



Munich Personal RePEc Archive

**Review of models for the seasonal
adjustment of monthly and quarterly
series of economic activity. Year 2021.**

Frank, Luis

16 April 2021

Online at <https://mpra.ub.uni-muenchen.de/112278/>
MPRA Paper No. 112278, posted 08 Mar 2022 18:37 UTC

Revisión de modelos para la desestacionalización y proyección de series macroeconómicas trimestrales. Año 2021

Luis Frank ^{*†}

Resumen

El trabajo presenta una revisión de los modelos ARIMA estacionales utilizados para desestacionalizar y proyectar las principales variables macroeconómicas de Argentina. La revisión fue motivada principalmente por el impacto la pandemia de COVID iniciada en 2020. Los resultados confirmaron la necesidad de la revisión - a pesar de haber transcurrido sólo cuatro trimestres desde la última revisión - y la conveniencia de introducir variables indicadoras de valores atípicos durante la pandemia, salvo para series de comercio exterior. Sin embargo, una inspección detallada de los resultados muestra que los modelos debieron haberse revisado pasados dos o tres trimestres de iniciada la pandemia para garantizar que las especificación fueran coherentes con la intervención realizada sobre cada serie.

Palabras clave: desestacionalización de series, ARIMA estacional.

JEL: C820

Abstract

The paper presents a review of the seasonal ARIMA models used to deseasonalize and project the main macroeconomic variables of Argentina. The review was mainly motivated by the impact of the COVID pandemic starting in early 2020. The results confirmed the necessity of the review - despite only four quarters passed since the last revision - and the suitability of the outlier-variables introduced during the pandemic, except for foreign trade series. Nevertheless, a closer inspection of the results show that models should have been revised two or three quarters after the start of the pandemic to ensure that the specifications were consistent with the intervention done on each series.

Keywords: seasonal adjustment of series, seasonal ARIMA.

JEL: C820

1 Introducción

La especificación de modelos para el desestacionalizado de series requiere revisiones periódicas a fin de (a) verificar que la especificación definida oportunamente continúa vigente aún después de

*DNMyP. Secretaría de Política Económica. Ministerio de Economía. Av. Hipólito Yrigoyen 250, C1086AAB. Buenos Aires, Argentina.

†Universidad de Buenos Aires. Facultad de Agronomía. Av. San Martín 4453, C1417DSE. Buenos Aires, Argentina.

incorporar nuevos datos a la serie; (b) incorporar variables relacionadas con eventos novedosos y de una importancia tal que los vuelve imposibles de ignorar como por ejemplo la pandemia de COVID iniciada en 2020; (c) corregir la interpretación de ciertos eventos, en términos de duración (temporaria o permanente), que pudiera haber sido desmentida por evidencia más reciente; (d) actualizar la especificación de modelos en concordancia con la revisión de cifras realizada por la fuente original de datos. La frecuencia de dichas revisiones no se encuentra estandarizada aunque la práctica usual en nuestro país consiste en revisar los modelos toda vez que ocurran eventos trascendentes, o se dejen de satisfacer los criterios de bondad de ajuste que justificaron la selección de los modelos actuales. En vista de que se ha cumplido un año desde el comienzo de un evento tan relevante como la pandemia de COVID, y que recientemente se ha completado el año 2020 del Sistema de Cuentas Nacionales (SCN) trimestrales, creemos oportuno revisar los modelos propuestos a principios del año pasado [3, 4] para proyectar agregados del SCN y de intercambio comercial.¹

2 Objetivos

El objetivo de este informe es revisar las especificaciones propuestas a principios de 2020 para desestacionalizar y proyectar indicadores trimestrales de actividad económica, y exportaciones por grandes rubros e importaciones por uso económico. La revisión se hará siguiendo el procedimiento de selección de modelos del programa X-13 ARIMA-SEATS de U.S. Census Bureau pero con la incorporación de variables exógenas que han probado ser relevantes en el pasado para modelar las series en cuestión. Por brevedad, no reiteraremos en este informe la descripción de estas variables, ni el procedimiento de búsqueda y selección de modelos del programa X-13 ARIMA-SEATS, ya que éstos fueron descritos en detalle en informes anteriores (véase e.g. [3] y [4]). Una vez revisadas las especificaciones, las compararemos con las obtenidas en 2020 y extraeremos algunas conclusiones sobre la forma de proceder en caso de repetirse eventos como la pandemia de COVID que afecta al país desde el año pasado.

3 Metodología

La metodología de revisión consiste básicamente en ajustar y seleccionar mediante criterios de bondad de ajuste el modelo ARIMA estacional (con variables exógenas) que mejor predice los valores observados, especialmente en el tramo final de la serie. Esta selección se realiza dos etapas. En la primera, se ajusta el modelo $(0, 1, 1)(0, 1, 1)$ (también conocido como el *modelo de la aerolínea*) transformado logarítmicamente y sin transformar, a fin de evaluar, mediante el criterio de información de Akaike o AICC, si dicha transformación es necesaria. En la segunda etapa, se realiza una búsqueda exhaustiva del modelo “óptimo” ajustando sucesivamente las 576 especificaciones $ARIMA(p, d, q) (P, D, Q)$ que surgen de combinar de parámetros $p, q = \{0, 1, 2, 3\}$ y $P, Q = \{0, 1, 2\}$ hasta un orden máximo de integración de 1. De todos los modelos evaluados se retiene aquel que presenta el menor error porcentual absoluto medio o MAPE (*mean absolute percetage error*, en inglés) en los últimos tres años de la serie, siempre que éste no supere el 15 %, el valor- p asociado al estadístico Q de Ljung-Box Q-statistic test sea mayor a 5 %, y no haya evidencia de sobrediferenciación.² El lector hallará una descripción completa de este procedimiento en la bibliografía antes mencionada al final del informe. Las series modeladas de esta forma fueron las siguientes:

¹El esquema de revisiones anuales que iniciamos en este informe sigue la llamada *partial concurrent adjustment strategy* de revisión de modelos (ver [5, cap. 7, § 70-80])

²Ver [7, 140-141] para mayores detalles.

- Agregados macroeconómicos trimestrales del Sistema de Cuentas Nacionales a precios constantes de 2004 (período I-2004 a IV-2020) elaborados por la DNCN de INDEC.
- Series trimestrales de valor agregado bruto (VAB) a nivel de letra (secciones de la clasificación CIU rev. 3.1) a precios constantes de 2004 (período I-2004 a IV-2020) elaboradas por la DNCN.
- Índices trimestrales de cantidades y valor FOB de exportaciones por grandes rubro e importaciones por uso económico (período I-2004 a I-2021) elaborados por la DNESE de INDEC, y exportaciones e importaciones de servicios reales, en dólares elaborados por la DNCI de INDEC, período I-2006 a IV-2020.³

Las variables exógenas incluidas en los modelos ARIMA fueron: días de la semana en relación a la cantidad de domingos (efecto calendario), cantidad de feriados y huelgas generales (siempre que no cayeran en domingo) en el trimestre, y variables indicativas de valores atípicos o *outliers*.⁴ Los días de la semana y la cantidad de feriados y huelgas se incorporaron a través de una matriz de datos específica elaborada por el autor, pero los *outliers* debieron ser identificados a través del procedimiento de identificación automático (basado en pruebas *t*) del programa X13 ARIMA (ver [7, §4.6, pp. 39-40]). En este punto nos apartamos de la práctica sugerida e.g. por [1, 6] de incorporar variables indicadoras específicas para tomar en cuenta el efecto de la pandemia de COVID, ya esta práctica implicaría (i) confundir un fenómeno determinista como es la cuarentena con un fenómeno aleatorio como la evolución de contagios y su impacto en la fuerza laboral, hábitos de consumo, etc.; y (ii) establecer de antemano el carácter fijo y atípico de la pandemia, descartando toda posibilidad de un cambio de régimen en la serie. A decir verdad, este tema ha sido motivo de controversia entre los expertos ya que a principios de la pandemia era imposible saber si el efecto de la misma sería transitorio, en cuyo caso correspondía agregar variables indicadoras, o permanente, en cuyo caso correspondía revisar la especificación completa del modelo una vez que se dispusiera de una cantidad suficiente de períodos post-pandemia. En los primeros meses de 2020 el criterio predominante entre los expertos fue considerar los efectos de la pandemia/cuarentena como *shocks* aditivos [1] para no alterar la componente de tendencia-ciclo de la serie en cuestión, al menos hasta tener mayor evidencia sobre los efectos de largo plazo de la pandemia sobre la actividad económica. En nuestro país, la DNCN mantuvo este criterio durante los dos primeros trimestres de pandemia, pero “a partir del tercer trimestre, no se continuó aplicando la intervención mediante outliers aditivos” [6] en vista de la reapertura gradual de actividades económicas y del otorgamiento de permisos de circulación a trabajadores. Más allá de la racionalidad de criterio, no se desprende de informes posteriores que esta forma de intervención haya sido validada mediante la revisión de modelos y la identificación de *outliers* a través del procedimiento habitual, es decir, sometiendo a todos los trimestres del año 2020 al mismo mecanismo de identificación aplicado en años anteriores. Por esta razón optamos por revisar los modelos y chequear luego si la intervención de II-2020 y III-2020 estuvo realmente justificada.

4 Resultados

Los cuadros 1 a 4 muestran las especificaciones obtenidas con el programa X-13 ARIMA-SEATS. Se puede observar que en prácticamente todas las series de agregados del SCN y VAB sectoriales, el segundo trimestre de 2020 aparece identificado como cambio temporario o

³Las cifras anteriores a 2006 estaban clasificadas según MBP 5. La DNCI migró en los últimos años a MBP 6 pero reclasificó las series desde 2006.

⁴Omitimos la inclusión de variables climáticas como el índice ONI porque ésta no resultó explicativa en la revisión 2020 para series trimestrales, no así para series mensuales.

TC.^{5,6} En cambio, las series de exportaciones, salvo las MOI, y las importaciones no parecen haber sufrido *shocks* relevantes durante la pandemia de COVID. Cabe aclarar, sin embargo, que la identificación de *outliers* en algunas series no fue confiable debido a que ninguno de las 576 modelos ajustados satisfizo los criterios de selección antes mencionados por lo cual se retuvo el modelo de menor MAPE en los último tres años aunque éste superara el umbral de 15%. Ante casos como éstos, el programa X13 ARIMA-SEATS interrumpe la rutina de identificación de *outliers*, en cuyo caso cualquier curso de acción que se siga, como por ejemplo definir *outliers* a priori o forzar la identificación de *outliers* con un modelo arbitrario conduce a resultados incomparables. Las series que no pudieron ser modeladas en la forma habitual fueron la variación de existencias (cuadro 1), el VAB de servicio doméstico (cuadro 2), y las importaciones de combustible y lubricantes, y de vehículos de pasajeros, en cantidades y en valor FOB (cuadro 4). En todos estos casos, salvo uno, el estadístico Q sin M_2 permitió validar nuestra selección *ad hoc*. Solamente el modelo de servicio doméstico (letra P del cuadro 2) fue rechazado por presentar un estadístico Q sin $M_2 > 1, 2$.

5 Conclusión

El informe presenta una amplia revisión de modelos ARIMA estacionales útiles para desestacionalizar y proyectar las principales variables económicas de Argentina. Esta revisión fue motivada principalmente por la controversia surgida a principios de la pandemia de COVID respecto del curso de acción a seguir para desestacionalizar series contemporáneamente con *shocks* cuya naturaleza (transitoria o permanente) resulta imposible establecer de antemano. La simple inspección de resultados y su comparación con los de la revisión de 2020 confirma que las especificaciones obtenidas hace un año perdieron vigencia. Los resultados confirman además que la introducción temprana de variables indicadoras relacionadas con la pandemia fue acertada, salvo para las series de comercio exterior. No obstante este acierto, los modelos utilizados hasta II-2020 debieron haberse revisado pasados dos o tres trimestres desde el inicio de la pandemia para garantizar que la especificación completa del modelo fuera coherente con la intervención realizada con variables indicadoras de valores atípicos. Dicho de otro modo, si los modelos no se revisan integralmente, no hay garantía que las (nuevas) variables indicadoras que resultaron “significativas” en el “viejo” modelo lo sigan siendo bajo la nueva especificación. En línea con este razonamiento, sugerimos que de manifestarse una segunda ola de COVID durante 2021 se introduzca inmediatamente una variable indicadora de cambio transitorio (TC) y se revise el modelo pasados dos trimestres desde la intervención pero con identificación automática de *outliers*.⁷ Si bien esta estrategia resuelve el problema de la desestacionalización ante *shocks* contemporáneos con la misma, no resuelve el problema de la proyección de cifras ante eventos posiblemente recurrentes como sucesivas olas de contagios de COVID.⁸ En caso de recurrencia, no vemos otra alternativa, a los fines de proyectar cifras, que crear una variable exógena específica para COVID, emulando las de los *outliers*, y replicarla con la frecuencia esperada en los sucesivos trimestres que podrían verse afectados. Por ejemplo, para proyectar el año 2021 se debería incluir una variable TC en II-2020 (para estimar el efecto pandémico) y II-2021 (para proyectar) bajo el supuesto de que el impacto de la pandemia en 2021 sería similar que en 2020. O bien modelar el impacto relativo de la pandemia en olas sucesivas a partir de eventos similares

⁵El manual de X13 ARIMA-SEATS [7, p. 35] define: “TCs allow for an abrupt increase or decrease in the level of the series that returns to its previous level exponentially rapidly”.

⁶La Central Statistics Office de Irlanda [2] halló que en la mayoría de las series irlandesas el impacto de la pandemia se modelaba mejor como una secuencia de saltos de nivel (*level shift*) o LS.

⁷Por construcción el agregado de variables TC al final de la serie es indistinguible del agregado de variables de cambio aditivo puntual o AO.

⁸Sabemos de esta recurrencia por lo ocurrido en países del hemisferio norte

Cuadro 1: Especificaciones para la proyección de agregados macroeconómicos del SCN a precios constantes de 2004.

Descrip.	Transf.	Modelo	const.	Día sem.	Feriados, huelgas	Año bis.	<i>Outliers</i>	Q s/M_2
PIB	log	(2 1 3) (1 1 1)	si	si	no	si	TC2012.3 TC2020.2	0,24
Importación	–	(2 0 1) (1 0 2)	si	no	no	si	–	0,50
Consumo privado	log	(2 0 3) (1 0 2)	si	si	no	no	AO2005.2 AO2009.2 TC2020.2	0,29
Consumo público	log	(2 1 2) (1 1 0)	si	si	no	no	TC2020.2	0,28
Exportación	log	(1 0 3) (0 0 2)	si	si	no	no	–	0,67
FBKF	–	(3 0 0) (0 0 2)	si	no	no	no	AO2020.2	0,42
Var. exist.	–	(0 0 3) (0 1 1)	no	si	no	si	–	0,88

ocurridos en el pasado, como la gripe aviar H1N1, y utilizar el impacto relativo estimado para la segunda ola como regresor en II-2021, el impacto relativo de la tercera ola como regresor de II-2022, etc.

Cuadro 2: Especificaciones para la proyección de VAB sectoriales a precios constantes de 2004

Letra	Descrip.	Transf.	Modelo	const.	Día sem.	Feriados, huelgas	Año bis.	<i>Outliers</i>	Q s/ M_2
A	Agricultura, ganadería	log	(1 0 2) (0 1 2)	si	si	no	si	AO2004.2 LS2008.4 AO2009.2 LS2010.2 AO2012.2 AO2015.2 TC2018.2 AO2020.2	0,40
B	Pesca	log	(3 0 1) (1 0 2)	si	si	no	no	LS2005.4	0,73
C	Minería	–	(3 0 3) (0 1 2)	si	no	si	no	LS2020.2	0,46
D	Industria	–	(0 0 2) (2 0 1)	si	no	no	si	TC2007.4 AO2020.2	0,27
E	Elec., gas y agua	log	(2 1 1) (0 0 0)	si	no	no	no	–	0,63
F	Construcción	–	(0 1 0) (0 1 2)	si	no	no	no	TC2020.2	0,42
G	Comercio	–	(2 0 3) (1 1 2)	no	si	no	no	AO2020.2	0,42
H	Hoteles y restaurant.	–	(0 0 1) (2 0 2)	no	si	no	no	LS2008.2 TC2009.3 TC2020.1 LS2020.2	0,31
I	Transp. y comun.	log	(3 0 2) (1 0 0)	si	si	no	si	LS2020.2	0,35
J	Intermed. financiera	log	(3 0 2) (0 0 0)	si	no	no	no	–	0,56
K	Actividades empresar.	log	(3 1 2) (0 1 1)	si	si	si	no	TC2020.2	0,16
L	Administ. pública	log	(1 1 3) (2 1 2)	si	si	si	si	AO2009.2 TC2020.2 LS2020.3	0,43
M	Enseñanza	log	(0 0 1) (0 0 2)	si	si	si	no	LS2008.2 LS2009.4 LS2015.4	0,73
N	Salud	log	(3 0 3) (2 0 1)	si	si	no	no	LS2006.1 AO2020.1 TC2020.2	0,20
O	Servicios personales	–	(0 1 2) (2 1 1)	no	si	si	no	LS2020.2 TC2020.2 AO2020.4	0,63
P	Serv. dom.	–	(3 0 3) (0 0 0)	si	si	no	no	TC2020.2	1,18
–	Imp. netos de subs.	–	(2 0 0) (0 0 2)	si	si	no	no	LS2006.4 LS2010.2 LS2011.2 TC2020.2	0,74

Cuadro 3: Especificaciones para la proyección de índices (base 2004 = 100) de cantidades y valores FOB de exportaciones

	Descrip.	Transf.	Modelo	const.	Día sem.	Feriados, huelgas	Año bis.	<i>Outliers</i>	Q s/ M_2
Cant.	Produc. primar.	–	(3 1 3) (2 1 1)	no	no	no	no	–	0,33
	MOA	–	(3 0 2) (0 0 0)	si	no	no	no	–	0,70
	MOI	–	(3 1 3) (0 1 2)	no	si	no	no	TC2020.2	0,24
	Comb., energía	–	(2 1 2) (0 1 1)	no	si	no	no	–	0,58
	Servicios	–	–	–	–	–	–	–	–
Valor FOB	Produc. primar.	log	(3 0 3) (2 0 2)	si	si	si	si	LS2009.4 AO2013.4	0,35
	MOA	log	(0 0 0) (0 0 2)	si	si	no	no	LS2007.3 AO2008.3 LS2010.3	1,11
	MOI	–	(3 1 2) (0 1 1)	no	si	no	no	TC2020.2	0,23
	Comb., energía	–	(3 0 3) (0 0 2)	si	si	si	no	LS2008.4	0,87
	Servicios	–	(3 0 2) (1 0 2)	si	si	no	no	LS2009.4 LS2020.2	0,31

Cuadro 4: Especificaciones para la proyección de índices (base 2004 = 100) de cantidades y valores FOB de importaciones

	Descrip.	Transf.	Modelo	const.	Día sem.	Feriados, huelgas	Año bis.	<i>Outliers</i>	Q s/ M_2
Cant.	Bs. cap.	–	(3 0 2) (1 0 1)	si	si	no	no	AO2012.2	0,31
	Bienes interm.	–	(3 0 3) (0 1 0)	si	si	si	si	AO2007.4 TC2011.3 AO2012.3	0,55
	Comb.	log	(1 1 2) (1 1 0)	no	no	no	si	–	0,52
	Piezas	log	(2 0 0) (0 0 2)	si	no	no	no	–	0,79
	Bienes consumo	log	(0 0 3) (0 0 2)	si	si	no	si	LS2005.2 LS2006.2 TC2009.1 LS2015.3	0,36
	Vehículos	log	(2 1 3) (2 1 2)	si	si	no	si	–	0,87
	Servicios	–	–	–	–	–	–	–	–
Valor FOB	Bs. cap.	–	(2 0 2) (1 0 1)	no	si	si	si	–	0,34
	Bienes interm.	log	(3 1 1) (1 1 2)	si	si	no	no	LS2009.1	0,45
	Comb.	log	(2 1 1) (2 1 2)	no	si	no	no	–	0,42
	Piezas	log	(3 0 0) (1 0 2)	si	si	si	si	–	0,76
	Bs. cons.	log	(3 0 3) (2 0 0)	si	si	si	si	LS2009.1	0,28
	Vehículos	log	(0 0 3) (1 0 1)	si	si	no	si	–	0,93
	Servicios	–	(3 0 2) (1 1 1)	si	no	no	no	LS2018.3 LS2020.2	0,71

Referencias

- [1] EUROSTAT, 2020. Guidance on treatment of COVID-19 crisis effect on data. Methodological note. Disponible en: <https://ec.europa.eu/eurostat/web/main/publications/manuals-and-guidelines>
- [2] Foley P., 2021. Seasonal Adjustment of Irish Official Statistics During the COVID-19 Crisis. *Statistical Journal of the IAOS* 37(1):57-66.
- [3] Frank L., 2020a. Revisión de modelos para la desestacionalización de series mensuales y trimestrales de actividad económica. MPRA Paper No. 111423. Disponible en <https://mpra.ub.uni-muenchen.de/111423/>
- [4] Frank L., 2020b. Nota metodológica sobre la proyección de exportaciones de bienes y servicios reales. MPRA Paper No. 111424. Disponible en <https://mpra.ub.uni-muenchen.de/111424/>
- [5] IMF, 2017. Quarterly National Accounts Manual. 2017 Edition. Disponible en <https://www.imf.org/external/pubs/ft/qna/>.
- [6] Instituto Nacional de Estadística y Censos, 2020. Informe de avance del nivel de actividad. Cuarto trimestre de 2020. Cuentas nacionales vol. 5, nro. 4/Informes técnicos vol. 5, nro 51. ISSN 2545-6636.
- [7] U.S. Census Bureau, 2017. X-13ARIMA-SEATS Reference Manual. Accessible HTML Output Version. Version 1.1. Disponible en: <http://www.census.gov/srd/www/x13as/>