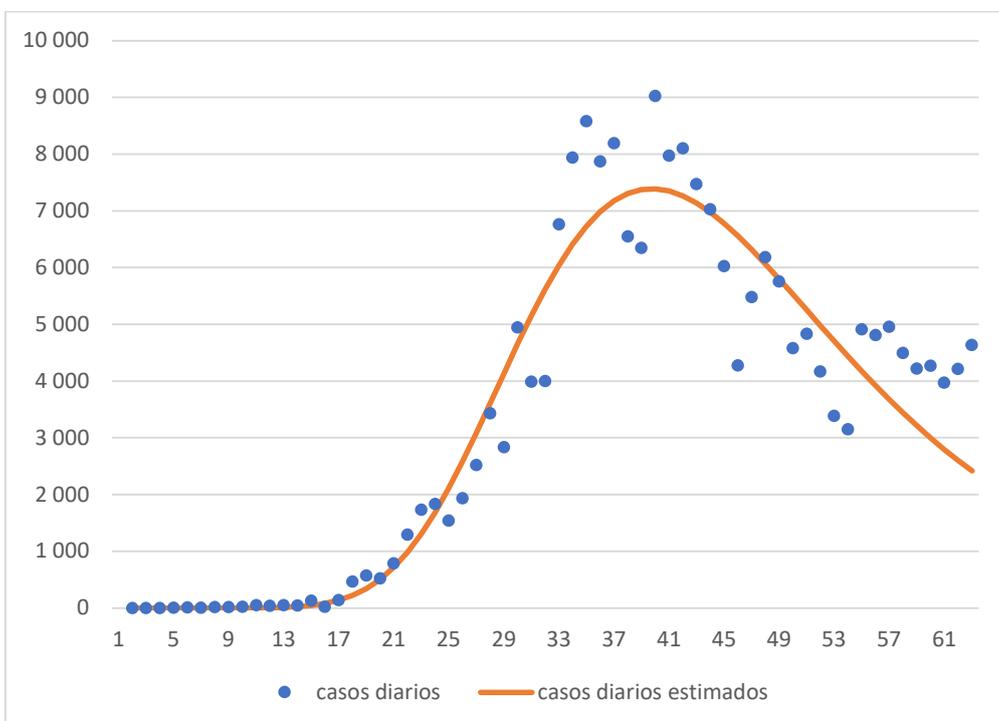
Fuente: elaboración propia a partir de los datos del Instituto de Salud Carlos III

**Figura 2.** Modelo de Gompertz aplicado a España

**Casos acumulados**



**Casos diarios**

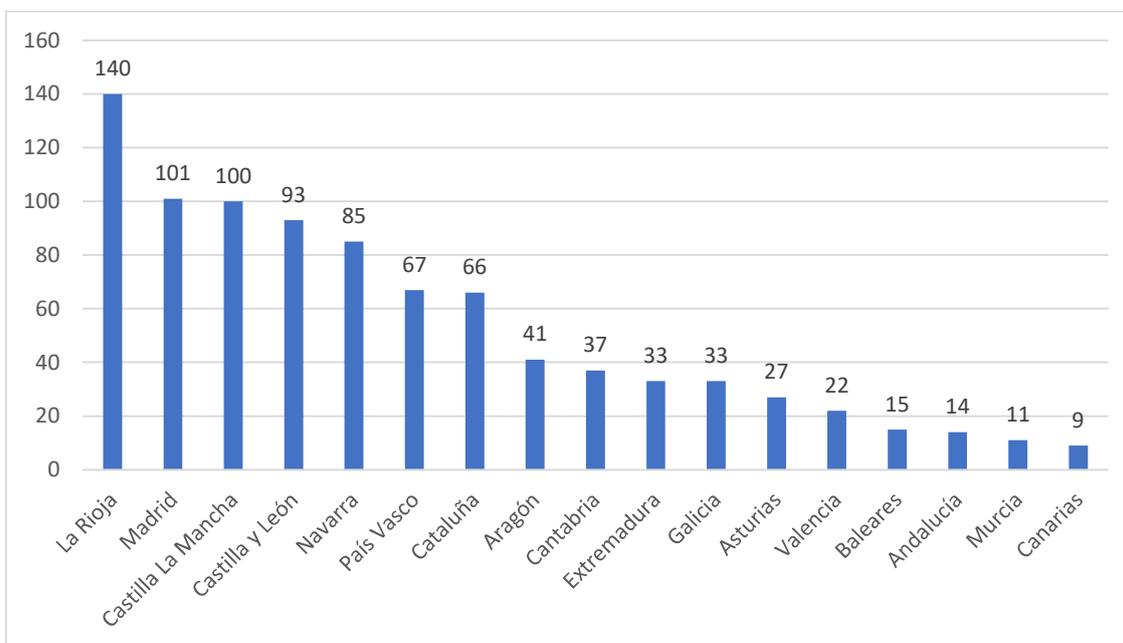


Fuente: elaboración propia a partir de los datos del Instituto de Salud Carlos III

Las diferencias observadas al comparar las estimaciones del modelo de Verhulst y del modelo de Gomerzt se mantienen en las estimaciones obtenidos para las distintas comunidades autónomas. En el modelo de Verhulst el tamaño final de la epidemia es inferior al obtenido en el modelo de Gomerzt, mientras que la tasa intrínseca de crecimiento en el número de casos confirmados es superior. Asimismo, el modelo de Verhulst estima siempre que el punto de inflexión se produce antes en el tiempo que el estimado mediante el modelo de Gomerzt. Atendiendo al criterio de información de Akaike, el modelo de Gomerzt presenta en todos los casos un mejor ajuste. En consecuencia, a continuación se comentan los resultados asociados a este último modelo.

El tamaño estimado final en las dos comunidades autónomas de mayor prevalencia (Madrid y Cataluña) se sitúa en los 68 mil y los 50,3 mil casos respectivamente. El límite superior del intervalo de confianza al 95% se sitúa en 71 mil y 52,6 mil casos respectivamente. El gráfico 1 muestra por su parte la prevalencia final estimada de la enfermedad por coronavirus relativa a la población (casos acumulados por 10.000 hab.). Como puede apreciarse, La Rioja es la comunidad autónoma con mayor prevalencia relativa, con 140 casos acumulados por cada 10.000 hab. A esta le siguen Madrid (101 casos por 10.000 hab.), Castilla La Mancha (100 casos por 10.000 hab.) y Castilla y León (93 casos por 10.000 hab.).

**Gráfico 1.** Prevalencia final estimada por cada 10.000 habitantes.



Fuente: elaboración propia a partir de los datos del Instituto de Salud Carlos III y del INE

La tasa de crecimiento estimada se sitúa en promedio en un 10%. Las comunidades con mayores tasas de crecimiento son Canarias (12,1%), Galicia (11,6%) y Baleares (11,3%) y Murcia (11,1%). El punto de inflexión o cresta de la curva se produce en todos los casos a finales de marzo y principios de abril.

**Cuadro 2.** Resultados de la estimación por MCNL del modelo logístico de crecimiento de Verhulst.

Comunidad Autónoma	Fecha inicial	$K$ (d.t.)	$\gamma$ (d.t.)	$\tau$ (d.t.)	N	R <sup>2</sup>	AIC	Punto de inflexion
Andalucía	26/02/2020	11398.7919*** (134.2164)	0.1822*** (0.0044)	33.9911*** (0.1999)	57	0.9991	770.7489	30/03/2020
Aragón	04/03/2020	5009.4443*** (92.0024)	0.1724*** (0.0067)	28.9370*** (0.3569)	50	0.9978	636.8885	02/04/2020
Asturias	29/02/2020	2417.6763*** (47.8206)	0.1415*** (0.0061)	31.4836*** (0.4260)	54	0.9969	625.3617	01/04/2020
Cantabria	29/02/2020	2026.8832*** (50.5712)	0.1824*** (0.0107)	30.5906*** (0.4850)	54	0.9947	641.2557	31/03/2020
Castilla y León	28/02/2020	18077.2653*** (476.5097)	0.1375*** (0.0059)	37.8020*** (0.5403)	55	0.9971	833.4468	06/04/2020
Castilla La Mancha	29/02/2020	17708.5952*** (214.7208)	0.1658*** (0.0036)	34.7560*** (0.2033)	54	0.9991	765.8075	04/04/2020
Canarias	20/02/2020	2055.9537*** (13.5681)	0.1977*** (0.0040)	37.7723*** (0.1602)	63	0.9994	612.6891	29/03/2020
Cataluña	25/02/2020	43739.0385*** (1058.6554)	0.1551*** (0.0067)	37.7789*** (0.4334)	58	0.9972	989.6219	03/04/2020
Extremadura	29/02/2020	3217.5695*** (69.5712)	0.1545*** (0.0084)	32.2744*** (0.5331)	54	0.9945	685.2385	01/04/2020

Notas: significatividad estadística al 1% \*\*\*, 5% \*\*, 10% \*. Desviaciones típicas de los parámetros estimados robustos a problemas de heterocedasticidad entre paréntesis. Todos los modelos han sido estimados mediante el comando “nl” de stata.

**Cuadro 2.** Resultados de la estimación por MCNL del modelo logístico de crecimiento de Verhulst (CONTINUACIÓN)

Comunidad Autónoma	Fecha inicial	$K$ (d.t.)	$\gamma$ (d.t.)	$\tau$ (d.t.)	N	R <sup>2</sup>	AIC	Punto de inflexion
Galicia	03/03/2020	8381.2204*** (108.5744)	0.1975*** (0.0056)	28.8202*** (0.2188)	51	0.9989	672.3239	01/04/2020
Baleares	20/02/2020	1730.8745*** (33.0024)	0.1879*** (0.0089)	38.2012*** (0.3646)	63	0.9965	696.7884	29/03/2020
Murcia	07/03/2020	1635.6500*** (24.2370)	0.1851*** (0.0069)	22.5170*** (0.3194)	47	0.9974	514.4841	30/03/2020
Madrid	25/02/2020	58706.8925*** (1348.7750)	0.1484*** (0.0049)	36.4347*** (0.4278)	58	0.9981	1005.861	01/04/2020
Navarra	29/02/2020	4889.5967*** (83.1085)	0.1537*** (0.0049)	32.3663*** (0.3449)	54	0.9983	667.9633	01/04/2020
País Vasco	28/02/2020	13062.3369*** (204.6099)	0.1571*** (0.0052)	32.9016*** (0.3325)	55	0.9982	719.135	01/04/2020
La Rioja	01/03/2020	3899.7474*** (20.1973)	0.1562*** (0.0021)	30.9391*** (0.1286)	53	0.9996	555.2687	01/04/2020
Valencia	25/02/2020	10175.9172*** (152.1081)	0.1758*** (0.0066)	34.9078*** (0.3229)	58	0.9979	8156993	31/03/2020

Notas: significatividad estadística al 1% \*\*\*, 5% \*\*, 10% \*. Desviaciones típicas de los parámetros estimados robustos a problemas de heterocedasticidad entre paréntesis. Todos los modelos han sido estimados mediante el comando “nl” de stata.

**Cuadro 3.** Resultados de la estimación por MCNL del modelo logístico de crecimiento de Gompertz.

Comunidad Autónoma	Fecha inicial	K (d.t.)	$\gamma$ (d.t.)	$\tau$ (d.t.)	N	R <sup>2</sup>	AIC	Punto de inflexion
Andalucía	26/02/2020	12402.8798*** (111.5924)	0.1050*** (0.0025)	31.2163*** (0.1464)	57	0.9996	727.3655	28/03/2020
Aragón	04/03/2020	5559.7693*** (-91.9826)	0.0977*** (-0.0031)	26.2901*** (-0.2497)	50	0.9992	583.5797	30/03/2020
Asturias	29/02/2020	2757.4353*** (-46.7932)	0.0771*** (-0.0024)	28.6324*** (-0.3017)	54	0.9991	555.3538	29/03/2020
Cantabria	29/02/2020	2184.5925*** (-57.3864)	0.1091*** (-0.0058)	27.7969*** (-0.3826)	54	0.9976	598.7262	28/03/2020
Castilla y León	28/02/2020	22498.0924*** (-611.1077)	0.0675*** (-0.0023)	36.2643*** (-0.5073)	55	0.9993	758.5217	04/04/2020
Castilla La Mancha	29/02/2020	20384.8001*** (-154.4173)	0.0883*** (-0.0014)	32.4313*** (-0.1287)	54	0.9997	699.8176	01/04/2020
Canarias	20/02/2020	2172.1450*** (-9.0767)	0.1210*** (-0.0021)	34.9344*** (-0.0985)	63	0.9996	580.894	26/03/2020
Cataluña	25/02/2020	50305.1144*** (-1127.3609)	0.0833*** (-0.0027)	35.2879*** (-0.3318)	58	0.9993	908.4395	31/03/2020
Extremadura	29/02/2020	3594.0196*** (-82.4497)	0.0875*** (-0.0041)	29.3999*** (-0.4587)	54	0.9974	644.2201	30/03/2020

Notas: significatividad estadística al 1% \*\*\*, 5% \*\*, 10% \*. Desviaciones típicas de los parámetros estimados robustos a problemas de heterocedasticidad entre paréntesis. Todos los modelos han sido estimados mediante el comando “nl” de stata.

**Cuadro 3.** Resultados de la estimación por MCNL del modelo logístico de crecimiento de Gompertz (CONTINUACIÓN).

Comunidad Autónoma	Fecha inicial	$K$ (d.t.)	$\gamma$ (d.t.)	$\tau$ (d.t.)	N	R <sup>2</sup>	AIC	Punto de inflexion
Galicia	03/03/2020	9038.6328*** (-92.8207)	0.1160*** (-0.0027)	26.1832*** (-0.1415)	51	0.9996	621.1751	29/03/2020
Baleares	20/02/2020	1852.2362*** (-32.8634)	0.1127*** (-0.0043)	35.3938*** (-0.2548)	63	0.9986	696.7884	26/03/2020
Murcia	07/03/2020	1752.2723*** (-20.2702)	0.1110*** (-0.003)	19.6749*** (-0.2198)	47	0.999	470.0195	27/03/2020
Madrid	25/02/2020	68026.3970*** (-1519.3352)	0.0783*** (-0.0026)	33.9589*** (-0.3699)	58	0.9992	957.2796	30/03/2020
Navarra	29/02/2020	5552.6401*** (-68.1326)	0.0837*** (-0.0016)	29.6734*** (-0.2038)	54	0.9997	569.8583	30/03/2020
País Vasco	28/02/2020	14637.5532*** (-179.4577)	0.0875*** (-0.002)	30.1039*** (-0.2109)	55	0.9995	717.8637	29/03/2020
La Rioja	01/03/2020	4417.1851*** (-74.5411)	0.0846*** (-0.0032)	28.2493*** (-0.2645)	53	0.9987	618.6956	29/03/2020
Valencia	25/02/2020	11090.1412*** (-149.1171)	0.1019*** (-0.0029)	32.0747*** (-0.2219)	58	0.9994	745.2869	28/03/2020

Notas: significatividad estadística al 1% \*\*\*, 5% \*\*, 10% \*. Desviaciones típicas de los parámetros estimados robustos a problemas de heterocedasticidad entre paréntesis. Todos los modelos han sido estimados mediante el comando “nl” de stata.

## Referencias

Gompertz B. (1825). "On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies". Philosophical Transactions of the Royal Society of London, vol. 115, pp. 513-583.

Verhulst, P. (1845). "Recherches mathématiques sur la loi d'accroissement de la population", Mémoires de l'Académie Royale des Sciences, des Lettres et des Beaux-Arts de Belgique 18, Bruselas, pp. 1-42.

WHO (2020a). "Pneumonia of unknown cause – China", Emergencies preparedness, response, Disease outbreak news, World Health Organization (WHO). Disponible en: <https://www.who.int/csr/don/05-january-2020-pneumonia-of-unkown-cause-china/en/>.

WHO (2020b). "WHO Director-Genral's opening remarks at the media briefing on COVID-19", WHO Director-General, Speeches, Detail, World Health Organization (WHO). Disponible en: <https://www.who.int/dg/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020>

WHO (2020b). "Coronavirus disease 2019 (COVID-19): situation report - 94", Health Organization (WHO). Disponible en: [https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200423-sitrep-94-covid-19.pdf?sfvrsn=b8304bf0\\_4](https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200423-sitrep-94-covid-19.pdf?sfvrsn=b8304bf0_4)