CrossMark

On the invertibility of seasonally adjusted series

Yuliya Lovcha[1] (iD) · **Alejandro Perez-Laborda[1]** ·
Luis Gil-Alana[2]

Abstract This paper examines the implications of the seasonal adjustment by an
ARIMA model based (AMB) approach in the context of seasonal fractional integra-
tion. According to the AMB approach, if the model identified from the data contains
seasonal unit roots, the adjusted series will not be invertible that has serious impli-
cations for the posterior analysis. We show that even if the ARIMA model identified
from the data contains seasonal unit roots, if the true data generating process is sta-
tionary seasonally fractionally integrated (as it is often found in economic data), the
AMB seasonal adjustment produces dips in the periodogram at seasonal frequencies,
but the adjusted series still can be approximated by an invertible process. We also per-
form a small Monte Carlo study of the log-periodogram regression with tapered data
for negative seasonal fractional integration. An empirical application for the Spanish
economy that illustrates our results is also carried out at the end of the article.

Keywords Seasonality · Invertibility · Fractional integration · TRAMO-seats ·
Tapering

JEL Classification C15

✉ Yuliya Lovcha
   yuliya.lovcha@gmail.com

[1] Departament d'Economia i CREIP, Universitat Rovira-i-Virgili, Av. Universitat 1, 43204 Reus,
   Spain

[2] Departamento de Economia, Universidad de Navarra, Campus Universitario, 31009 Pamplona,
   Spain

1 Introduction

Given the seasonal nature of many macroeconomic time series, seasonal adjustment is a widespread practice and millions of series are routinely adjusted, some of them are not even publicly available in the nonadjusted version. Seasonal adjustment is believed to remove undesirable fluctuations at seasonal frequencies without producing significant changes at other frequencies (especially at the low part of the spectrum) making the data easily tractable, thereby simplifying posterior modeling and analysis. However, the properties of the adjusted series crucially depend on the method used for the adjustment and the initial properties of the series, and they may result just as unattractive for analysts as seasonality itself.

In this paper, we examine one of the important features of the adjusted data: dips in the periodogram at seasonal frequencies and the resulting noninvertibility of the adjusted series. The spectral dips (or zeros) are produced by all seasonal adjustment methods used in practice, regardless of whether it is a naive adjustment by seasonal dummies or a sophisticated ARIMA-model-based (AMB) signal extraction produced by specialized programs. Ooms and Hassler (1997) point out that the regression on seasonal dummies generates zeros in the periodogram at seasonal frequencies that can lead to the singularities in the log-periodogram regression. Nerlove (1965) applies Census X-11 and the modified 'Hannan' method and concludes that both methods remove more than just the seasonal component. Grether and Nerlove (1970) show that the phenomenon observed in Nerlove (1965), namely dips created near the seasonal frequencies after adjustment, is obtained as a result of 'optimal' adjustment procedure as well. The consequent seasonal adjustment routines of Census, X-11-ARIMA, and X-12 produce the same result by construction. Gomez and Maravall (2001) call attention to the fact that TRAMO-Seats generates dips at seasonal frequencies whenever the model identified for the data contains seasonal unit roots. According to Gomez and Maravall (2001), the spectral zeros are the frequency counterpart of the unit MA roots and therefore the adjusted series is not invertible and does not accept autoregressive AR (or VAR) approximations to its Wold representation. Although often ignored, this is, perhaps, the most important practical implication of AMB adjustment, since AR (and VAR) approximations to seasonally adjusted data are typically carried out in the applied econometric work.

In this work we analyze in detail the dips at the seasonal frequencies and the apparent noninvertibility produced by the AMB approach, within the fractional integration (FI) framework, which admits a wider definition of the invertibility condition than the one applied by Gomez and Maravall (2001). In particular, a fractionally integrated (FI) process is (seasonally) invertible whenever the FI coefficients at seasonal frequencies are higher than -0.5. In addition, notice that the negative seasonal FI parameters correspond to the spectral zeros at seasonal frequencies. Thus, the process can have spectral zeros at seasonal frequencies but still remain invertible.

We choose TRAMO-Seats for Windows (TSW) as the representative AMB seasonal adjustment program. TSW is a pair of the data adjustment programs developed by the Bank of Spain that have been intensively employed by Eurostat since 1994, and

nowadays their use has been extended to various European countries (Gomez and Maravall 2001; Eurostat 2009).[1]

To check for the invertibility of the series adjusted by TSW, we produce simulations for a set of processes. We do not restrict the analysis to processes with integer orders of integration at zero and the seasonal frequencies since it has been shown by many authors that FI at seasonal frequencies is a widespread phenomenon in economics (Porter-Hudak 1990; Gil-Alana and Robinson 2001; etc.). However, we also make simulations for a set of Airline models, which are the default models in TSW. For each model, we simulated 500 series, we adjust them by TSW and then, we estimate the fractional differencing parameters at the seasonal frequencies in the adjusted series with the log-periodogram regression with tapered data.

We find that if the data generating process (DGP) follows the default Airline model, TSW always identifies seasonally nonstationary ARIMA models for the data, and the adjusted series produced by TSW are indeed noninvertible, which is in line with the results of Gomez and Maravall (2001). However, if the original series is FI at the seasonal frequencies, which is less restrictive and very plausible in many cases according to the empirical evidence, the adjusted series may be approximated by an invertible process depending on the stationarity of the original series. If the DGP is a seasonally stationary FI model, TSW is less prone to identifying a seasonally nonstationary model. Moreover, even if the model chosen by the program is a seasonally nonstationary, the adjusted series does contain dips at the seasonal frequencies but these dips correspond to negative seasonal FI with coefficients greater than -0.5. Hence the adjusted series still can be approximated by an invertible process. On the contrary, the adjustment of a series generated from a model with nonstationary FI seasonality results in noninvertible negative FI coefficients. Note that this last result is not straightforward since overdifferencing is expected to be larger for data generated from seasonal stationary models when a nonstationary model is employed for adjustment.

The paper is organized as follows. Section 2 describes the problem. Section 3 briefly introduces the ideas behind the concept of seasonal FI. The simulation set-up and the results are presented in Sect. 4. Section 5 contains a small empirical application illustrating the results reported in Sect. 4. Section 6 concludes the paper.

2 The problem

SEATS is an "ARIMA-model-based" (AMB) seasonal adjustment routine. Within the AMB approach, the program TSW starts by identifying an ARIMA model to the observed data

$$\Phi(B)x_t = \Theta(B)a_t, \tag{1}$$

where B is backward shift operator, $B^i x_t = x_{t-i}$, a_t is an iid $N\left(0, \sigma^2\right)$ innovation, the polynomial $\Phi(B) = F(B)\left(1 - B\right)^D \left(1 - B^\tau\right)^{D_\tau}$ contains nonseasonal and seasonal roots respectively, τ is the number of observations per year, D is the (integer) order of integration at frequency zero, while D_τ is the (integer) seasonal order of differentiation.

The polynomials $F(B)$ and $\Theta(B)$ are finite in B, the first one includes stationary seasonal and nonseasonal AR roots and the second is an invertible MA polynomial.

If the aim of the application of the TSW is the seasonal adjustment, the observed series x_t is decomposed into the mutually orthogonal seasonally adjusted (SA) component n_t and seasonal component s_t:

$$x_t = n_t + s_t$$

The processes for the SA (or signal) and the seasonal components follow ARIMA processes:

$$\Phi_n(B)n_t = \Theta_n(B)a_{nt}, \quad a_{nt} \sim niid(0, \sigma_n^2) \tag{2}$$
$$\Phi_s(B)s_t = \Theta_s(B)a_{st}, \quad a_{st} \sim niid(0, \sigma_s^2) \tag{3}$$

such that $\Phi(B) = \Phi_n(B)\,\Phi_s(B)$ and $\Theta(B)a_t = \Phi_n(B)\,\Theta_s(B)\,a_{st} + \Phi_s(B)\,\Theta_n(B)\,a_{nt}$.

Thus, the seasonal component captures the peaks around the seasonal frequencies, which may be subtracted by the filter.

For the seasonal adjustment, the purpose is, given x_t, to obtain the estimator of $\hat{n}_t$ such that $E\left[(n_t - \hat{n}_t)^2 \,|x_t\right]$ is minimized, i.e. the MMSE estimator of n_t.

Denote $n_t = \Psi_n(B)\,a_{nt}$, with $\Psi_n(B) = \Theta_n(B)/\Phi_n(B)$; x_t and s_t are defined in the same way. As it is shown in Whittle (1963), $\hat{n}_t$ is obtained by means of the Wiener-Kolmogorov (WK) filter as the MMSE estimator of the signal given the observed series:

$$\hat{n}_t = \left[\frac{\sigma_n^2\Psi_n(B)\Psi_n(F)}{\sigma^2\Psi(B)\Psi(F)}\right]x_t = \nu(B, F)x_t \tag{4}$$

where F is a forward-shift operator (i.e. $F^i x_t = x_{t+i}$). The estimator given by (4) is called historical estimator. The WK filter can be expressed after simplification as

$$\nu(B, F) = \frac{\sigma_n^2}{\sigma^2}\frac{\Theta_n(B)\,\Phi_s(B)\Theta_n(F)\,\Phi_s(F)}{\Theta(B)\,\Theta(F)} \tag{5}$$

From (1), (4) and (5) it can be obtained

$$\Phi_n(B)\hat{n}_t = \Theta_n(B)\frac{\Theta_n(F)\Phi_s(F)\sigma_n^2}{\Theta(F)\sigma^2}a_t. \tag{6}$$

It is clear that the process for the SA component (2) is different from the process for its historical estimator (6). If the process (1) is seasonally stationary, the seasonal component (3) will not contain the seasonal unit roots and, as a result, the polynomial $\Phi_s(F)$ will be stationary. In this case $\hat{n}_t$ given by (6) is going to be invertible. If the seasonal component is a nonstationary I(1) process (i.e. it contains unit roots at seasonal frequencies) $\Phi_s(F) = 1 + F + F^2 + \cdots + F^{r-1} = S(F)$, these unit roots will show up as MA in the model generating $\hat{n}_t$ and will produce spectral zeros for the associated seasonal frequencies.

Thus, if the seasonal component identified within AMB approach is nonstationary, the historical estimator of the series will not be invertible. An important implication of this result according to Gomez and Maravall (2001) is that the estimator of the SA series "…*will not accept, in general, an AR (or VAR) approximation to its Wold representation*".[2]

3 Seasonal fractional integration

The AMB approach assumes that the data follows an ARIMA-type process. This assumption restricts the DGP to be either stationary I(0) or, alternatively integrated of order one, I(1), at zero and/or the seasonal frequencies. In this article, we extend the seasonal I(1)/I(0) approach to the fractional case and examine cases where the original series have noninteger orders of integration at seasonal frequencies. In such a case, the process is said to be seasonally fractionally integrated or seasonal I(d).

For the purpose of the present work, we first define an I(0) process as a covariance stationary process with a positive and bounded spectral density at all frequencies in the spectrum. Then, we say that a process x_t is seasonal I(d) if it can be represented as:

$$(1 - B^\tau)^d x_t = a_t, \tag{7}$$

where B^τ is the seasonal lag operator (i.e., $B^\tau x_t = x_{t-\tau}$) and τ represents the number of periods per year (e.g., $\tau = 4$ with quarterly data, $\tau = 12$ in case of monthly data, etc.), d is a real value and a_t is an I(0) process that may include seasonal and nonseasonal weakly autocorrelated (e.g., ARMA) terms. If $d > 0$ in (7), x_t is said to be a seasonal long memory process, so-named because of the strong degree of association between observations widely (seasonally) separated in time. However, the specification in (7) is rather restrictive in the sense that it imposes the same degree of integration at all frequencies, noting that $(1 - B^\tau)$ can be decomposed into $(1 - B)S(B)$ where $S(B) = 1 + B + B^2 + \cdots + B^{\tau-1}$ refers exclusively to the seasonal frequencies. Thus, for example, in the case of the polynomial $\left(1 - B^4\right)^d$, it can be expressed as $(1 - B)^d \left(1 + B + B^2 + B^3\right)^d = (1 - B)^d \left(1 + B\right)^d \left(1 + B^2\right)^d$, implying the same degree of integration, d, at zero and the seasonal frequencies $\pi, \pi/2$ $(3\pi/2)$ (of a 2π cycle). Extending this model, we may consider a more general specification that permits different degrees of integration at each of the frequencies. In particular, for the case of quarterly data, in the paper we will examine models of the form:

$$(1 - B)^{d_0}(1 + B)^{d_2}(1 + B^2)^{d_1} x_t = a_t, \tag{8}$$

where d_0 refers to the order of integration at the long run or zero frequency; d_2 is the order of integration at the semiannual frequency π, and d_1 corresponds to the annual frequencies $\pi/2$ and $3\pi/2$. Applications using the flexible model (8) can be found in

---

[2] It is applied to the historical or final estimator of the SA series. However, only the central observations of the SA series are produced by the historical estimator. For the periods close to the beginning or the end, the filter cannot be completed and some preliminary estimator has to be used. If the series is long its properties are dominated by the final estimator. However, if it is short it will have a highly nonlinear structure.

Arteche and Robinson (2000), Gil-Alana and Robinson (2001), Arteche (2003) and
Hassler et al. (2009).

If the true process for (quarterly) data is fractionally integrated and it is given by
(8), then TSW will find the best possible integer framework approximation to model
seasonality, which can be stationary or not.

Theoretically, the process for the historical estimator of the SA series has the fol-
lowing form:

$$\nabla^{d_0} \hat{N}_t = \frac{\sigma_n^2}{\sigma^2} \frac{\Theta_n(B)\Phi_s(B)}{\Theta(B)} \frac{\Theta_n(F)\Phi_s(F)}{\Theta(F)} (1+B)^{-d_2} \left(1+B^2\right)^{-d_1} a_t \quad (9)$$

where $\nabla = 1 - B$. If we separate seasonal and non-seasonal MA roots in the pro-
cess, TSW chooses for the data, $\Theta(k) = \Theta_{SS}(k)\Theta_{NN}(k)$, $k = B, F$, then the terms
$\frac{\Phi_s(B)}{\Theta_{SS}(B)}$ and $\frac{\Phi_s(F)}{\Theta_{SS}(F)}$ are parts of the two-sided WK filter (with backward and for-
ward operators correspondingly) aimed to subtract the seasonal component given by
$(1+B)^{d_2}\left(1+B^2\right)^{d_1}$. In other words, $\frac{\Phi_s(B)}{\Theta_{SS}(B)}$ is the TSW integer framework approx-
imation to the fractionally ingenerated seasonal component.

Note that the (pseudo) spectrum of the process in (9) is the same as that of

$$\nabla^{d_0} \breve{N}_t = \frac{\sigma_n^2}{\sigma^2} \frac{\Theta_n^2(B)}{\Theta_{NN}^2(B)} \left[\frac{\Phi_s(B)}{\Theta_{SS}(B)}\right]^2 (1+B)^{-d_2} \left(1+B^2\right)^{-d_1} \alpha_t$$

with $\text{var}(a_t) = \text{var}(\alpha_t)$.[3]

If TSW chooses a nonstationary SARIMA to fit the data, then $\Phi_s(B) = (1+B)\left(1+B^2\right)$ and therefore:

$$\nabla^{d_0} \breve{N}_t = \frac{\sigma_n^2}{\sigma^2} \frac{\Theta_n^2(B)}{\Theta_{NN}^2(B)} \frac{(1+B)^{2-d_2}\left(1+B^2\right)^{2-d_1}}{\Theta_{SS}^2(B)} \alpha_t \quad (10)$$

If $d_1, d_2 < 1$, the adjusted series should have a seasonal fractionally integrated MA
polynomial with coefficients larger than 1, thus being, hypothetically, not invertible.
In the following section, we investigate this issue in practice by means of simulations.

4 Simulation study

4.1 Simulation set-up

To study invertibility of the time series adjusted by TSW, we generate quarterly data
from different specifications of seasonal fractionally integrated ARIMA (SARFIMA)
models. The parameters for the simulated SARFIMA processes of the form as in (8)

---

[3] We analyse the process (10) instead of (9) because of two reasons. First, in empirical econometrics, two-sided processes are very rare and are not estimated in practice. Second, given that the estimation techniques employed in the paper are based on the (pseudo) spectrum of a process, then (10) cannot be virtually distinguished from (9).

are $d_0 = \{0.3, 0.7, 1, 1.5\}$, $d_i = \{0.1, 0.3, 0.5, 0.7\}$, $i = \{1, 2\}$, and $\sigma^2 = 1$. For $d_0 = 1$ we simulate additionally $d_i = 1$. The choice of the values is justified by the empirical evidence. The number of observations for each series is set $T = 500$.

To generate the data, the long memory polynomials in (8) have to be expanded. We choose the lag truncation 1000 for each of the three polynomials $(1 - B)^{d_0}$, $(1 + B^2)^{d_1}$, and $(1 + B)^{d_2}$. Thereafter, we multiply the expanded long memory polynomials and, following Bhardwaj and Swanson (2006), we truncate the resulting polynomial when the coefficients become smaller than $1.0e-004$ (the truncation lag is always smaller than 1000). All observations are generated using standard normal errors. For each process and each replication, we generate 3000 observations and we use just the last T observations to avoid the initial values problem, which is especially important when taking into account the long-memory properties of the DGP.

To each simulated series, we apply TSW. If TSW chooses a seasonally nonstationary ARIMA model for this series, we collect it for the future analysis. If the model chosen by TSW contains stationary seasonality we discard the simulated series. We proceed until we have $I = 500$ simulated series for each specification identified by TSW as seasonally nonstationary.

In addition to SARFIMA, we produce simulations for a set of quarterly Airline models of Box and Jenkins (1970) which are believed to approximate reasonably well the stochastic properties of many series

$$(1 - B)(1 - B^\tau)x_t = (1 + Q_1 B)(1 + Q_\tau B^\tau)a_t \tag{11}$$

with $\tau = 4$ and negative values for Q_1 and Q_4: $Q_i = \{-0.8, -0.6, -0.3\}$, $i = 1, 4$ and $[Q_1, Q_4] = \{[-1, -0.8], [-1, -0.6], [-1, -0.3]\}$. In the same way as for SARFIMA, we collect $I = 500$ series for each specification identified by TSW as seasonally nonstationary.

Thereafter, each selected series for each specification is adjusted by TSW and coefficients of FI at seasonal frequencies are estimated.[4] It is important to remark that, even if several series are simulated from the same SARFIMA specification, the TSW may choose distinct SARIMA models for each of the simulated series. Moreover, even if the model chosen is the same, the estimated SARIMA parameters may be very different. Since the seasonal filters applied to the data are based on the identified SARIMA model, different filters may be applied to each of the series simulated from the same SARFIMA specification. In this way, the mean of the estimated parameters of FI at seasonal frequencies of the adjusted series does not have statistical meaning. Therefore, to build conclusions on the invertibility of the adjusted series we propose the following testing procedure. After estimating the FI parameters at seasonal frequencies, we test if we can reject the null hypothesis $d_i \geq -0.5$ in favor of the alternative $d_i < -0.5$ at least for one of the two parameters of seasonal FI. If this is the case, we say that the adjusted series is *statistically noninvertible*. Also, we test if the series is *statistically invertible*, i.e., if we can reject the null hypothesis $d_i \leq -0.5$ in favor of the alternative $d_i > -0.5$ for both estimated parameters of seasonal FI.

---

[4] We exclude 40 observations from the beginning and from the end of series to eliminate nonlinearities produced by preliminary estimator of SA.

As a result, for each Airline and SARFIMA specification, we can compute both the percentage of statistically noninvertible and the percentage of statistically invertible series (in the adjusted $I = 500$ series chosen by TSW to be seasonally nonstationary before adjustment).

To estimate the coefficients of FI at the seasonal frequencies, we use the log-periodogram regression with a complex-valued taper proposed by Hurvich and Chen (2000):

$$h_t = 0.5 \left[1 - \exp \left\{ \frac{i2\pi (t - 0.5)}{T} \right\} \right], \quad t = 1, \ldots, T.$$

The choice of the log-periodogram regression is justified by several reasons. Since we do not know what the correct specification after adjustment is, we avoid the parameterization of the whole spectrum by choosing a local estimation method. Tapering is particularly suitable when the estimated coefficients of FI are expected to be negative, possibly smaller than -0.5. In these circumstances, the estimation results based on the nontapered data will have a strong positive bias, making the method not appropriate for the purposes of this work. As Hurvich and Ray (1995) and Velasco (1999) point out, the use of a taper can alleviate the negative effects of overdifferencing, reducing the bias in FI estimates. Finally, tapering also reduces the bias that appears due to contamination of the periodogram from the short memory component of the spectral density and allows for a less restrictive trimming of frequencies in presence of asymmetries, as it happens at the frequency $\pi/2$. A comprehensive discussion of the performance of the method for seasonal and cyclical time series with asymmetric long memory properties is presented in Arteche and Velasco (2005). Nevertheless, we also perform a small Monte Carlo study to check the performance of the estimation method in the presence of negative seasonal fractional integration at seasonal frequencies in the following way. For each specification, after simulating the data and before applying the TSW (i.e. when we still know the true DGP), we take yearly differences, making sure that the resulting series are over-differenced at seasonal frequencies having negative FI coefficients. We estimate these coefficients computing the mean for each specification to assess the estimation bias.

Also, note that the classification of an adjusted series as statistically noninvertible or statistically invertible requires the one-tailed testing of two different seasonal FI parameters (one for each seasonal frequency), and thus is subject to a multiple-testing problem. Given that independence of the estimates at different frequencies is satisfied asymptotically, we can control for the experiment-wide type-I error.[5] Statistical noninvertibility requires the rejection of at least one of the two nulls, so it faces the increase in type I error that occurs when statistical tests are used repeatedly. If two independent tests are performed, the experiment-wide significance level is given

---

[5] The construction of a joint test is difficult. Note that even if one assumes independence, the testing procedures would still require a one-tailed joint test. Wald tests based on chi-squared distributions cannot easily accommodate directional tests. More important, *statistically invertibility* requires the rejection of the null $H_0 : d_i \leq -0.5$ for the two seasonal FI parameters. In a standard joint test, the alternative would take the form $H_1 : d_1 > -0.5$ or $d_2 > -0.5$ thus the rejection of the null does not guarantee statistical invertibility.

by $\bar{\alpha} = 1 - (1 - \alpha)^2$, being α the significance level employed testing at each frequency (Sidak 1967). Thus, an $\bar{\alpha} = 10\%$ implies the use of the significance level $\alpha = 5.01\% \approx 5\%$ at individual tests. As for statistical invertibility, it requires the rejection of exactly the two one-tailed nulls. Although the experiment-wide significance level cannot be computed exactly, it can be bounded by α under independence. Note that the process would be incorrectly classified as statistically invertible if the two nulls are rejected when in fact, either both hypotheses are true or only one is true. In the former case, the experiment-wide significance level under independence is $\bar{\alpha} = \alpha^2$. In the latter case, the experiment-wide significance level can be bounded as $\alpha^2 \leq \bar{\alpha} = \alpha(1 - \beta) \leq \alpha$, where β is the probability of type II error. Therefore, we remain conservative by testing each frequency at significance level $\alpha = 10\%$, ensuring that the experiment-wide significance level $\bar{\alpha}$ does not exceed 10%.

All the simulations and estimations were produced in Matlab. The programs are available from the authors upon request. For seasonal adjustment, we use the last release of the TSW for Matlab developed by the Bank of Spain.

4.2 Simulation results

The results of the simulation study for the different Airline and SARFIMA specifications are presented in Tables 1 and 2 respectively. In both tables, the particular specification, chosen as DGP, appears in the first column (i.e., the values for the MA parameters Q_1 and Q_4 for the Airline model and the parameters of FI at seasonal frequencies d_1 and d_2 for SARFIMA).

4.2.1 Monte Carlo results for the tapered log-periodogram regression for seasonal frequencies with negative FI

The Monte Carlo study is presented in the columns 2 and 3 of Tables 1 and 2. The coefficients in the table are $\hat{d}_i = 1 + \hat{d}_i^*$, where $\hat{d}_i^*$ denotes the mean of the estimates obtained with yearly differenced data.[6]

The estimated parameters are slightly positively biased. The bias is higher when the negative FI is greater in absolute value. Thus, the highest bias is observed for the process $\{d_0 = 0.3, d_1 = 0.1, d_2 = 0.1\}$. Recall that after taking yearly difference it becomes $\{d_0^* = -0.7, d_1^* = -0.9, d_2^* = -0.9\}$.

The simulation results for the Airline specifications are useful to study the performance of the method in the presence of short memory components. Results are presented in Table 1. The estimation method always detects seasonal unit roots. The precision of the estimates depends on the value of the seasonal MA parameter Q_4: the larger this value in absolute terms, the larger is the bias.

---

[6] Recall that the original data was yearly differenced in the Monte Carlo analysis to ensure that the resulting series are overdifferenced at seasonal frequencies with negative coefficients of fractional integration. For the processes with the fractional order of integration at zero $d_0 = 1.5$, we take a first difference in addition to the yearly difference. The same applies to the Airline model.

Table 1 Simulation results, airline model

Before adjustment	Mean		After adjustment				Median [68 band]	
DGP			%					
Q_1, Q_4	$\hat{d}_1$	$\hat{d}_2$	NS	I	SNI	SI	$\tilde{d}_1$	$\tilde{d}_2$
{−0.8,−0.3}	1.043	1.062	100.0	16.4	81.0	10.6	−0.790 [−0.734, −0.883]	−1.004 [−0.843, −1.100]
{−0.8,−0.6}	1.124	1.089	100.0	14.6	74.8	11.0	−0.713 [−0.613, −0.801]	−1.102 [−0.828, −1.258]
{−0.8,−0.8}	1.195	1.112	100.0	10.8	86.6	9.8	−0.760 [−0.677, −0.980]	−1.214 [−1.069, −1.320]
{−0.6,−0.3}	1.040	1.071	100.0	6.2	92.6	4.8	−0.827 [−0.778, −0.870]	−1.069 [−0.982, −1.133]
{−0.6,−0.6}	1.113	1.110	100.0	25.2	72.4	20.2	−0.740 [−0.591, −0.814]	−1.067 [−0.787, −1.213]
{−0.6,−0.8}	1.203	1.096	100.0	19.5	68.0	14.0	−0.648 [−0.530, −0.730]	−1.002 [−0.726, −1.327]
{−0.3,−0.3}	1.024	1.056	100.0	4.0	94.2	3.4	−0.837 [−0.791, −0.878]	−0.963 [−0.878, −1.036]
{−0.3,−0.6}	1.116	1.071	100.0	6.0	92.1	4.4	−0.785 [−0.732, −0.859]	−0.909 [−0.774, −1.118]
{−0.3,−0.8}	1.192	1.083	100.0	15.0	62.8	7.8	−0.601 [−0.515, −0.675]	−0.766 [−0.638, −0.902]
{−1,−0.3}	1.025	1.064	100.0	4.0	94.6	2.6	−0.889 [−0.839, −0.935]	−1.002 [−0.943, −1.066]
{−1,−0.6}	1.114	1.089	100.0	4.2	95.4	4.0	−1.008 [−0.963, −1.044]	−1.013 [−0.950, −1.075]
{−1,−0.8}	1.213	1.106	100.0	3.8	95.2	3.4	−1.096 [−1.045, −1.137]	−1.008 [−0.944, −1.074]

In the Airline model the true seasonal FI parameters before adjustment are $d_1 = d_2 = 1$. The bandwidth parameter employed in the estimation is the square root of T

Table 2 Simulation results, SARFIMA model

Before adjustment			After adjustment					
DGP	Mean		%				$\tilde{d}_1$	$\tilde{d}_2$
d_1, d_2	$\hat{d}_1$	$\hat{d}_2$	NS	I	SNI	SI		
$d_0 = 0.3$								
{0.1,0.1}	0.173	0.191	31.7	99.4	0.2	91.4	−0.116 [−0.041,−0.178]	−0.096 [−0.028,−0.162]
{0.1,0.3}	0.166	0.366	49.1	97.0	0.2	78.6	−0.240 [−0.176,−0.285]	−0.119 [−0.055,−0.180]
{0.1,0.5}	0.173	0.539	58.1	84.6	1.0	46.0	−0.272 [−0.212,−0.341]	−0.273 [−0.207,−0.356]
{0.1,0.7}	0.168	0.734	35.8	45.0	10.6	11.0	−0.258 [−0.192,−0.327]	−0.511 [−0.434,−0.583]
{0.3,0.1}	0.360	0.169	60.5	86.2	1.0	46.8	−0.200 [−0.147,−0.260]	−0.327 [−0.249,−0.400]
{0.3,0.3}	0.353	0.362	64.7	92.4	0.0	56.0	−0.274 [−0.227,−0.328]	−0.259 [−0.197,−0.335]
{0.3,0.5}	0.348	0.540	80.5	69.8	3.2	30.6	−0.333 [−0.275,−0.404]	−0.302 [−0.231,−0.397]
{0.3,0.7}	0.352	0.731	74.7	27.2	21.4	6.4	−0.340 [−0.272,−0.407]	−0.563 [−0.452,−0.625]
{0.5,0.1}	0.544	0.158	95.8	66.8	6.4	26.2	−0.301 [−0.257,−0.349]	−0.403 [−0.312,−0.497]
{0.5,0.3}	0.529	0.342	94.9	68.2	4.8	22.6	−0.341 [−0.284,−0.401]	−0.399 [−0.329,−0.481]
{0.5,0.5}	0.533	0.533	96.9	63.4	6.0	17.2	−0.401 [−0.339,−0.457]	−0.404 [−0.333,−0.476]
{0.5,0.7}	0.530	0.730	99.6	32.8	23.4	10.2	−0.478 [−0.397,−0.561]	−0.440 [−0.375,−0.531]
{0.7,0.1}	0.741	0.156	100.0	38.8	20.0	11.6	−0.476 [−0.420,−0.526]	−0.465 [−0.340,−0.574]
{0.7,0.3}	0.736	0.332	100.0	30.0	21.0	10.4	−0.489 [−0.427,−0.548]	−0.531 [−0.447,−0.613]
{0.7,0.5}	0.720	0.530	99.8	27.0	26.0	8.0	−0.518 [−0.461,−0.575]	−0.536 [−0.456,−0.628]
{0.7,0.7}	0.721	0.708	99.8	13.2	42.0	7.6	−0.590 [−0.528,−0.658]	−0.598 [−0.519,−0.680]
$d_0 = 0.7$								
{0.1,0.1}	0.184	0.187	16.7	98.2	0.2	91.0	−0.031 [0.012,−0.096]	−0.046 [0.014,−0.118]
{0.1,0.3}	0.174	0.363	22.1	96.2	0.2	86.6	−0.116 [−0.056,−0.172]	−0.069 [0.004,−0.155]
{0.1,0.5}	0.142	0.569	72.4	81.0	1.6	41.2	−0.186 [−0.104,−0.269]	−0.298 [−0.237,−0.373]
{0.1,0.7}	0.145	0.745	77.8	47.8	13.6	18.6	−0.290 [−0.181,−0.389]	−0.490 [−0.410,−0.575]

Table 2 continued

DGP	Before adjustment Mean		After adjustment %					
d_1, d_2	$\hat{d}_1$	$\hat{d}_2$	NS	I	SNI	SI	$\tilde{d}_1$	$\tilde{d}_2$
{0.3,0.1}	0.358	0.164	73.3	76.4	6.8	42.2	−0.236 [−0.172,−0.300]	−0.335 [−0.238,−0.421]
{0.3,0.3}	0.348	0.349	76.9	80.8	3.4	51.0	−0.286 [−0.201,−0.368]	−0.278 [−0.200,−0.380]
{0.3,0.5}	0.335	0.534	74.2	73.6	6.2	39.0	−0.275 [−0.189,−0.364]	−0.325 [−0.263,−0.400]
{0.3,0.7}	0.332	0.731	97.3	28.2	29.6	11.0	−0.527 [−0.412,−0.628]	−0.488 [−0.401,−0.560]
{0.5,0.1}	0.534	0.159	98.2	52.0	13.6	19.2	−0.400 [−0.320,−0.475]	−0.457 [−0.373,−0.556]
{0.5,0.3}	0.522	0.335	99.2	30.0	33.2	12.2	−0.465 [−0.395,−0.548]	−0.585 [−0.475,−0.680]
{0.5,0.5}	0.527	0.532	99.0	31.0	22.4	9.2	−0.524 [−0.453,−0.590]	−0.493 [−0.424,−0.575]
{0.5,0.7}	0.524	0.726	99.2	17.2	52.8	8.8	−0.638 [−0.549,−0.727]	−0.535 [−0.457,−0.617]
{0.7,0.1}	0.728	0.142	100.0	22.2	51.4	10.4	−0.617 [−0.523,−0.731]	−0.457 [−0.327,−0.618]
{0.7,0.3}	0.717	0.324	100.0	13.6	52.2	4.8	−0.590 [−0.523,−0.652]	−0.590 [−0.451,−0.736]
{0.7,0.5}	0.712	0.522	100.0	8.0	60.2	3.0	−0.610 [−0.563,−0.678]	−0.689 [−0.583,−0.770]
{0.7,0.7}	0.712	0.718	100.0	5.6	69.6	3.8	−0.677 [−0.625,−0.731]	−0.639 [−0.569,−0.715]
$d_0 = 1$								
{0.1,0.1}	0.164	0.176	24.8	100.0	0.0	98.4	−0.042 [−0.003,−0.083]	−0.035 [0.019,−0.095]
{0.1,0.3}	0.177	0.356	26.3	98.6	0.0	90.8	−0.134 [−0.072,−0.191]	−0.025 [0.038,−0.089]
{0.1,0.5}	0.165	0.545	28.9	88.8	1.0	56.4	−0.098 [−0.045,−0.155]	−0.280 [−0.175,−0.375]
{0.1,0.7}	0.148	0.748	57.6	38.2	21.4	15.2	−0.086 [−0.034,−0.173]	−0.562 [−0.453,−0.658]
{0.1,1.0}	0.117	1.037	64.6	9.6	83.8	9.0	−0.056 [−0.005,−0.110]	−0.876 [−0.814,−0.937]
{0.3,0.1}	0.355	0.168	58.1	94.0	1.0	71.0	−0.145 [−0.093,−0.195]	−0.212 [−0.143,−0.295]
{0.3,0.3}	0.335	0.348	70.0	92.4	1.2	70.0	−0.211 [−0.144,−0.262]	−0.214 [−0.139,−0.288]
{0.3,0.5}	0.339	0.533	74.4	86.8	1.2	48.2	−0.211 [−0.140,−0.297]	−0.280 [−0.215,−0.345]

Table 2 continued

Before adjustment			After adjustment					
DGP	Mean		%					
d_1, d_2	$\hat{d}_1$	$\hat{d}_2$	NS	I	SNI	SI	$\tilde{d}_1$	$\tilde{d}_2$
{0.3,0.7}	0.328	0.728	89.9	38.6	13.2	12.6	−0.208 [−0.145,−0.281]	−0.514 [−0.438,−0.578]
{0.3,1.0}	0.316	1.037	91.4	18.2	71.2	14.0	−0.459 [−0.244,−0.607]	−0.799 [−0.715,−0.888]
{0.5,0.1}	0.533	0.159	96.2	69.0	8.8	29.8	−0.355 [−0.286,−0.415]	−0.351 [−0.247,−0.443]
{0.5,0.3}	0.527	0.338	97.8	40.4	24.4	14.8	−0.474 [−0.379,−0.561]	−0.503 [−0.394,−0.607]
{0.5,0.5}	0.514	0.530	98.0	42.6	26.4	14.2	−0.484 [−0.405,−0.580]	−0.483 [−0.381,−0.566]
{0.5,0.7}	0.516	0.722	98.4	35.8	26.4	11.0	−0.386 [−0.323,−0.513]	−0.530 [−0.452,−0.603]
{0.5,1.0}	0.505	1.017	100.0	11.4	85.8	9.2	−0.768 [−0.496,−0.866]	−0.812 [−0.727,−0.887]
{0.7,0.1}	0.728	0.142	100.0	29.6	32.0	11.8	−0.557 [−0.485,−0.628]	−0.351 [−0.221,−0.460]
{0.7,0.3}	0.711	0.327	100.0	15.8	59.6	7.6	−0.654 [−0.578,−0.727]	−0.630 [−0.513,−0.722]
{0.7,0.5}	0.714	0.516	100.0	15.2	72.2	6.6	−0.721 [−0.633,−0.801]	−0.792 [−0.637,−0.916]
{0.7,0.7}	0.700	0.716	100.0	8.2	76.8	3.4	−0.773 [−0.670,−0.851]	−0.727 [−0.637,−0.820]
{0.7,1.0}	0.699	1.010	100.0	8.6	89.4	5.8	−0.942 [−0.866,−1.003]	−0.824 [−0.757,−0.895]
{1.0,0.1}	1.017	0.117	100.0	7.6	91.0	6.8	−0.947 [−0.888,−0.992]	−0.153 [−0.66,−0.254]
{1.0,0.3}	1.009	0.314	100.0	8.6	90.4	5.0	−0.944 [−0.901,−0.989]	−0.403 [−0.278,−0.587]
{1.0,0.5}	1.004	0.506	100.0	6.2	93.6	4.8	−0.942 [−0.900,−0.979]	−0.596 [−0.479,−0.738]
{1.0,0.7}	1.010	0.691	100.0	7.2	90.8	4.6	−0.914 [−0.867,−0.968]	−0.795 [−0.655,−0.983]
{1.0,1.0}	1.005	0.991	100.0	6.4	93.0	4.6	−0.919 [−0.878,−0.962]	−0.899 [−0.833,−0.956]
$d_0 = 1.5$								
{0.1,0.1}	0.177	0.188	18.3	94.8	0.6	76.4	−0.155 [−0.104,−0.215]	−0.171 [−0.095,−0.248]
{0.1,0.3}	0.142	0.367	49.1	96.2	0.4	77.0	−0.208 [−0.142,−0.269]	−0.147 [−0.083,−0.220]

Table 2 continued

| DGP | Before adjustment Mean | | After adjustment % | | | | | |
d_1, d_2	$\hat{d}_1$	$\hat{d}_2$	NS	I	SNI	SI	$\tilde{d}_1$	$\tilde{d}_2$
{0.1,0.5}	0.137	0.541	72.5	82.0	4.0	41.0	−0.307 [−0.239,−0.375]	−0.279 [−0.214,−0.363]
{0.1,0.7}	0.132	0.739	69.2	55.2	9.4	19.2	−0.304 [−0.245,−0.372]	−0.464 [−0.384,−0.548]
{0.3,0.1}	0.332	0.153	75.1	63.0	7.6	31.8	−0.269 [−0.200,−0.336]	−0.410 [−0.302,−0.531]
{0.3,0.3}	0.331	0.343	69.8	73.4	2.6	34.0	−0.345 [−0.277,−0.415]	−0.341 [−0.279,−0.427]
{0.3,0.5}	0.325	0.535	84.5	63.4	6.6	20.2	−0.408 [−0.359,−0.475]	−0.358 [−0.285,−0.437]
{0.3,0.7}	0.318	0.728	80.4	38.6	24.4	12.8	−0.430 [−0.356,−0.502]	−0.486 [−0.390,−0.622]
{0.5,0.1}	0.519	0.147	89.6	53.6	19.2	19.2	−0.344 [−0.272,−0.410]	−0.444 [−0.322,−0.573]
{0.5,0.3}	0.516	0.336	96.5	49.2	15.4	13.0	−0.383 [−0.330,−0.433]	−0.489 [−0.404,−0.576]
{0.5,0.5}	0.505	0.526	96.3	50.8	19.8	8.8	−0.427 [−0.382,−0.481]	−0.441 [−0.380,−0.528]
{0.5,0.7}	0.502	0.717	97.8	31.4	20.8	6.8	−0.497 [−0.435,−0.560]	−0.477 [−0.391,−0.565]
{0.7,0.1}	0.709	0.134	99.6	32.4	41.8	14.6	−0.479 [−0.388,−0.574]	−0.436 [−0.241,−0.682]
{0.7,0.3}	0.705	0.323	99.8	26.8	39.4	7.4	−0.456 [−0.386,−0.554]	−0.582 [−0.444,−0.749]
{0.7,0.5}	0.703	0.508	100.0	20.8	35	5.0	−0.488 [−0.427,−0.546]	−0.601 [−0.516,−0.705]
{0.7,0.7}	0.700	0.717	100.0	16.4	37.8	4.8	−0.547 [−0.490,−0.616]	−0.605 [−0.510,−0.681]

The bandwidth parameter employed in the estimation is the square root of T

Finally, note that no matter the specification, the coefficient d_1 tends to be estimated less precisely than the coefficient d_2 due to the asymmetries presented in the periodogram around frequency $\pi/2$.

Overall, the performance of the method at seasonal frequencies is similar to its performance at frequency zero, documented in previous studies. The log-periodogram regression with tapered data performs well for negative seasonal FI even for coefficients from the noninvertible region and also for the estimation of the parameter at frequency $\pi/2$, where the spectrum is not symmetric. The method works also well in the presence of short memory components, as shown in Table 1. Thus, the results from the Monte Carlo study confirm that the log-periodogram regression with tapered data is appropriate for the purposes of the present work.

4.2.2 Results on the invertibility of the adjusted series

The column four of Tables 1 and 2 presents the percentage of cases for which TSW chooses a seasonally nonstationary model to fit the data for each of the simulated processes (NS). As can be seen in Table 1, TSW always chooses a nonstationary model when the true DGP follows the Airline model. As expected for the SARFIMA specifications (Table 2) this percentage increases together with the magnitude of both d_1 and d_2.

Next three columns of Tables 1 and 2 contain results of the statistical testing described in the simulation set-up. Column five presents the percentage of cases in which the seasonally adjusted processes are estimated invertible (I), i.e., with the two estimated coefficients of FI at seasonal frequencies greater than -0.5. Thereafter we compute the percentage of replications in which the SA series have at least one estimated coefficient of seasonal FI statistically smaller than -0.5—that is to say, the series is statistically noninvertible (SNI). Thereafter, we test statistical invertibility: both estimated coefficients are statistically greater than -0.5. The percentage of statistically invertible results is given in column seven (SI). When the data are simulated from the Airline model (Table 1), the estimated coefficients of FI at frequencies $\pi/2$ ($3\pi/2$) and π are almost always smaller than -0.5, which indicates the (possible) noninvertibility of the corresponding SA series. Moreover, in a high percentage of cases, this noninvertibility is statistically significant. This result is not surprising and it is completely in line with the implications of the TSW for this class of models (Gomez and Maravall 2001). For the SARFIMA specifications (Table 2), the result of the application of TSW depends on the initial properties of the simulated data. Thus, if the two coefficients of the seasonal FI are within the stationary region ($d_i < 0.5$, $i = \{1, 2\}$), even if the TSW identifies a nonstationary seasonal model (this occurs in a relatively small percentage of cases), the estimated coefficients of seasonal FI of the SA series are greater than -0.5 in most of the cases. Only a very small percentage of series are (possibly) statistically noninvertible. The percentage of statistically invertible results decreases as the seasonal FI coefficients of the original series approach the nonstationary region. For example, for $d_0 = 0.3$, if the original series have both coefficients of seasonal FI $d_i = 0.1$, TSW only selects a seasonal nonstationary representation in 32% of the cases. In addition, even if this is the case and a nonstationary SARIMA is chosen, the estimated coefficients of the SA series almost always lie (99%) in the invertible

region. Moreover, in 91% of the cases both parameters are statistically greater than -0.5 and the series are statistically invertible. The percentage of statistically invertible results decreases to 45% if d_1 and d_2 are equal to 0.3. Still, the estimated parameters are greater than -0.5 in 92 % of the cases and the percentage of statistically invertible results is 56%.

On the contrary, if one of the coefficients of seasonal FI in the DGP is greater than 0.5, TSW fits a seasonally nonstationary model a higher percentage of times, and for these cases, the SA series are often estimated to be noninvertible. Once more, the percentage of statistically noninvertible results increases with the parameters of seasonal FI of the original series. Thus, (again for $d_0 = 0.3$) if $d_i = 0.7$ for the two coefficients, TSW selects a nonstationary representation in almost 100% of the cases. Only in 13.2% of the cases, the SA series are invertible (and only in 7.6% the invertibility is statistically significant) whereas in 42% of the replicas the SA series were found to be statistically noninvertible. It is also interesting to note that, although the parameter of fractional integration at zero is not neutral, the same conclusions are obtained for all simulated d_0.

If d_1 and d_2 in the DGP are very different in magnitude, the TSW chooses the model with nonstationary seasonality for the process with higher d_1 more often. After the adjustment, these processes present lower percentage of invertible and statistically invertible cases. For example, if parameters in the DGP are $d_0 = 0.3$, $d_1 = 0.1$ and $d_2 = 0.5$, the TSW chooses models with non-stationary seasonality 58.1% of times, while this percentage reaches 96% when $d_1 = 0.5$ and $d_2 = 0.1$. The SA data is invertible in 85% of the cases (and statistically invertible in 46%) for the initial process with $d_1 = 0.1$ and $d_2 = 0.5$, and in 67% (26%) of the cases for data generated with $d_1 = 0.5$ and $d_2 = 0.1$. Thus, we find that the TSW seems to react strongly to long memory at the yearly frequency.

The previous results indicate that invertibility may not be a severe issue in many circumstances. However, they contradict the theoretical findings derived in Sect. 3. Recall that the simulations are based on data identified by TSW as seasonally nonstationary and, according to the Eq. (10), the adjusted series should contain seasonal unit MA roots. To aid in the explanation of this apparent puzzle, the columns eight and nine in Tables 1 and 2 report the median with the 16th and 84th percentiles (68% band) of the estimated seasonal FI parameters after adjustment. Several results emerge from these columns:

1. If the DGP contains stationary seasonality with relatively low parameters of FI, the adjusted series is not only usually estimated (statistically) invertible, but also the estimated parameters of FI are indeed very small in magnitude, with no sign of unit MA roots. For example, if the data is generated from an SARFIMA with $d_0 = 0.3$, $d_1 = 0.1$ and $d_2 = 0.1$, the median of the estimated values after adjustment is $\tilde{d}_1 = -0.116$ and $\tilde{d}_2 = -0.096$.
2. If the DGP contains seasonal FI with equal parameters ($d_1 = d_2$), the medians of parameter estimates after adjustment are also close to each other and very similar to the values employed to generate the data but with opposite sign.
3. If d_1 and d_2 in the DGP are very different in magnitude, the median values after adjustment are negative but smaller in magnitude than the larger parameter in the

Table 3 Model for the signal and its approximation

SARFIMA: $(1-B)^{2-0.8-0.7^{1/4}}(1+B)^{1-0.7^{1/2}}\left(1+B^2\right)^{1-0.7^{1/4}}x_t = \varepsilon_{1,t};\ \text{var}\left(\varepsilon_{1,t}\right)=1,$

Airline: $\nabla\nabla^4 x_t = (1-0.8B)\left(1-0.7B^4\right)a_t,\ \text{var}\left(a_t\right)=1,$

Models for the separate components

$x_t = n_t + s_t$

$\left(1+B+B^2+B^3\right)s_t = (1-B)(1+0.8290B+0.3519B)a_{st},\ \text{var}\left(a_{st}\right)=0.0167$

$\nabla^2 n_t = (1-0.9147B)(1-0.8000B)a_{nt},\ \text{var}\left(a_{nt}\right)=0.7733$

$\nabla^{0.8}\overset{\smile}{N}_t = \left[\dfrac{(1-0.9147B)(1-0.8000B)}{(1-0.8B)\left(1-0.7^{1/4}B\right)}\right]^2\left[\dfrac{(1+B)\left(1+B^2\right)}{\left(1+0.7^{1/4}B\right)\left(1+0.7^{1/2}B^2\right)}\right]^2\Psi$

$\Psi = (1+B)^{-\left(1-0.7^{1/2}\right)}\left(1+B^2\right)^{-\left(1-0.7^{1/4}\right)}\dfrac{0.7733}{1}\varepsilon_t$

Approximation

$\nabla^{0.8}\overset{\smile}{N}_t^* = \left[\dfrac{(1-0.9147B)(1-0.8000B)}{(1-0.8B)\left(1-0.7^{1/4}B\right)}\right]^2(1+B)^{1-0.7^{1/2}}\left(1+B^2\right)^{1-0.7^{1/4}}\dfrac{0.7733}{1}\varepsilon_t'$

DGP. For example, if parameters in the DGP are $d_0 = 0.3$, $d_1 = 0.7$ and $d_2 = 0.1$, the median estimated values after adjustment are $\tilde{d}_1 = -0.476$ and $\tilde{d}_2 = -0.465$.

The most likely explanation for these results is that for a given data length, frequency and bandwidth employed in the simulations, the integer framework process resulting from TSW seasonal adjustment is difficult to distinguish from FI integration at seasonal frequencies, i.e., $\left[\dfrac{\Phi_s(B)}{\Theta_{SS}(B)}\right]^2(1+B)^{-d_2}\left(1+B^2\right)^{-d_1}\approx(1+B)^{\delta_2}\left(1+B^2\right)^{\delta_1}$.[7] Therefore (10) can be approximated by:

$$\nabla^{d_0}\overset{\smile}{N}_t^* = \frac{\sigma_n^2}{\sigma^2}\frac{\Theta_n^2(B)}{\Theta_{NN}^2(B)}(1+B)^{\delta_2}\left(1+B^2\right)^{\delta_1}\alpha_t \qquad (12)$$

To illustrate the approximation of (10) by (12), we provide a numerical example in Table 3.

Consider that the true DGP is an SARFIMA model with stationary FI at seasonal frequencies and that TSW identifies an Airline model to fit the data. The respective parameterizations appear in the first part of Table 3.[8] The table also shows the models for the separate components and the model for the SA data derived by TSW if the program uses the Airline specifications in the table for the seasonal adjustment. Finally,

---

[7] The quality of the approximation depends on the DGP and the model chosen by TSW to fit the data. According to the simulation results, this approximation is expected to be reliable for $d_1, d_2 \in [0, 1]$, which are the values employed in our study. For these values, δ_1 and δ_2 usually belong to the $[0, 1]$ interval (see Table 2).

[8] We have selected the parameterizations of the SARFIMA and Airline models for illustrative purposes. The parameterization of the Airline selected by the TSW depends on the particular realization of the data. Thus, there might be another parameterization of the Airline model that constitutes a better integer approximation to the SARFIMA in the table.
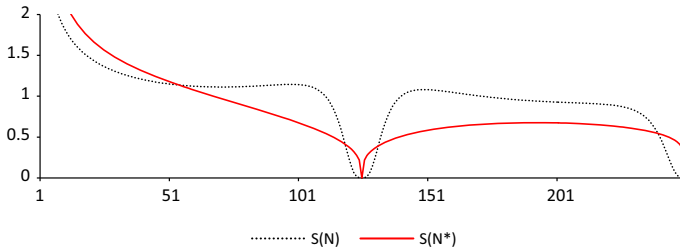
Fig. 1 (pseudo) Spectrums of the signal $S(N)$ and its approximation $S(N^*)$ of the process of the illustrative example of Table 3

the last row of the table presents the approximation $\widbreve{N}_t^{\,*}$ to the historical estimator of the SA series $\widbreve{N}_t$.

Figure 1 plots the (pseudo) spectrums of both, $\widbreve{N}_t$ and $\widbreve{N}_t^{\,*}$. Note that the (pseudo) spectrum of $\widbreve{N}_t$ presents strong dips at the two seasonal frequencies which correspond to the MA roots of a noninvertible process. These dips are also present in the spectrum of the approximation $\widbreve{N}_t^{\,*}$. However, in this last case the dips correspond to invertible negative seasonal fractional integration. To assess the quality of the approximation of $\widbreve{N}_t$ by $\widbreve{N}_t^{\,*}$, we perform the following exercise. From the spectrum of the historical estimator of the SA series, $S\left(\widbreve{N}_t, \omega\right)$, we draw $k = 1, \ldots, 1000$ periodograms $I^{(k)}\left(\widbreve{N}, \omega_j\right) = \frac{1}{2}S\left(\widbreve{N}, \omega_j\right)^* v_j^k$, where $v_j^{(k)}$ are independently and identically distributed χ_2^2 errors, $j = 1, \ldots, T/2 - 1$.[9] If we fit $\widbreve{N}_t^{\,*}$ instead of $\widbreve{N}_t$ to the simulated periodograms, we can test the null hypothesis that $2I^{(k)}\left(\widbreve{N}, \omega_j\right)/S\left(\widbreve{N}^{\,*}, \omega_j\right) = v_j^{(k)*}$ in the neighborhood of the two seasonal frequencies is distributed as χ_2^2. If for a given sample size, frequency and bandwidth ($T = 500$, quarterly data, the same bandwidth as we use in simulations), we cannot reject this null, the processes $\widbreve{N}_t$ and $\widbreve{N}_t^{\,*}$ are statistically not distinguishable around the seasonal frequencies, that is to say that the approximation of (10) by (12) is good. In this example, the null is rejected only a 13% and a 6% of times for the frequencies $\pi/2$ and π respectively.[10] This empirical example suggests that in many cases the noninvertible process resulting from the seasonal adjustment by TSW may not be distinguished in practice from an invertible process with seasonal fractional integration.

This explanation can also accommodate the second and third findings. Our results show that the integer SARIMA framework where TSW operates is quite successful

---

[9] For $j = 0$ and $j = T/2$, v_j is distributed χ_1^2. However frequencies zero and π are not required for this test. See Berkowitz and Diebold (1998) for the details on the bootstrap in frequency domain.

[10] These percentages depend both on T and the bandwidth. For larger T and smaller bandwidth, the null is rejected more often. However, practical applications with $T > 500$ (more than 125 years of data) are rare. A smaller bandwidth implies fewer points for testing, which influences significantly the quality of the test. Thus, it has sense only if T is very large.

approximating processes with equal orders of FI at seasonal frequencies. However, the integer framework is not flexible enough to correctly approximate SARFIMA processes with different (and not related) orders of FI at seasonal frequencies. In these cases, TSW reacts more to the FI coefficient at a frequency $\pi/2$ and selects an SARIMA model with stronger seasonality for data simulated with $d_1 >> d_2$ than for data simulated with $d_1 << d_2$. However, the main result still holds: if the true DGP is stationary FI at seasonal frequencies, then the TSW adjusted data usually is virtually not distinguishable from a process with invertible seasonal FI MA roots.

Overall, simulation results are in line with Gomez and Maravall (2001), but using a more flexible definition of invertibility: if the process contains strong nonstationary seasonality (including FI) then the SA series estimated by TSW will be in general noninvertible. However, if the original series were stationary fractionally integrated at seasonal frequencies, TSW will choose a nonstationary representation in a smaller percentage of cases and, even if a nonstationary seasonal model is chosen, the resulting SA series is likely to be not distinguishable from an invertible process. This result is important because an econometrician never knows what the DGP for the real data is, and always works with approximations which fit the data reasonably well according to the results of statistical testing. We illustrate our results with real data in the following section.
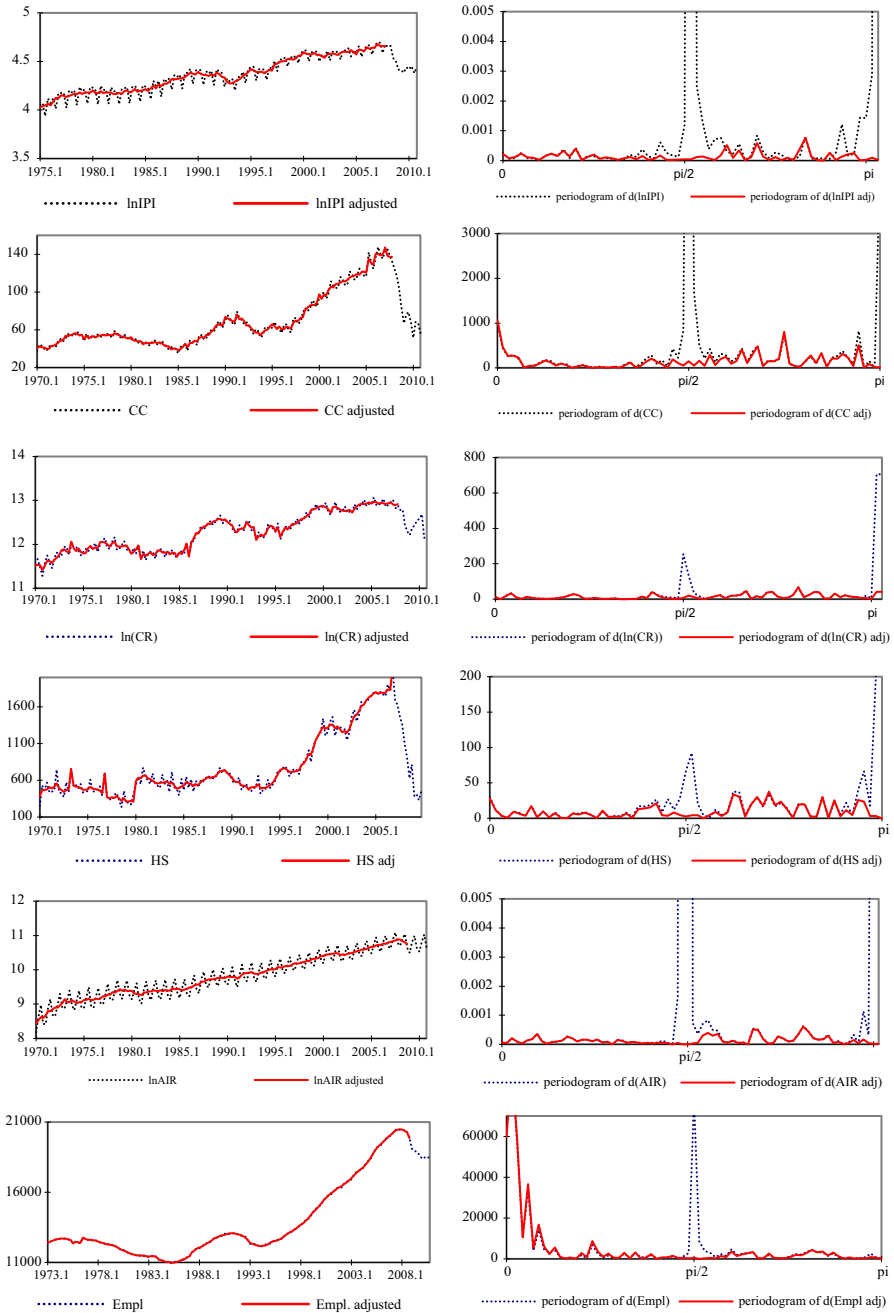
5 Empirical examples

To illustrate the simulation results, we consider several quarterly series of the Spanish economy: Industrial Production Index (IPI), airline passengers (AIR), employment (EMP) and three quarterly cyclical economic indicators, namely: cement consumption (CC), car registrations (CR) and housing starts (HS). The IPI and these three indicators are considered to be the cycle drivers for an economy in Leamer (2009) and have been recently used by Bujosa et al. (2013) to construct a composite leading indicator for the Spanish economy. All the series start in the late 60s or early 70s and they are nonstationary in the mean and strongly seasonal. Monthly data for IPI, CC, CR, HS and AIR are obtained from the Bank of Spain. To convert the IPI to quarterly, we use the simple average of the monthly observations inside each quarter. The other series are converted to quarterly by adding the observations inside the quarter. Employment has been obtained from the OECD stats database. We apply a logarithmic transformation to IPI, CR, and AIR to stabilize volatility.

Figure 2 (left panels) plots the original series and the series after adjustment by TSW. A strong seasonal pattern is observed in all the original series. We exclude the last years of data from the analysis to circumvent the strong nonlinearities consequence of the late 2000s recession.[11] This nonlinear pattern is clearly visible in Fig. 2 and may disturb the interpretation of the results.

---

[11] In particular the analysed series have the following lengths. HS: 1970:1-2006:4; CR: 1970:1-2007:4; CC: 1970:1-2007:4; ln(IPI): 1975:1-2007:4; ln(AIR): 1970:1-2008:4; EMP: 1964:1-2007:4. Note that for HS, an additional year is taken since the effects of the crisis manifest earlier. On the contrary, we have been able to use an additional year for airline passengers (AIR) and employment (EMP). Nevertheless, the estimation results including the period of the crisis are available under request.

Fig. 2 Nonadjusted and adjusted data (*left panel*) and their respective periodogram (*right panel*) from the empirical application

Table 4 Empirical example: models identified by TSW

Variable	ARIMA model chosen by TSW:
ln(IPI)	$(1 + 0.2275B)\,\nabla\nabla^4 y_t = \left(1 - 0.7146B^4\right)\varepsilon_t$, changed by seats to:
	$\nabla\nabla^4 y_t = (1 - 0.2275B)\left(1 - 0.7146B^4\right)\varepsilon_t$
CC	$\nabla\nabla^4 y_t = \left(1 - 0.7054B^4\right)\varepsilon_t$
ln(CR)	$\nabla\nabla^4 y_t = \left(1 - 0.6732B^4\right)\varepsilon_t$
HS	$\left(1 - 0.4836B^4\right)\nabla y_t = (1 - 0.5360B)\,\varepsilon_t$
ln(AIR)	$\nabla\nabla^4 y_t = (1 - 0.1686B)\left(1 - 0.6158B^4\right)\varepsilon_t$
EMP	$(1 - 0.7923B)\,\nabla\nabla^4 y_t = \left(1 - 0.6320B^4\right)e_t$

The right panel in Fig. 2 depicts the periodogram of both differenced series: original and adjusted by TSW. As expected, the periodograms of the differenced original series have strong peaks at the two seasonal frequencies, while those of the differenced adjusted series present dips at the same frequencies. The models identified by TSW for the original series are provided in Table 4.

As can be seen in the table, all the series except ln(IPI), HS and EMP follow a standard Airline model. For CC and ln(CR) the trend is very strong and Q_1 is equal to zero. The model identified for ln(IPI) does not accept the admissible decomposition and is modified by SEATS. Given that AR(1) polynomials with Φ_1 in the interval $(-0.2, -0.4)$ are practically indistinguishable from the MA(1) with $Q_1 = -\Phi_1$, SEATS replaces the original model with the Airline model (Maravall 2009). TSW chooses a stationary seasonal model for HS and SARIMA for EMP.

We estimate the coefficients of FI at seasonal frequencies before and after the adjustment by TSW. These results are presented in Table 5.

It is interesting to note that the examined series seem to follow the SARIMA model. Note that the estimated seasonal FI coefficients before the adjustment (columns four and five) are both statistically different from one in all the cases. In particular, CC, ln(CR) and EMP seem to be seasonally stationary before adjustment. Overall, in line with the simulation results, even though TSW has selected seasonally nonstationary models for the three series, the estimated coefficients after the adjustment are substantially higher than -0.5, suggesting that the adjusted series can be approximated by an invertible process. That also seems to be the case of HS, for which TSW has selected a stationary representation before adjustment. For the ln(IPI) and ln(AIR) series, the estimated coefficients of FI at frequency $\pi/2$ are larger than 0.5, albeit we cannot reject the null of $d_1 = 0.5$ at any significance level. After adjustment, the estimated coefficient of FI at the frequency π is smaller than -0.5 and the adjusted series may then be noninvertible, although we cannot reject the null of $d_2 \geq -0.5$ in favor of the alternative $d_2 < -0.5$. Overall, the empirical results are in line with the results of the simulation study.

Table 5 Empirical example: estimation results after adjustment

	T	m_1/m_2	Before adjustment		After adjustment	
			$\hat{d}_1$	$\hat{d}_2$	$\tilde{d}_1$	$\tilde{d}_2$
ln(IPI)	128	32/16	0.617	0.231	−0.153	−0.646
			(0.144)	(0.229)	(0.144)	(0.229)
CC	148	36/18	0.276	0.075	−0.127	−0.284
			(0.133)	(0.209)	(0.133)	(0.209)
CR	148	36/18	0.222	0.212	−0.317	−0.178
			(0.143)	(0.168)	(0.115)	(0.153)
HS	144	36/18	0.391	0.474	−0.266	−0.132
			(0.134)	(0.210)	(0.134)	(0.2100)
ln(AIR)	148	36/18	0.532	0.244	−0.195	−0.537
			(0.133)	(0.209)	(0.133)	(0.209)
EMP	140	36/18	0.380	0.054	−0.232	−0.226
			(0.134)	(0.210)	(0.134)	(0.210)

m_1, m_2—bandwidth parameters employed for the estimation at frequencies $\pi/2$ and π respectively selected based on an examination of the log-log plot of the tapered periodogram of the data in differences

6 Conclusions

In this paper, we have examined the invertibility property of seasonal series adjusted by TSW. According to Gomez and Maravall (2001) whenever the process chosen by TSW to fit the data contains seasonal unit roots, the adjusted series estimated by the program has MA unit roots and, as a result, it is not invertible and cannot be approximated by an AR (VAR) process as it is ordinarily done in practice.

In the simulation study carried out in this work, we found that the invertibility issue may not be in many circumstances a strong concern. In particular, we found that if the true DGP follows the Airline model, the adjusted series produced by TSW are indeed noninvertible. However, if the series is fractionally integrated at the seasonal frequencies, which is less restrictive and very plausible in some cases according to the empirical evidence, the adjusted series still can be approximated by an invertible process, depending on the stationarity of the original series. Thus, if the original series is seasonally stationary with coefficients of FI at seasonal frequencies smaller than 0.5, the SA series adjusted by TSW is likely to be statistically invertible or indistinguishable from an invertible process even if the program chose a nonstationary model for the data, therefore still admitting AR (or VAR) approximation. This approximation is more plausible the further the seasonal FI parameters of the original series are from the nonstationary region. On the contrary, if the original series is seasonally nonstationary, the resulting adjusted series are expected to be noninvertible. As shown in the empirical examples, these results are interesting since stationary FI seasonality is not a rare event in economic data.

References

Arteche J (2003) Semiparametric robust tests on seasonal or cyclical long memory time series. J Time Ser Anal 23:251–285

Arteche J, Robinson PM (2000) Semiparametric inference in seasonal and cyclical long memory processes. J Time Ser Anal 21:1–25

Arteche J, Velasco C (2005) Trimming and tapering semi-parametric estimates in asymmetric long memory time series. J Time Ser Anal 26:581–611

Berkowitz J, Diebold FX (1998) Bootstrapping multivariate spectra. Rev Econ Stat 80:664–666

Bhardwaj G, Swanson NR (2006) An empirical investigation of the usefulness of ARFIMA models for predicting macroeconomic and financial time series. J Econom 131:539–578

Box GEP, Jenkins GM (1970) Time series analysis: forecasting and control. Holden-Day, San Francisco

Bujosa M, García-Ferrer A, de Juan A (2013) Predicting recessions with factor linear dynamic harmonic regressions. J Forecast 32:481–499

Eurostat (2009) ESS guidelines on seasonal adjustment. European Communities, Luxembourg

Gil-Alana LA, Robinson PM (2001) Testing of seasonal fractional integration in the UK and Japanese consumption and income. J Appl Econom 16:95–114

Gomez V, Maravall A (2001) Seasonal adjustment and signal extractions in economic time series. In: Peña D, Tiao DG, Tsay RS (eds) A course in advanced time series analysis. Wiley, New York, pp 171–201

Grether DM, Nerlove M (1970) Some properties of optimal seasonal adjustment. Econometrica 38:682–703

Hassler U, Rodrigues P, Rubia A (2009) Testing for general fractional integration in the time domain. Economet Theor 25:1793–1828

Hurvich CM, Ray BK (1995) Estimation of the memory parameter for non-stationary or non-invertible fractionally integrated processes. J Time Ser Anal 16:17–41

Hurvich CM, Chen WW (2000) An efficient taper for potentially overdifferenced long-memory time series. J Tim Ser Anal 21:155–180

Leamer EE (2009) Macroeconomic patterns and stories. Springer, Berlin

Maravall A (2009) identification of Reg-ARIMA models and of problematic series in large scale applications program TSW. Mimeo, Bank of Spain, Statistic and Econometric Software

Nerlove M (1965) A comparison of a modified' Hannan' and the BLS seasonal adjustment filters. J Am Stat Ass 60:442–491

Ooms M, Hassler U (1997) On the effect of seasonal adjustment on the log-periodogram regression. Econ Lett 56:135–141

Porter-Hudak S (1990) An application of the seasonal fractionally differenced model to the monetary aggregates. J Am Stat Ass 85:338–344

Sidak ZK (1967) Rectangular confidence regions for the means of multivariate normal distributions. J Am Stat Ass 62:626–633

Velasco C (1999) Non-stationary log-periodogram regression. J Econom 91:325–371

Whittle P (1963) Prediction and regulation by linear least-square methods. English Universities Press, London