BANCA D'ITALIA
EUROSISTEMA

Questioni di Economia e Finanza

(Occasional Papers)

Seasonal adjustment of credit time series in the Bank of Italy

by Simone Di Paolo and Danilo Liberati

March 2024

Number 835

BANCA D'ITALIA
EUROSISTEMA

Questioni di Economia e Finanza

(Occasional Papers)

Seasonal adjustment of credit time series in the Bank of Italy

by Simone Di Paolo and Danilo Liberati

Number 835 – March 2024

The series Occasional Papers *presents studies and documents on issues pertaining to the institutional tasks of the Bank of Italy and the Eurosystem. The* Occasional Papers *appear alongside the* Working Papers *series which are specifically aimed at providing original contributions to economic research.*

The Occasional Papers *include studies conducted within the Bank of Italy, sometimes in cooperation with the Eurosystem or other institutions. The views expressed in the studies are those of the authors and do not involve the responsibility of the institutions to which they belong.*

The series is available online at www.bancaditalia.it .

ISSN 1972-6643 (online)

Designed by the Printing and Publishing Division of the Bank of Italy

SEASONAL ADJUSTMENT OF CREDIT TIME SERIES IN THE BANK OF ITALY

by Simone Di Paolo* and Danilo Liberati*

Abstract

This paper presents the seasonal adjustment process applied to the credit time series produced by the Bank of Italy. Due to the methodological and computational improvements made in recent times, making the most suitable choice for estimating the seasonal component has become more challenging. First, the paper discusses the seasonal adjustment methods most widely adopted within the community of central banks and statistical institutes. Second, it compares the output of the different approaches with reference to the credit time series currently published by the Bank of Italy and explains the main reasons underlying the adoption of X-13ARIMA-SEATS. Finally, the paper extends the new approach in order to derive new seasonally adjusted time series for banks' loans to non-financial corporations and to households classified by economic sector of activity (NACE) and by purpose of lending respectively.

JEL Classification: C22, C87.
Keywords: seasonal adjustment, X-13ARIMA-SEATS, TRAMO-SEATS, banking data.
DOI: 10.32057/0.QEF.2023.0835

---

* Bank of Italy, Economic Research and International Relations.

1 Introduction[1]

Seasonal adjustment is a statistical technique that consists in the estimation and removal of seasonal patterns from a time series. Seasonal patterns refer to regular fluctuations that systematically occur at specific times of the year. Seasonal adjustment is applied to facilitate the analysis and interpretation of economic phenomena, making it possible to compare adjacent periods and identify short-term dynamics without alterations due to the interference of such regular, exogenous factors.

Seasonality affects many economic phenomena and short-term indicators: summer holidays cause sharp decline in industrial production every August, Christmas shopping boosts retail sales every December, heating costs rise in the winter and fall in the summer, and so on. Money and credit aggregates are also affected by fluctuations of this nature. For instance, during the holiday season, people tend to withdraw more cash from their bank accounts to make purchases, which can lead to a temporary increase in currency in circulation. In addition, deposits can exhibit a seasonal pattern related to tax season or quarterly business cycles: when people receive tax refunds or companies receive payments at the end of each quarter, they may deposit the money in their bank accounts, thus determining an increase in deposits.

The removal of the seasonal component from a time series is important for a proper interpretation of the data; however, great care and caution must be paid in evaluating series net of the seasonal component. First, one should not forget that the seasonal component is unknown and therefore it must be *estimated*; second, such estimates can be affected (even to a considerable extent) by sharp and sudden shocks in economic fundamentals, as happened, for example, during the Covid-19 pandemic crisis[2].

The estimate of the seasonal component is strictly related to the adjustment method used as well as to other methodological choices, such as the pre-treatment of the raw series to remove deterministic components (outliers and calendar effects), or the time span chosen for carrying out the estimation process. Matas Mir and Rondonotti (2003) compare the performance of one of the most widely used

[1] We would like to thank Roberto Sabbatini, Laura Mellone, Emanuela Basili, Giorgio Nuzzo and Giuseppe Marinelli for their useful comments. The views expressed herein are solely those of the authors and do not necessarily reflect the views of the Bank of Italy.

[2] In this respect, Eurostat (2020) provided a guidance on time series treatment in the context of the pandemic shock in which it suggests not to model the data point hit by the shock as a seasonal outlier. Abeln and Jacobs (2022) studied the impact of Covid-19 on seasonal adjustment by focusing on different models and verifying if revisions occur once new observations become available after the shock. See also OECD (2020) for an overview of the impact of COVID-19 crisis on official statistics.

software for seasonal adjustment, the US Census Bureau's X-12-ARIMA, on a "short time series" comprising only 5 years of monthly data, against a benchmark situation (a "long time series") where 15 additional years of historical data are available. The results indicate that adjusted data based on short time series may suffer from significant distortion and great caution must be exerted when studying monthly developments from such estimates. Recently, Poměnková and Klejmova (2019) compare findings of a wavelet analysis applied to monthly data of bank loans to non-financial corporations and households in the euro area and show that using different techniques to compute seasonally adjusted time series may alter their interpretation significantly[3].

Turning to the focus of this paper, i.e. the estimate of the seasonal component of credit data, the Bank of Italy has being publishing seasonally adjusted data on growth rates of loans and deposits ever since the 1980s, and provided several methodological contributions on this topic. Cividini and Cottarelli (1989) highlight the strong seasonality in monthly series of deposits and compare the output produced by different seasonal adjustment procedures on these data. Battipaglia and Focarelli (1996) study the information content of diagnostic statistics in X-11-ARIMA by means of Monte-Carlo simulation. More recently, Silvestrini (2009) studies the results in seasonal adjustment of banks' loans and deposits obtained with TRAMO-SEATS, a model-based procedure largely used in the Bank of Italy.

Since July 1999 the European Central Bank (ECB) has being publishing data net of seasonal component and calendar effects in the Monthly Bulletin (later replaced by the Economic Bulletin) of all components and counterparts of the monetary aggregate M3[4] for the euro area as a whole. Similar data are also disseminated monthly in the statistical press release related to the monetary developments in the euro zone[5]. As for the national components of such aggregates, national central banks (NCBs) are free to resort to the seasonal adjustment method they prefer. With respect to the four largest NCBs, Banco De España (Cabrero, 2000) has a long tradition in the development and use of the TRAMO-SEATS procedure; Bank of Italy has used TRAMO-SEATS since 2009 up to now; Banque de France (Fonteny, 2006) uses a seasonal adjustment procedure which implements both TRAMO-SEATS and X-12-ARIMA for each seasonal adjusted time series[6], and the Deutsche

[3] See Grether and Nerlove (1970) for a first assessment of the effects of the application of different types of seasonal adjustment procedures to economic time series in the spectral domain.
[4] The monetary aggregate M3, sometimes also referred to as *Broad money*, includes currency, deposits with an agreed maturity of up to two years, deposits redeemable at notice of up to three months, repurchase agreements, money market fund shares/units and debt securities up to two years.
[5] The press releases can be found at the following link: https://www.ecb.europa.eu/press/pr/stats/md/html/index.en.html.
[6] In particular, TRAMO is initially employed to identify any anomalous data points in the series and to select an appropriate SARIMA model. Subsequently, both SEATS and X12 are utilized to seasonally adjust each series, using the

Bundesbank publishes the "seasonally adjusted business statistics" based on the X-13ARIMA-SEATS method[7]. Outside of the euro area, the Bank of England seasonally adjusts its data using X-13ARIMA-SEATS[8], and Danmarks Nationalbank (Pedersen and Fæste, 2006) uses the X-12-ARIMA procedure.

Within this framework, this work outlines the methodological change introduced at the beginning of 2023 by the Bank of Italy for the seasonal adjustment of monetary and credit aggregates, related to the adoption of X-13ARIMA-SEATS in place of TRAMO-SEATS. This change was made primarily for two reasons: to ensure greater alignment with the method used by the ECB and by other NCBs, and to leverage the greater versatility of X-13ARIMA-SEATS, which allows performing seasonal adjustment also by means of the algorithms that make up TRAMO-SEATS. The main goal of the paper is to highlight the differences between the two procedures, thereby explaining the rationale for making this change.

The rest of this paper is structured as follows. Section 2 describes the seasonal adjustment methods more frequently used in the community of central banks and statistical institutions. Section 3 explains in detail the new approach introduced by the Bank of Italy for estimating the time series for banks' balance sheet items net of the seasonal component, which are released in official publications, and presents new series for banks' loans to non-financial corporations and to households classified, respectively, by sector of economic activity (NACE) and by purpose of lending, that are computed with the new procedure but for the moment are used for internal purposes only. Section 4 concludes.

2 Seasonal adjustment methods

The basic idea behind seasonal adjustment methods is that every raw time series can be divided into individually unobservable components. Seasonal adjustment can thus be interpreted simply as a

same SARIMA model identified by TRAMO. The selection of the final estimates is determined based on indicators that allow identifying the series that guarantee the best combination between the quality of the data and the stability of the figures compared to the estimates made in previous months. See https://publications.banque-france.fr/sites/default/files/medias/documents/working-paper_147_2006.pdf for further details.

[7] See https://www.bundesbank.de/resource/blob/808770/9b0f24d26620cac67ed57286e6445ab9/mL/erlaeuterungen-data.pdf for further details.

[8] See https://www.bankofengland.co.uk/statistics/details/further-details-about-seasonal-adjustment-data for further details.

signal extraction procedure. Typically, the unobservable components identified in the literature are the following:

- a long-term or trend component;
- a cyclical dynamic;
- a seasonal movement;
- an irregular component, which represents the unexplained residual.

Based on the preferred decomposition scheme, these four components can then be added or multiplied to form the observed time series.

Research in seasonal modelling has a long tradition that can be traced back to the 19th century[9], and the different seasonal adjustment procedures that have been introduced over time are traditionally classified into two categories:

1. *non-parametric methods*, that use smoothing techniques based on linear filtering (especially moving averages) in extracting the different components of a time series;
2. *parametric methods*, in which the unobserved components are explicitly specified.

Recent developments have also led to mixing the features of the two approaches and to the development of the so-called *semi-parametric methods*.

Non-parametric methods were developed *empirically* and, as documented by Bell et al. (2012), their first software implementation can be traced back to Julius Shiskin, who developed the so-called *Method I* on the US Census Bureau's UNIVAC I computer in 1954. Eleven experimental versions of *Method II* (called *X-0*, *X-1*, *X-2*…) followed, until in the mid-1960s the *X-11* program was introduced. X-11 rapidly became a benchmark for official statistical offices all over the world, and it is still nowadays the most famous non-parametric seasonal adjustment approach. It also generated a great variety of alternative methods that are usually referred to as *X-11 style procedures* (e.g. STL - *Seasonal-Trend decomposition based on Loess* or SABL - *Seasonal Adjustment at Bell Laboratories*) that share its basic principles.

Shortly after, Box and Jenkins introduced a new ARIMA model-based procedure for seasonal adjustment (Box and Jenkins, 1970) and, following their work, a new variant of X-11 was developed,

[9] See Buys Ballot (1847), whose works in meteorology are often considered the first applications of seasonal adjustment procedures, or Jevons (1862).

called *X-11-ARIMA* (Dagum, 1980). The novelty of the approach proposed by Box and Jenkins laid in the use of an ARIMA model to extend the current time series by replacing the missing data both at the beginning and at the end with forecasts and backcasts. This allowed for the use of less asymmetric filters. Until then, X-11 extrapolated missing values arbitrarily, so the new approach led to much less significant revisions once the missing data became available. X-11-ARIMA represents the first major example of a semi-parametric method, employing the linear filters of X-11 in combination with the ARIMA-model filters adjusted globally to the extended series. However, in semi-parametric methods, ARIMA models cannot be applied directly to observed time series, and a preliminary analysis is necessary to address issues of stationarity and deterministic effects.

In 1998, *X-12-ARIMA*, a further evolution of X-11-ARIMA, was developed by the US Census, and it is, nowadays, one of the most popular and widely used software in seasonal adjustment alongside its latest version, called *X-13ARIMA-SEATS*, which was introduced in 2005.

However, some authors have criticized the X-11 approach based on the use of linear filters and moving averages, showing that these can introduce artificial cycles into the data, and generally deploring the empirical approach typical of non-parametric (and partially also semi-parametric) methods. For this reason, seasonal adjustment methods based on explicit models have received an increasing attention over time by the statistical institutions.

Parametric methods can be distinguished between *methods based on deterministic models* (DAINTES, BV4) and *methods based on stochastic models* (ARIMA-model based, Structural Time Series approach).

Among the ARIMA-model based methods (AMB, in short), there is the famous *TRAMO-SEATS* procedure, developed in the Bank of Spain by Víctor Gómez and Augustín Maravall in the early 1990s, which is today, along with X-13ARIMA-SEATS, the other most widely used software in the production of seasonally adjusted economic statistics. AMB methods model the observed series through a seasonal ARIMA model (SARIMA), and derive the components from the models structure by using spectral estimations. To obtain a unique decomposition from the ARIMA model fitted to the original series, AMB methods make use of what is called the *canonical decomposition*, which aims at maximizing the variance of the irregular component and minimizing that of the seasonal component. Nevertheless, just like for semi-parametric methods, it remains the fact that ARIMA models are very sensitive to outliers and fail to accurately estimate deterministic components, so that pre-treatment of the series is required.

In conclusion, TRAMO-SEATS and X-13ARIMA-SEATS are currently recognized and accepted as the best procedures for producing seasonally adjusted data in official statistics, and are also the recommended methods in the ESS guidelines on seasonal adjustment (Eurostat, 2015).

Although based on different theoretical approaches, X-13ARIMA-SEATS and TRAMO-SEATS have the common characteristic of being aimed at the semi-automatic production of estimates, with undeniable operational advantages when dealing with a large number of time series. For both procedures, the seasonal adjustment process mainly consists of two phases: the "pre-treatment" and the decomposition. In the pre-treatment, a regression model is defined for the observed series, and are estimated:

- the deterministic regression effects (outliers and calendar effects) and their corresponding coefficients;
- an ARIMA model, used to represent the so-called "linearized" series, which is the regression error.

Once the pre-treatment is completed, the linearized series is decomposed into three components: cycle-trend, seasonal, and irregular. The previously removed regression effects are then assigned to each individual components based on their nature, thus leading to the final decomposition of the observed series.

We will now proceed and analyze in greater detail these two phases, trying to highlight the specificities of TRAMO-SEATS and X-13ARIMA-SEATS.

2.1 Pre-treatment

The decomposition routines implemented in TRAMO-SEATS and X-13ARIMA-SEATS require the input series to be stochastic, i.e. clean of deterministic effects, in order to be represented by an ARIMA model. Therefore, the first step of these seasonal adjustment procedures consist in the *pre-treatment* of the observed time series aimed at adjusting it for deterministic effects such as outliers, missing observations or calendar effects, that would otherwise impede the identification of an optimal ARIMA model. This process is also known as *linearization*, since it results in a series that, by assumption, can be represented using an ARIMA linear model.

The pre-treatment in TRAMO-SEATS is performed by *TRAMO* (Time Series Regression with ARIMA Noise, Missing Observations, and Outliers), while in X-13ARIMA-SEATS, just like in its predecessor X-12-ARIMA, the *RegARIMA* algorithm is used. The two routines have slight differences, but they both make use of *regARIMA models* (i.e. regression models with ARIMA errors).

A regARIMA model for a time series Z_t can be written in the following form:

$$Z_t = \sum_i \beta_i X_{i,t} + Y_t. \tag{1}$$

Here, the $X_{i,t}$ are the deterministic regression variables observed concurrently with Z_t, and Y_t is the time series of regression errors (the linearized series mentioned above). Y_t is assumed to follow a multiplicative seasonal ARIMA model:

$$\varphi(L)\,\Phi(L^s)(1-L)^d(1-L^s)^D\,Y_t = \theta(L)\Theta(L^s)\varepsilon_t \tag{2}$$

where:

- L is the lag operator, such that $L\,Y_t = Y_{t-1}$;
- s is the seasonal period ($s = 12$ for monthly data, and $s = 4$ for quarterly data);
- $\varphi(L) = (1 - \varphi_1 L - \varphi_2 L^2 - \cdots - \varphi_p L^p)$ and $\Phi(L^s) = (1 - \Phi_1 L^s - \Phi_2 L^{2s} - \cdots - \Phi_P L^{Ps})$ are the nonseasonal and seasonal autoregressive polynomials of order p and P respectively;
- $(1-L)^d$ and $(1-L^s)^D$ are the non-seasonal and seasonal differencings of order d and D respectively. If $d = D = 0$ (i.e. no differencing), it is common to replace Y_t in (2) with the deviations from its mean, that is, with $Y_t - \mu$ where $\mu = \mathbb{E}(Y_t)$;
- $\theta(L) = (1 - \theta_1 L - \theta_2 L^2 - \cdots - \theta_q L^q)$ and $\Theta(L^s) = (1 - \Theta_1 L^s - \Theta_2 L^{2s} - \cdots - \Theta_Q L^{Qs})$ are the nonseasonal and seasonal moving average polynomials of order q and Q respectively;
- ε_t is a white noise process with 0 mean and variance equal to σ^2.

The general regARIMA is obtained by combining the equations (1) and (2):

$$\varphi(L)\,\Phi(L^s)(1-L)^d(1-L^s)^D\left(Z_t - \sum_i \beta_i X_{i,t}\right) = \theta(L)\Theta(L^s)\varepsilon_t. \tag{3}$$

From (3) we see that in a regARIMA model the regression variables are first subtracted from the observed series Z_t, to get the zero mean series Y_t, which is then differenced to get a series, say v_t that follows the stationary ARIMA model $\varphi(L)\ \Phi(L^s)\ v_t = \theta(L)\Theta(L^s)\varepsilon_t$.

Therefore, in the pre-treatment phase, the first objective is to identify the deterministic regression variables and then proceed with estimating the ARIMA model for the linearized series. There are two main types of regression variables:

- calendar effects, such as the *day-of-week* or the *moving holidays effects*;
- outliers.

Day-of-week effects are related to the fact that monthly economic time series (such as monthly sales or inventories at month's end) are usually compiled by aggregating daily values. In these cases, there can be a pattern of variations within the week that causes the monthly series to change with the day-of-week composition. For example, May 2023 began on a Monday, so in this month there have been five Mondays, Tuesdays and Wednesdays, and four of each of the other days. On the other hand, in May 2020 there have been five Fridays, Saturdays and Sundays, and four of each of the other days. The different numbers of each weekday may affect monthly data on economic phenomena such as production, sales and turnover. This kind of situation can be detected by observing a particularly distorted sample autocorrelation function.

To deal with these effects, it is possible to use as regression variables the counting of the number of each weekday i within a month t. The total *trading day effect* can then be written as:

$$td(\beta_1, \dots, \beta_7, W_{1,t}, \dots, W_{7,t}) = \sum_{i=1,7} \beta_i W_{i,t}. \tag{4}$$

Here, $W_{1,t}$ is the number of Mondays in month t, $W_{2,t}$ is the number of Tuesdays, ..., and $W_{7,t}$ is the number of Sundays, while β_i, $i = 1, \dots, 7$ represents the effect of each weekday. To avoid multicollinearity and to consider only the non-seasonal component of the trading day effects, the average of $W_{\mathrm{i},t}$is set to be zero:

$$\frac{1}{336} \sum_{t=1,336} \sum_{i=1,7} \beta_i W_{i,t} = \frac{1}{336} \sum_{i=1,7} \beta_i \sum_{t=1,336} W_{i,t} = 0 \tag{5}$$

where $336 = 12 \times 28$ is used at the denominator because the calendar is periodic of 28 years. Since $W_{i,t} > 0$, equation (5) implies that $\sum_{i=1,7} \beta_i = 0$, and thus $\beta_7 = -\sum_{i=1,6} \beta_i$. Therefore, equation (4) can be written as:

$$td(\beta_1, \dots, \beta_7, W_{1,t}, \dots, W_{7,t}) = \sum_{i=1,6} \beta_i W_{i,t} - \sum_{i=1,6} \beta_i W_{7,t} = \sum_{i=1,6} \beta_i \left(W_{i,t} - W_{7,t}\right)$$

If we define $D_{i,t} = W_{i,t} - W_{7,t}$[10], $i = 1, \dots, 6$, the trading day effect becomes:

$$TD(\beta_1, \dots, \beta_6, D_{1,t}, \dots, D_{6,t}) = \sum_{i=1,6} \beta_i D_{i,t}. \tag{6}$$

An alternative, and more parsimonious, representation of the effects due to the composition of the calendar is provided by the so-called *working day effect*:

$$WD(\beta, D_t) = \beta \cdot D_t = \beta \left(\sum_{i=1,5} W_{i,t} - \frac{5}{2} \cdot \sum_{i=6,7} W_{i,t} \right). \tag{7}$$

In (7) there is only one regression variable D_t, with the key assumption that days from Monday to Friday have similar effects, while Saturday and Sunday are treated as contrast variables.

Besides the day-of-week effects, there are also holidays that need to be managed in a specific way as they are based on the lunar calendar and fall on different days (or even months) each year. The most important example of such a holiday is Easter, which can fall between March 22nd and April 25th, and its effect is felt in the days leading up to it (for example, sales turnover usually increases before Easter). The variable used to model the Easter effect is the following:

$$E_{\gamma,t} = \frac{1}{\gamma} \cdot n_t,$$

where γ represents the number of days of the Easter effect before Easter Sunday, and n_t is the number of the γ days before Easter that fall in month t. For example, if Easter falls on April 3 and

[10] The use of Sunday to define the contrast variables $D_{i,t}$ is usual in the literature, but clearly, each day of the week can be used.

$\gamma = 8$ (that is, we are assuming Easter affects the 8 days preceding the Easter Sunday), then $E_{\gamma,t} = 5/8$ for $t = 3$ (March), $E_{\gamma,t} = 3/8$ in April, and $E_{\gamma,t} = 0$ for all other months.

In modeling calendar effects, many other regressors could be taken in consideration, for example length-of-month and leap-year variables. For more details and references see e.g. (Eurostat, 2017, 2018; Findley and Monsell, 2009).

Outliers are instead unusual observations that have a substantial impact on the time series and are related to external events and unexpected changes, such as new regulations or recording errors. Following the work by Fox (1972), outliers have been traditionally classified in three different types:

1. *additive outliers* (AO), that detect isolated spikes in a time series, following events that have an influence for only one period, $t = t_0$. They can be represented as:

$$P_{t_0}(t) = \begin{cases} 1, & t = t_0, \\ 0, & t \neq t_0; \end{cases}$$

2. *level shifts* (LS), that constitute actual "steps" that permanently change the level of a series from time $t = t_0$ onward. They can be represented by a step function:

$$S_{t_0}(t) = \begin{cases} -1, & t < t_0, \\ 0, & t \geq t_0; \end{cases}$$

3. *transitory changes* (TC), representing an isolated jump at time $t = t_0$, followed by an exponential decay back to the starting level. Their representation is as follows:

$$T_{t_0}(t) = \begin{cases} 0, & t < t_0, \\ \delta^{t-t_0}, & t \geq t_0 \end{cases}$$

where $\delta \in (0, 1)$ is called the dampening factor.

Other types of outliers (such as Seasonal or Innovative Outliers) have been proposed over time and are included both in TRAMO-SEATS and X13-ARIMA-SEATS. For further details and references, see (Eurostat, 2017, 2018; Fox, 1972; Chen and Liu, 1993; Kaiser and Maravall, 1992).

TRAMO-SEATS and X13-ARIMA-SEATS feature an automatic method for outlier detection based on iterative procedures first proposed by Chang et al. (1988) and Tsay (1988). The goal is to identify the timing of the anomalous observation, determine its most likely type, and adjust the series for its effects. According to these procedures, outlier detection occurs by comparing the absolute values of t-statistics with a critical value C[11], which is set as an input parameter for the programs. The default critical value is determined by the number of observations in the interval where outliers are searched. Indicatively, based on Monte Carlo simulations, Chang et al. (1988) recommended three possible values for this parameter: $C = 3$ for high sensitivity to outliers, $C = 3.5$ for medium sensitivity, and $C = 4$ for low sensitivity. The issue of whether to model or not a potential outlier can be interpreted in terms of errors associated with statistical hypothesis testing. A critical value too low can lead to an increase in the rate of false positives, i.e. an observation is modelled as anomalous when it is not, resulting in an increased probability of type I error. Conversely, low sensitivity increases the rate of false negatives and thus the probability of type II errors, i.e. no outlier is modelled when the observation is actually anomalous. From a purely statistical perspective, false negatives represent more severe errors because in the case of false positives the estimation of the parameter associated with the outlier should be statistically non-significant. Moreover, false negatives end up in the error term, which may no longer satisfy the conditions for unbiased parameter estimation, leading to biased forecasts and invalid calendar adjustments. However, if too many outliers are identified when they are actually valid observations, degrees of freedom are lost in the modelling.

In general, modelling outliers relies on the co-presence of statistical significance of an identified outlier and additional information that can support such an interpretation of the data. Indeed, a change in the level of a time series never occurs out of nowhere, and very often, the analyst has access to additional information that provides details about what happened. Based on these elements, it is possible to assess both the most appropriate type of outlier and its expected effects. The direct intervention of the analyst is also particularly important in managing outliers at the current end of the time series. In these cases, it is not possible to understand from data only whether the effect of the anomalous value is permanent or temporary, and it is not possible to tell the difference between an additive outlier, a level shift, and a temporary change. The choice must therefore be guided by considerations based on the knowledge of the treated phenomenon and must be validated upon the arrival of new observations.

Once the deterministic effects related to the calendar and anomalous values have been removed from the raw time series, the second step of the pre-treatment consists, as mentioned earlier, in the identification of an ARIMA model for the regression residuals. The approach used for this purpose,

[11] Outlier identification depends on the model and its parameters. In fact, an observation is marked as outlier if it is further away from the expectation than the critical value C times the standard deviation of the residuals. If the model changes with the course of time, expectation and standard deviation vary as well, leading to a different outlier detection and revisions of seasonally adjusted data.

both in TRAMO and in RegARIMA, is based on the one proposed by Box and Jenkins (1970), which consists of an iterative scheme composed of three steps:

1. model identification: a tentative model is selected, fixing the order of the ARIMA model $(p, d, q)(P, D, Q)$. In this phase, it is crucial to consider models as parsimonious as possible (i.e., with a small number of parameters) to minimize autocorrelations and cross-correlations among variables that could lead to spurious relationships in the model;
2. estimation of the parameters of the selected model and of the variance of the white noise process ε_t in (2);
3. diagnostic checking: the goodness of the model is evaluated, considering, for example, the residuals, which should, by assumption, be normally, independently, and identically distributed, and the statistical significance of the estimated parameters.

In general, this model construction procedure is complex and time-consuming, and the choice of the optimal ARIMA model is made based on information criteria such as the Akaike Information Criterion (AIC, also in its corrected form AICC) or the Bayesian Information Criterion (BIC), which evaluate the trade-off between the models log -likelihood and number of parameters.

From an implementation standpoint, the differences between TRAMO and RegARIMA are mainly related to some technical aspect[12], and the two routines carry out the pre-treatment following the basic steps outlined above. Both procedures begin with a test for log-level specification, in which they evaluate whether to apply or not a logarithmic transformation of the data before proceeding with estimation. TRAMO and RegARIMA are capable of identifying *automatically* the systematic components related to the calendar and any anomalous values. They perform automatic selection of the ARIMA model and estimate its parameters, generating forecasts and backcasts used in the subsequent decomposition phase.

However, even though the capability of both software to proceed in the pre-treatment automatically is their major strength (enabling the massive treatment and decomposition of many

[12]For example, in choosing the parameters for the ARIMA model, RegARIMA uses the AICC while TRAMO employs the BIC, with the addition of some further constraints aimed at favoring balanced models. Another difference concerns the handling of fixed national holidays (those that fall on a specific date or on a particular working day of a given month). In RegARIMA, these holidays are not associated with specific regressors in the regARIMA model, as they are expected to have a fixed effect (not affecting other months) and thus should be incorporated into the seasonal component of the series. In contrast, in TRAMO, they are included in the regARIMA model and treated as Sundays. Differences of this kind, while not particularly significant from a theoretical and methodological perspective, can lead to important discrepancies in the final estimates, as they can lead to the selection of different ARIMA models, affecting parameter estimation, forecasts, and the subsequent decomposition.

series with clear benefits for the overall quality of data), it is important that the analyst maintains a certain level of control over the parameters used by these routines. For instance, when estimating calendar effects, the optimal choice of regressors should be made manually based on the users knowledge of the series being modeled, and, as for the ARIMA model for the linearized series, in general it is not ideal to let TRAMO or RegARIMA choose it automatically. Indeed, with the arrival of new data, the automatic model selection could change from month to month, resulting over time in highly variable and unstable final estimates. As we will see in Section 2.3, a proper balance between data quality and estimation stability should always be sought.

For an in-depth description of TRAMO and RegARIMA algorithms see (Gómez and Maravall, 1992; Maravall and Caporello, 2004; Findley, et al., 1988, 1998; US Bureau of the Census, 2000, 2017).

2.2 The decomposition

The second step of the seasonal adjustment process consists, as mentioned above, in the decomposition of the linearized series in its unobserved stochastic components. This is the phase where the most significant differences emerge between the parametric AMB approach of TRAMO-SEATS and the non-parametric approach of the X-11 algorithm, which has also been adopted (with some improvements over time) by the more recent procedures X-12-ARIMA and X13-ARIMA-SEATS.

In TRAMO-SEATS, the decomposition is carried out by SEATS (*Signal Extraction in ARIMA Time Series*) through an *AMB* procedure based on the MSX (*Minimum Seasonal eXtraction*) method developed by Baruman (1980) at the Bank of England. The basic assumption, as stated above, is that the linearized series Y_t produced by TRAMO follows the ARIMA model (2), that we can also write as:

$$\varphi(L)\, Y_t = \theta(L)\varepsilon_t, \tag{8}$$

where $\varphi(L)$ contains both stationary and nonstationary roots[13]. Let us consider the case of an additive decomposition of the linearized series into a trend T_t, a seasonal component S_t and an irregular component I_t

$$Y_t = T_t + S_t + I_t. \tag{9}$$

Apart from the irregular component, which is assumed to be white noise, T_t and S_t follow, by assumption, an ARIMA model of the same type as the one in Equation (8), with their own AR and MA polynomials:

$$\begin{aligned} \varphi_T(L)\, T_t &= \theta_T(L)\varepsilon_{Tt}, \\ \varphi_S(L)\, S_t &= \theta_S(L)\varepsilon_{St}. \end{aligned} \tag{10}$$

Since the combination of ARIMA models produces ARIMA models, from Equations (8), (9), and (10) we can derive the following decomposition for Y_t:

$$Y_t = \frac{\theta(L)}{\varphi(L)}\,\varepsilon_t = \frac{\theta_T(L)}{\varphi_T(L)}\,\varepsilon_{Tt} + \frac{\theta_S(L)}{\varphi_S(L)}\,\varepsilon_{St} + I_t.$$

SEATS starts from the ARIMA model identified for the series Y_t and decomposes it into the models for each component of the series. In general, there is no unique decomposition, and various assumptions (like the aforementioned *canonical decomposition*) are to be made to handle the problem of under-identification. Finally, for estimating the parameters of the trend-cycle and seasonal components, *Wiener-Kolmogorov* filters are employed[14].

In X-11, the decomposition is instead carried out through non-parametric smoothing procedures based on the use of moving averages. In general, given a time series w_t, we can define its moving average of *order* $p + f + 1$ with coefficients θ_i as follows:

$$MA(w_t) = \sum_{i=-p}^{f} \theta_i w_{t+i}.$$

[13]In (8), both φ and θ are express only as functions of L in order to make the notation more concise, but the generalized autoregressive and moving average polynomials are to be thought also as functions of L^s.

[14] For a complete account of the more technical aspects of SEATS, see (Maravall and Gómez, 1992; Gómez and Maravall, 1992, 1997; Maravall and Caporello, 2004; Silvestrini, 2009).

The value of w_t is replaced by a weighted average of p past values, f future values, and the current value at time t. If $p = f$, the moving average is said to be *centered*, and, if for every index i, $\theta_{-i} = \theta_i$, the moving average is *symmetric*.

In estimating the seasonal component, X-11 resorts to the iterative use of moving averages of order $m \times n$, obtained by composing a moving average of order m with coefficients all equal to $1/m$, and a moving average of order n, with all coefficients equal to $1/n$. The procedure followed by the X-11 algorithm can be divided in three steps[15]:

1. *the initial decomposition*. Starting from the linearized series Y_t, a first estimate for the cycle-trend component $T_t^{(1)}$ is obtained by using a 2×12 moving average[16]:
$$T_t^{(1)} = \frac{(Y_{t-6} + 2Y_{t-5} + 2Y_{t-4} + \cdots + 2Y_{t+4} + 2Y_{t+5} + Y_{t+6})}{24}.$$
Then, the *Seasonal-Irregular component* (also referred to as the *SI-ratio*) $SI_t^{(1)}$ is computed as:
$$SI_t^{(1)} = Y_t - T_t^{(1)},$$
in the case of an additive decomposition, while, for a multiplicative model, its formula takes the following form:
$$SI_t^{(1)} = \frac{Y_t}{T_t^{(1)}}.$$
Seasonal factors are then calculated through a 3×3 moving average:
$$\tilde{S}_t^{(1)} = \frac{\left(SI_{t-24}^{(1)} + 2SI_{t-12}^{(1)} + 3SI_t^{(1)} + 2SI_{t+12}^{(1)} + SI_{t+24}^{(1)}\right)}{9}$$
and are subsequently normalized by means of a 2×12 filter
$$\bar{S}_t^{(1)} = \frac{\left(\tilde{S}_{t-6}^{(1)} + 2\tilde{S}_{t-5}^{(1)} + 2\tilde{S}_{t-4}^{(1)} + \cdots + 2\tilde{S}_{t+4}^{(1)} + 2\tilde{S}_{t+5}^{(1)} + \tilde{S}_{t+6}^{(1)}\right)}{24}.$$
The seasonal component $S_t^{(1)}$ is obtained as
$$S_t^{(1)} = \tilde{S}_t^{(1)} - \bar{S}_t^{(1)},$$
for additive models, and as
$$S_t^{(1)} = \frac{\tilde{S}_t^{(1)}}{\bar{S}_t^{(1)}},$$

[15] See also https://sylwiagrudkowska.github.io/JDemetra-documentation/pages/theory/SA_X11.html.
[16] This holds true for monthly time series. In the case of quarterly data a 2× 4 average is used.

in the case of multiplicative decomposition. The first estimate of the seasonally adjusted series is therefore given by:

$$SA_t^{(1)} = Y_t - S_t^{(1)},$$

or

$$SA_t^{(1)} = \frac{Y_t}{S_t^{(1)}}.$$

The final estimate is obtained as an improvement of this initial decomposition;

2. *trend estimation in the absence of seasonality*. The *Henderson filters* are applied to the first estimate of the seasonally adjusted series $SA_t^{(1)}$ to obtain a second estimate of the trend component $T_t^{(2)}$. As mentioned earlier, in the initial trend estimation, equal weights are used for most months, but, when fitting nonlinear trends, different weights (and in particular filters with negative coefficients both at the beginning and the end) could be used. This is precisely the case with the Henderson filters, which have excellent smoothing properties. The second estimate of the trend component $T_t^{(2)}$ is then used to compute once again the SI-ratios $SI_t^{(2)}$;
3. *seasonal component estimation in the absence of trend*. In the third, and final step of the decomposition, the seasonal component $\tilde{S}_t^{(2)}$ is estimated by applying a moving average[17] to the SI-ratios $SI_t^{(2)}$, and it is normalized by means of a 2×12 filter. The final estimate of the seasonally adjusted series is obtained by following the same steps as above.

For a complete description of X-11, see (Ladiray and Quenneville, 2001).

Lastly, once the linearized series has been decomposed, in order to obtain the final components for the observed series, the deterministic effects (estimated in the pre-treatment phase) are to be added to the stochastic components based on their nature. In particular:

- calendar effects are assigned to the seasonal component (therefore they are not part of the seasonally adjusted series);
- level shifts, representing systematic changes in data, are assigned to the trend-cycle component;
- transitory changes and additive outliers are attributed to the irregular component.

[17]The filter's order is automatically chosen among 3×1, 3×3, 3×5, 3×9, and 3×15.

Figure 1. RegARIMA models

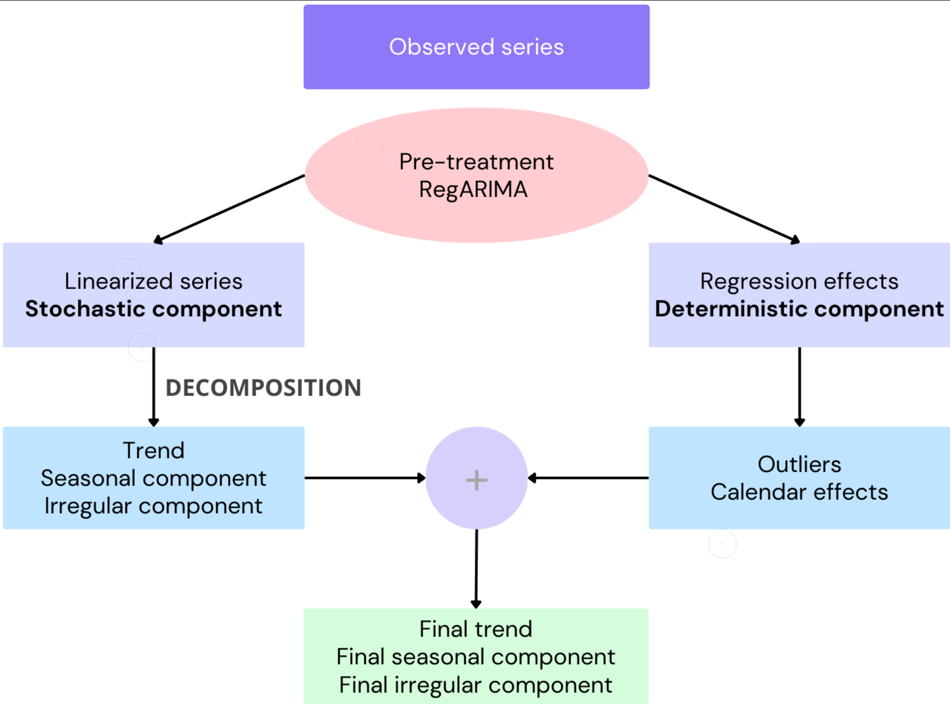


Figure 1 summarizes the main steps followed by the two procedures, X-13ARIMA-SEATS and TRAMO-SEATS.

2.3 Revision policies for seasonal adjusted data

A relevant aspect in the production of seasonal adjusted data concerns the revision policies being adopted. In general, to ensure the quality of the final series, it is always advisable to update the estimates whenever new data becomes available (whether it is new observations or revisions of past data). However, this can lead to issues related to the instability of the estimates. In fact, even a single new observation can result in changes to the estimated model, causing significant variations in the seasonal adjusted series for numerous years. These variations can then propagate to the rates of change commonly used in business cycle analysis, potentially confusing users.

Therefore, a trade-off must be found between ensuring the highest quality of estimates on the one hand, and the stability of data over time on the other, thus avoiding non-significant revisions that could later be reversed (especially in the case of new observations, which are often provisional or partially estimated, and could be revised before becoming final).

The European Statistical System identifies four main possible approaches in revising seasonal adjusted data:

1. *concurrent adjustment*: whenever new data becomes available, the seasonal adjustment procedure is carried out in its entirety, leading to new estimates for the ARIMA model, filters, deterministic regressors, and their respective parameters. This approach generates the most accurate estimates at any given time, but sacrifices the stability of the final series, as it produces continuous revisions, often of small magnitude, and sometimes even of opposite directions;
2. *partial concurrent adjustment*: the ARIMA model, filters, and deterministic regressors are estimated once a year and maintained for the estimates of the following twelve months. When new data becomes available, only the model coefficients and seasonal factors are re-estimated;
3. *current adjustment*: in this approach the identification of the ARIMA model and the estimate of its coefficients, forecasts, deterministic regression effects, and seasonal factors are carried out annually. For the next twelve months, they are all kept unchanged and used in the production of seasonal adjusted series. This revision policy ensures maximum stability, but may lead to poor accuracy of the final estimates;
4. *controlled current adjustment*: the seasonal factors estimated annually by the *current adjustment* are forecasted and used for seasonal adjustment in the following months. However, to ensure higher quality, when new data arrives, the factors are re-estimated, and seasonal adjusted series are calculated using the *partial concurrent adjustment* approach. If significant differences exist, the latter estimate is preferred. In practice, controlled current adjustment means that each series undergoes seasonal adjustment twice. For this reason, such approach is feasible only when dealing with a few and important time series.

Rather than the two extremes represented by current and concurrent adjustments, the ESS suggests using one of the other two revision policies, which can adapt to a broader class of time series and provide a reasonable balance between the number of revisions and the quality of the final series. Nevertheless, the optimal revision policy varies depending on the time series under consideration. For example, in dealing with a series characterized by deterministic or highly stable seasonal patterns, current adjustment may be the best choice, whereas, in the presence of moving seasonality, concurrent adjustment is to be preferred.

Silvestrini (2009) conducts an interesting experiment comparing the different revision policies listed above, by measuring the speed of convergence of the estimates obtained with each type of revision method towards a stable seasonal adjusted series that serves as a benchmark. The result of the experiment suggests following the method of partial concurrent adjustment.

Another important methodological aspect, partly related to revision policies, concerns the choice of the time span of the series to be seasonally adjusted. Seasonal adjustment procedures are generally designed to be applied to series that are sufficiently long. This is related both to technical factors, such as the properties of the symmetric filters used in the process, and non-technical aspects, such as the fact that unobserved components can be better identified over sufficiently long periods. Additionally, long series are also required to correctly interpret the seasonality test statistics, which might be less robust when conducted on short time series. Finally, when series are short, they tend to be more volatile, so the irregular component dominates the signals and complicates the correct identification and estimation of components.

On the other hand, performing seasonal adjustment on series that are too long can also be challenging. Over very long periods, series may exhibit a significant heterogeneity related to potential changes in economic fundamentals, data collection methods, social habits, and so on; which would lead to changes in the unobserved components and their structures. Seasonal adjustment over the entire series in these cases may produce suboptimal results, especially in the more recent and initial parts of the series.

Based on the ESS Guidelines on Seasonal Adjustment (2015), monthly series of 3 years or less are to be considered too short to be seasonally adjusted, and a good level of reliability in the estimates is provided only by series of at least 8 years. When a series is longer than 21 years, it is considered as too long, and it would be preferable to split it in order to obtain higher quality estimates.

3 The seasonal adjustment procedures adopted in the Bank of Italy

In the statistical and econometric literature, there is no unanimous consensus on the procedure that produces the “best” seasonally adjusted estimates. The Italian National Institute of Statistics (ISTAT) has adopted TRAMO-SEATS as its official method of seasonal adjustment, and the Bank of Italy too moved to this software in 2009 for producing its seasonally adjusted series in monetary and credit statistics.

Seasonal adjustment over the years has been carried out through a *partially concurrent adjustment*. The order of the ARIMA model, outliers, and deterministic regressors are identified once a year, while their parameters and the seasonal adjustment factors are estimated whenever new or revised data become available. Therefore, every month, the seasonally adjusted series are computed in their entirety (starting from January 2003) based on the new estimated factors. This approach, as mentioned earlier, was chosen because it can provide an adequate trade-off between the quality of estimates and their stability over time.

On the other hand, the European Central Bank uses the X-12-ARIMA software for producing its seasonally adjusted series. The revision policy adopted is the *controlled current adjustment*, aimed at ensuring the highest stability in the estimates. The RegARIMA modeling and estimation of seasonal factors are carried out annually using the information available when the models are executed, and the time series produced include forecasts of seasonal factors for 12 months, which are used throughout the following year. Every month, the ECB checks whether new available information implies significant revisions of seasonal patterns. If relevant differences emerge, seasonal factors are updated outside of the annually scheduled review.

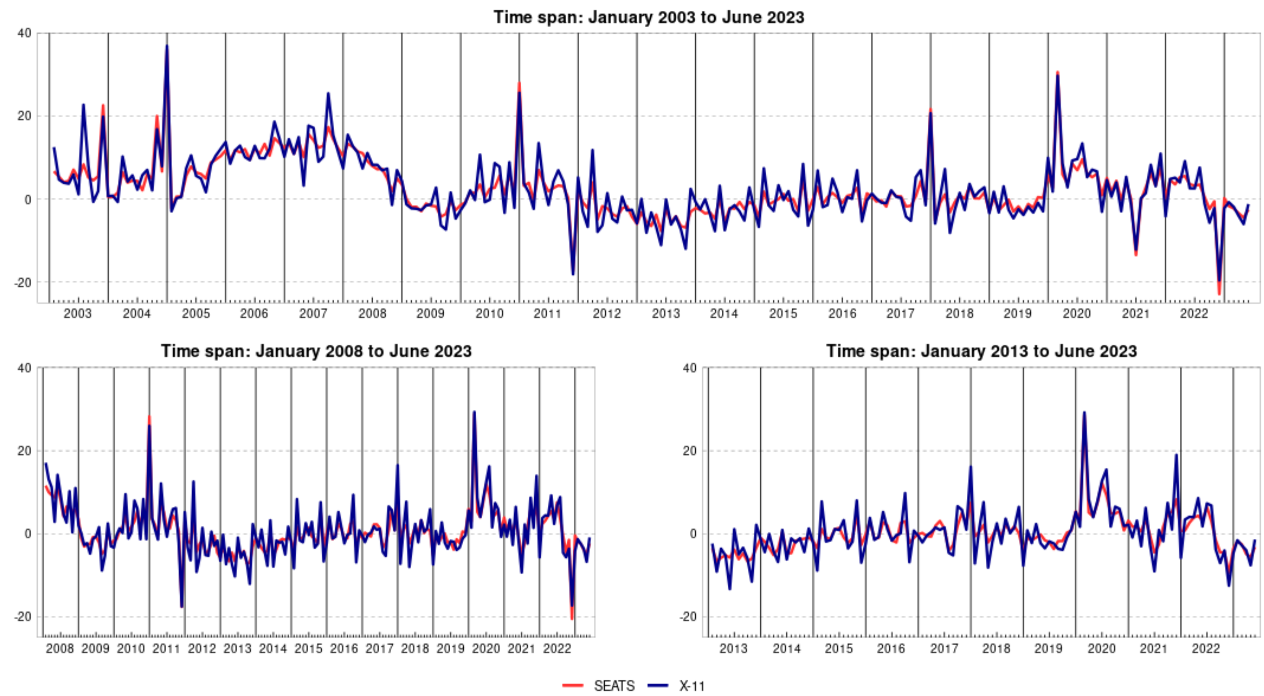
Recently, a change to the internal procedures of the Bank of Italy for producing seasonally adjusted series has been introduced and X-13ARIMA-SEATS is now used. This software has been chosen for its versatility, since it can perform seasonal adjustment both through the ARIMA model-based procedure SEATS and via the X-11 algorithm (in its semi-parametric version). X-13ARIMA-SEATS also has the advantage of providing a wide range of statistics and diagnostic tools useful for assessing the quality of estimates obtained.

As for the revision policy, it was decided to maintain a partially concurrent adjustment. In addition to ensuring greater continuity with the procedures previously in use, this approach brings some relevant operational advantages. As mentioned earlier, controlled current adjustment requires performing seasonal adjustment twice for each time series, which is not in line with internal operational needs, that require handling numerous time series within a short timeframe.

As for the decomposition phase, X-11 was preferred to SEATS in order to align the Institute with the methodology used by the ECB. Moreover, X-11 ensures high data quality in the case of particularly long series. Currently, the series produced by the Bank of Italy have a temporal depth of 20 years (the first observations are dated back to January 2003), and seasonal adjustment is performed on the whole series. Identifying a satisfactory ARIMA model for such long series may be challenging, and this can have a significant impact on the final estimates, especially when decomposition is performed based on a parametric procedure like SEATS. In the case of X-11, however, since decomposition is based on predetermined moving average filters that only use data from the years closest to the reference date, values from the past of the series provide a minimal contribution to current estimates.

The stability of X-11 can be observed comparing the seasonally adjusted one-month growth rate series of loans to non-financial corporations obtained using SEATS and X-11 in the decomposition phase. Figure 2 shows the comparison of the series over three different time spans: January 2003 – June 2023 (the entire time interval), January 2008 – June 2023 (the last 15 years), and January 2013 – June 2023 (the last 10 years). In addition to the graphical inspection, also the correlation between seasonally adjusted estimates obtained by applying X-11 on the entire series and on the last 10 years is greater than that obtained by applying SEATS.

Figure 2. Loans to non-financial corporations – one-month percentage changes on an annual base



These considerations have led us to choosing a partially concurrent adjustment over the entire time horizon of the time series, using RegARIMA for the pre-treatment and X-11 for the decomposition. The quality of seasonally adjusted data is monitored monthly and every year is evaluated the opportunity to revise the model.

3.1 The seasonally adjusted series produced by the Bank of Italy

In the monthly publication "*Banks and Money: National Data*", the Bank of Italy publishes seasonally adjusted percentage changes for the main balance sheet items of Italian banks. Specifically, the series pertain to one-month percentage changes on an annual basis of loans and deposits split by the different domestic counterparties (households, non-financial corporations, other financial intermediaries, and private sector), as well as the percentage changes of debt securities issued and held in banks portfolios [18].

The computation of the growth rates is performed through an *indirect method* for the series related to the private sector, and through a *direct method* for all others. The direct method is obtained by directly applying the seasonal adjustment algorithm to the raw time series, while the indirect method is derived from a linear combination of other directly adjusted series. The calculation of directly adjusted series involves the following steps:

1. *computation of the index of notional stocks*: the outstanding amount at the end of the month reported by the banks in their balance sheet does not reflect only the cumulative effect of transactions, but it is also influenced by reclassifications and other factors, such as variations in exchange rates and loan write-offs. In the analysis of credit trends, and particularly in the calculation of growth rates, it is convenient to remove these non-transaction-related variations from the stocks. The index of notional stocks is introduced in order to isolate the purely transactional part of the month-to-month variation in outstanding amounts and it is given by the chain formula:

$$\mathrm{Z}_t = \mathrm{Z}_{t-1} \cdot \left(1 + \frac{F_t}{L_{t-1}}\right), \tag{11}$$

where F_t and L_t represent the flow and the outstanding amount at time t respectively. By iterating (11), Z_t can also be expressed as:

$$\mathrm{Z}_t = \mathrm{Z}_0 \prod_{j=0}^{t-1} \left(1 + \frac{F_{t-j}}{L_{t-j-1}}\right). \tag{12}$$

Here, Z_0 is the index of notional stocks at the base period $t = 0$, and it is conventionally set to either 1 or 100.

[18] For the series of debt securities held in portfolios issued by all sectors excluding the banking sector, are also published the 12-month percentage changes.

The index of notional stocks is dimensionless, and its value depends exclusively on the period chosen as the basis for the chain calculation, the corresponding level, and the month-to-month growth rate;

2. *seasonal adjustment*: the X13-ARIMA-SEATS procedure takes in input the index of notional stocks Z_t and carries out the seasonal adjustment procedure. It is assumed a multiplicative decomposition scheme so that Z_t can be written as:

$$\mathrm{Z}_t = \mathrm{T}_t \cdot \mathrm{S}_t \cdot I_t. \tag{13}$$

In (13), T_t, S_t, and I_t represent the final trend, seasonal and irregular components, and we are assuming that they also include the deterministic regressors. S_t is also the *seasonal adjustment factor* and it is used in computing the *seasonally adjusted index of notional stocks*:

$$\mathrm{Z}_t^{SA} = \frac{\mathrm{Z}_t}{\mathrm{S}_t} = \mathrm{T}_t \cdot I_t. \tag{14}$$

3. *computation of seasonally adjusted stocks and flows*: in the case of a multiplicative decomposition, the seasonally adjusted series of outstanding amounts is obtained by dividing the raw series by the seasonal adjustment factor computed in the previous step:

$$L_t^{SA} = \frac{\mathrm{L}_t}{\mathrm{S}_t}. \tag{15}$$

The adjusted flow, on the other hand, is obtained as:

$$F_t^{SA} = (L_t^{SA} - L_{t-1}^{SA}) - \frac{[(\mathrm{L}_t - \mathrm{L}_{t-1}) - \mathrm{F}_t]}{\mathrm{S}_t}. \tag{16}$$

4. *computation of seasonally adjusted growth rates:* finally, the seasonally adjusted one-month percentage change on an annual basis are computed with the following formula:

$$R_1^{SA} = 100\left(\left(1 + \frac{\mathrm{Z}_t^{SA} - \mathrm{Z}_{t-1}^{SA}}{\mathrm{Z}_{t-1}^{SA}}\right)^{12} - 1\right). \tag{17}$$

The indirect method, on the other hand, is used for calculating the two series of one-month percentage changes on an annual basis of total deposits and total loans with the private sector. This approach is chosen primarily with the aim of preserving additivity between the seasonally adjusted aggregates and their respective components. The private sector is defined as the union of households, non-financial corporations, other financial intermediaries, insurance corporations, and pension funds. This means that the stock and flows of the series related to the private sector are obtained by summing the same series related to the different counterparties:

$$L_{t,PS} = \sum_i L_{t,i}, \qquad F_{t,PS} = \sum_i F_{t,i},$$

where i varies among the various components of the private sector PS. However, since seasonal adjustment is not a linear operator, this relationship could be not preserved when using a direct adjustment method, meaning that:

$$L_{t,PS}^{SA,direct} \neq \sum_i L_{t,i}^{SA,direct}, \qquad F_{t,PS}^{SA,direct} \neq \sum_i F_{t,i}^{SA,direct}.$$

By using the indirect approach instead, the seasonally adjusted stocks and flows of the series related to households, non-financial corporations, other financial intermediaries, insurance corporations, and pension funds, are first computed (with a direct approach), and then are summed up in order to define the adjusted series related to private sector:

$$L_{t,PS}^{SA,indirect} \stackrel{\text{def}}{=} \sum_i L_{t,i}^{SA,direct}, \qquad F_{t,PS}^{SA,indirect} \stackrel{\text{def}}{=} \sum_i F_{t,i}^{SA,direct}.$$

At this point, the seasonally adjusted index of notional stocks is computed based on adjusted stocks and flows:

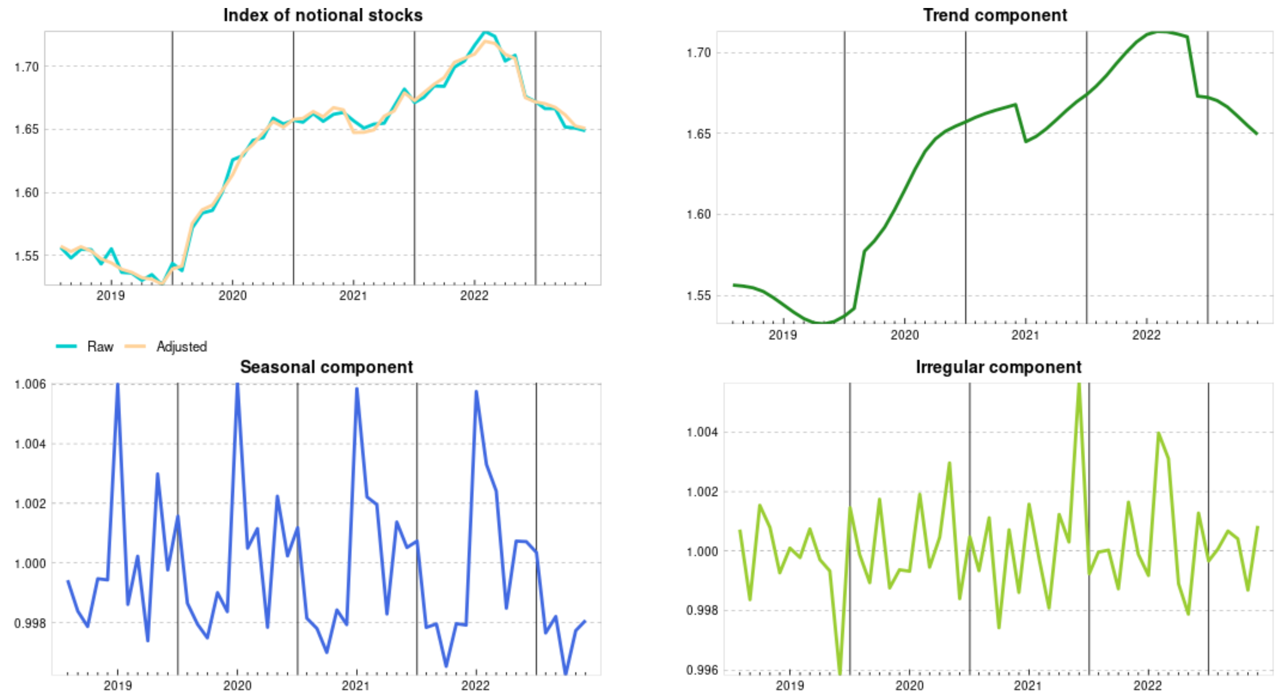
$$Z_{t,PS}^{SA} = Z_{t-1,PS}^{SA}\left(1 + \frac{F_{t,PS}^{SA,indirect}}{L_{t-1,PS}^{SA,indirect}}\right) = Z_{0,PS}^{SA}\prod_{j=0}^{t-1}\left(1 + \frac{F_{t-j,PS}^{SA,indirect}}{L_{t-j-1,PS}^{SA,indirect}}\right),$$

and adjusted growth rates can be calculated as in (17).

We now analyze the output of the seasonal adjustment performed by X-13ARIMA-SEATS on some of the series published by the Bank of Italy[19].

[19] The Bank of Italy, together with the central banks of the other three largest euro area countries (i.e. France, Germany and Spain), also transmits to the ECB further seasonal adjusted data (outstanding amounts and flows) on loans to households and non-financial corporation of the whole euro area. Those series are published by the ECB and extensively used by the ECB's DG Monetary Policy for their analysis.

Figure 3. Loans to domestic non-financial corporations



In order to show the results of the adjustment procedure, Figure 3 highlights the output related to the series of loans to domestic non-financial corporations by Italian banks[20]. The top-left panel compares the seasonally adjusted and raw series of index of notional stocks, starting from January 2019, while the other three graphs relate to the trend, seasonal, and irregular components identified by the procedure.

Some of the diagnostics produced by X-13ARIMA-SEATS and TRAMO-SEATS are presented in Table 1, whereas Figure 4 compares the raw and the seasonally adjusted growth rate outputted by the new procedure. As it is a one-month growth rate, the series exhibits some volatility. Figure 4, in particular, shows a spike at the beginning of 2020, coinciding with the outbreak of the Covid-19 pandemic. As mentioned earlier, at the end of March 2020, Eurostat released a methodological note on the treatment of pandemic crisis effects, suggesting the introduction of an outlier (at least additive) for March 2020. The specific type of outlier would have to be verified as subsequent months data became available. Analyses conducted for Italy indicated the use of a Level Shift, which indeed exhibits a high level of significance.

[20] Further details on the output of the seasonal adjustment procedure for the other series published in the monthly publication "*Banks and Money: National Data*" are available by the authors upon request.

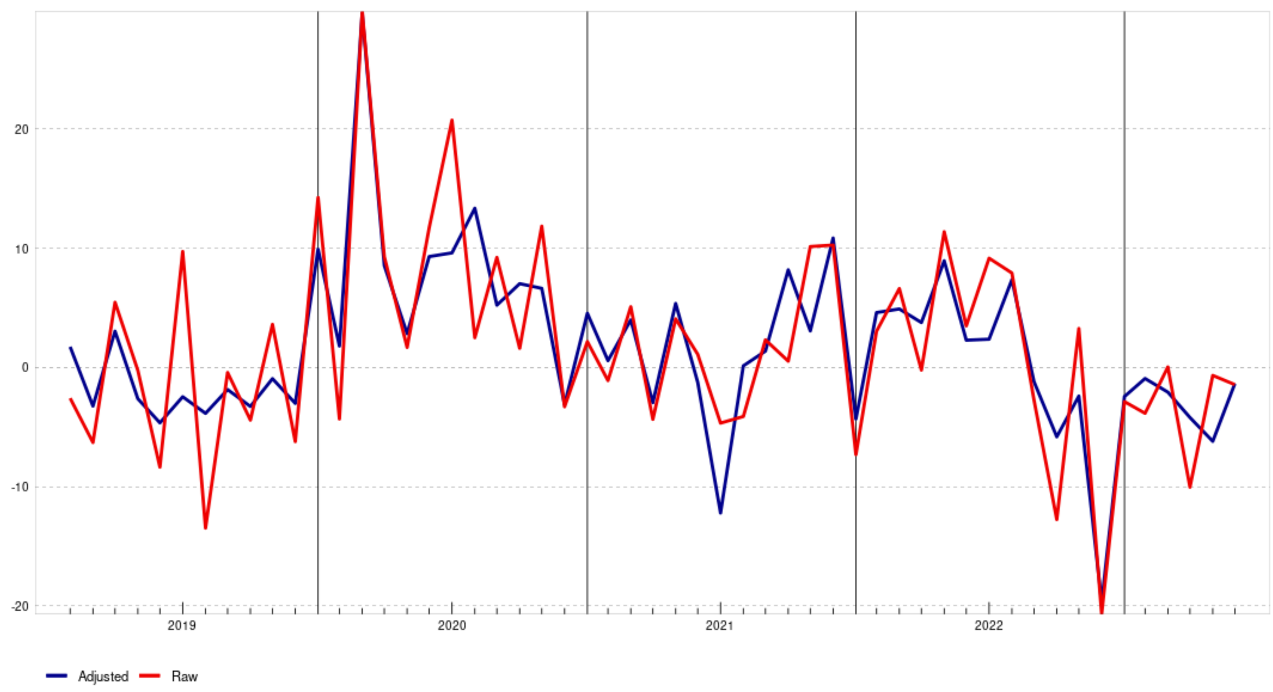
Table 1. Loans to domestic non-financial corporations

X13-ARIMA-SEATS	MODEL DEFINITION
	**Transformation:** Log(y) **ARIMA Model:** (3 1 1)(1 0 0) **Model Span:** 2003.Jan to 2023.Jun **Number of parameters estimated:** 16 **NORMALITY STATISTICS FOR RESIDUALS** **Skewness coeffi:** 0.2494 **Kurtosis:** 3.1844 **No indication of lack of normality.** **AUTOCORRELATION TEST FOR RESIDUALS** **Durbin-Watson:** 2.022167 **SEASONALITY TEST FOR RESIDUALS** **QS Statistic for Residuals:** 0.00 (P-Value = 1.0000) **SAMPLE AUTOCORRELATIONS OF THE RESIDUALS** **Ljung-Box (lag=24):** 20.47 (P-Value: 0.37)
TRAMO-SEATS	**MODEL DEFINITION**
	**Transformation:** Log(y) **ARIMA Model:** (3 1 1)(1 0 0) **Model Span:** 2003.Jan to 2023.Jun **Number of parameters estimated:** 16 **NORMALITY STATISTICS FOR RESIDUALS** **Skewness coeffi:** 0.27 **Kurtosis:** 3.0892 **No indication of lack of normality.** **AUTOCORRELATION TEST FOR RESIDUALS** **Durbin-Watson:** 1.9778 **SEASONALITY TEST FOR RESIDUALS** **QS Statistic for Residuals:** 0.00 (P-Value = 1.0000) **SAMPLE AUTOCORRELATIONS OF THE RESIDUALS** **Ljung-Box (lag=24):** 20.19 (P-Value: 0.38)

All in all, the diagnostics show the adequacy of both seasonal adjustment procedures, proving that, as reported in Section 1, the choice of adopting X-13ARIMA-SEATS in place of TRAMO-SEATS is to be attributed mainly to the willingness to align the procedures in place in the Bank of Italy to the best practices used in the ECB and other NCBs in the Euro area.

Figure 5, on the other hand, represents the difference between the one-month growth rates obtained through the new methodology based on X-13ARIMA-SEATS and those obtained using programs based on TRAMO-SEATS. Positive values in the graph are associated with higher values of X-13ARIMA-SEATS estimates, whereas negative values are obtained when TRAMO-SEATS leads to higher figures. The differences refer to the series of loans to households and non-financial corporations for the period between January 2019 and June 2023, and show how the estimates obtained through the two methods are overall quite similar.

Figure 4. Loans to domestic non-financial corporations – one-month percentage changes on an annual base



3.2 New adjusted series for loans to non-financial corporations by sector of economic activity and to households by purpose of lending

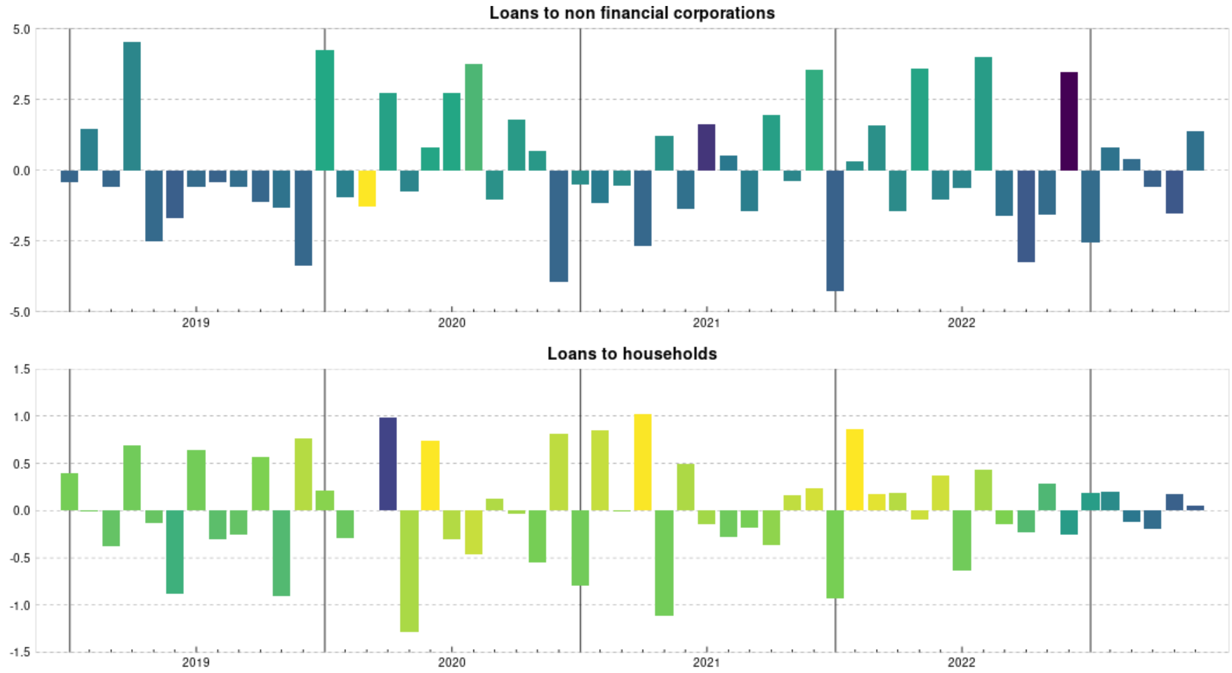
In addition to the production of the series already published by the Bank of Italy, the X-13ARIMA-SEATS procedure has also been adopted in compiling new seasonally adjusted series of loans to households categorized by purpose of lending, and to non-financial corporations, split by branch of economic activity, from which one-month percentage changes on an annual basis are derived.

These are more detailed series compared to the ones currently disseminated, and they may be of interest for studying the short and medium-term dynamics of credit disbursements to the private sector in Italy.

More precisely, for loans to households, three new series have been produced referred to:

1. loans for house purchase;
2. loans for consumption (e.g. the purchase of durable goods and cars);
3. loans for other purposes.

Figure 5. One-month percentage changes – difference between estimates obtained from the new methodology of X-13ARIMA-SEATS and TRAMO-SEATS



The computation of the index of notional stocks for these three series was made possible by the availability of the corresponding stocks and flows in the Bank of Italys internal statistical data warehouse, starting from the reference date of February 2003. As for the seasonal adjustment procedure, initially the same parameters defined for the series of total loans to households were used, and their diagnostics were compared with those obtained from the automatic X-13ARIMA-SEATS procedure. The best parametrization has been found by means of an iterative process of comparison and fine-tuning, that took into account both economic and business reasons (especially when defining the outliers for the ARIMA model) together with purely statistical metrics resulted in the fit of the model.

Figure 6 shows how the total amount of loans to households was broken down by purpose of lending as of June 2023. Since the largest share was represented by lending for house purchase, let us consider this subseries.

Figure 6. Loans to households by purpose of lending – stocks

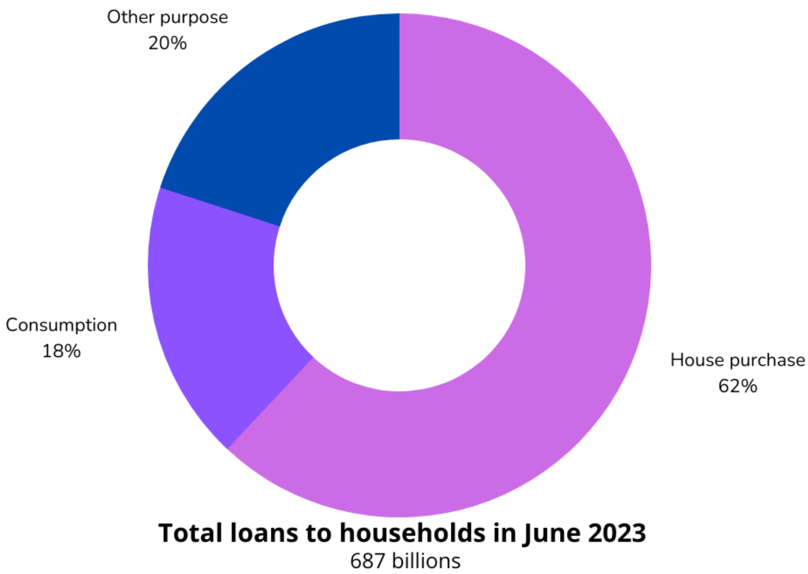


Figure 7 represents the comparison between the adjusted one-month percentage changes of this subseries and of total loans to households computed both through direct and indirect method, from January 2019 through June 2023. As can be seen, the two series have very similar dynamics, although it can be noted that, starting from 2021, the growth rate of loans for house purchase has consistently remained above the level of that of total loans to households. It is interesting to note how the seasonally adjusted estimates show a strong downward trend starting from mid-2022 up until the last observations under consideration. This dynamic is reasonably mainly related to the increase in the official rates carried out by the ECB Governing Council as countermeasure to the rise of inflation, and is not attributable to any seasonal patterns. Therefore, it is preserved in the final estimates produced by the seasonal adjustment procedure.

Figure 7. One-month percentage changes on an annual basis: loans for house purchase and total loans to households – seasonally adjusted data

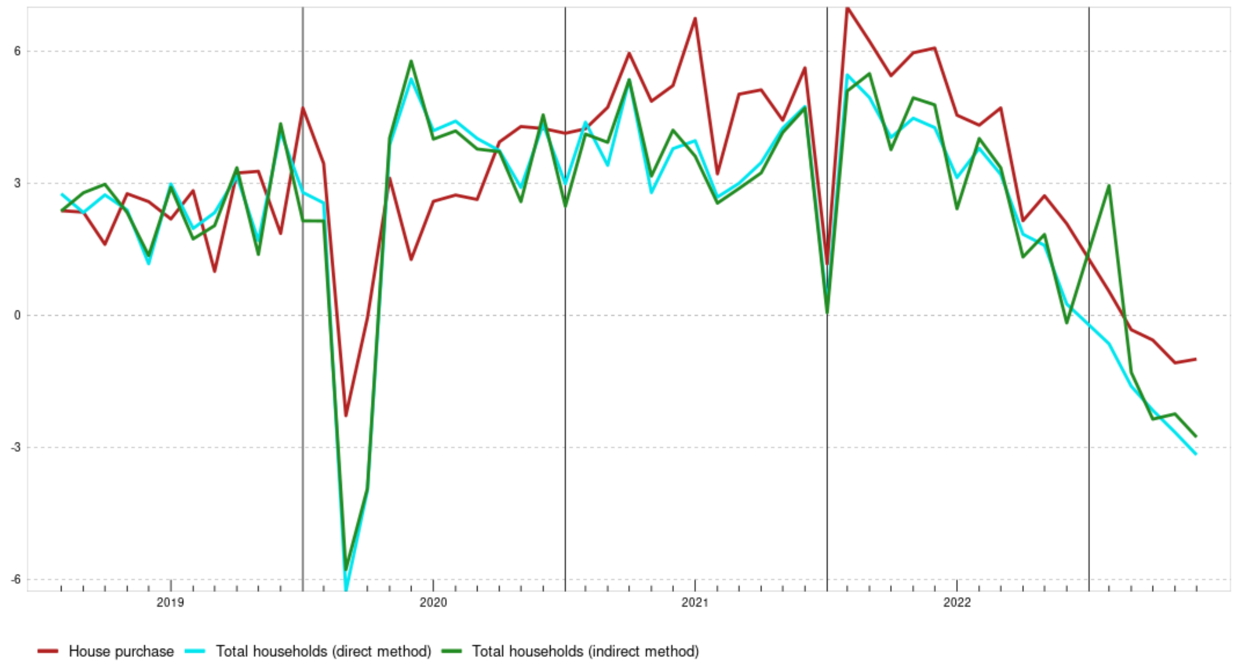


Table 2 summarizes some of the most relevant statistics and diagnostic checks produced by X-13ARIMA-SEATS that are used to assess the quality of the seasonally adjusted estimates. Among them, there are all the likelihood statistics of the ARIMA model, the Ljung-Box diagnostic in testing the sample autocorrelations of the residuals, and the normality and seasonality statistics for the RegARIMA model residuals. All these diagnostics report values within the pre-established significance thresholds, suggesting a satisfactory quality for the seasonally adjusted estimates produced.

As for loans to non-financial corporations, the categorization has been made in the following four branches of economic activity:

1. manufacturing;
2. construction;
3. real estate activities, transportation, storage, scientific activities, and other commercial, financial, legal and technical services;
4. agriculture, mining, electricity and water supply, and all remaining activities.

Table 2. Loans for house purchase

MODEL DEFINITION			
Transformation: Log(y)		**regARIMA Model Span:** 2003.Feb to 2023.Jun	
ARIMA Model: (1 2 2)(0 1 1)			
Nonseasonal AR	**Lag 1**	**Estimate:** -0.86925	**Standard Error:** 0.03706
Nonseasonal MA	**Lag 1**	**Estimate:** -0.41547	**Standard Error:** 0.05865
	Lag 2	**Estimate:** 0.58417	**Standard Error:** 0.05535
Seasonal MA	**Lag 12**	**Estimate:** 0.23889	**Standard Error:** 0.06324
LIKELIHOOD STATISTICS			
Log likelihood: 1188.9523	**AIC:** -1973.7015	**AICC:** -1971.4689	**BIC:** -1922.0652
DIAGNOSTIC CHECKING			
Sample Autocorrelations of the Residuals with the Ljung-Box diagnostic - Lag 24			
ACF: -0.08	**SE:** 0.07	**LB Q:** 35.95 (DF: 20)	**P:** 0.02
The P-values approximate the probability of observing a Q-value at least this large when the model fitted is correct. When DF is positive, small values of P, customarily those below 0.05, indicate model inadequacy.			
Normality Statistics for regARIMA Model Residuals			
Number of residuals: 231	**Skewness coeffi:** -0.0869	**Geary's a:** 0.7720	**Kurtosis:** 3.6804
No indication of lack of normality.			
QS Seasonality Statistic for regARIMA Model Residuals			
QS Statistic for regARIMA Model Residuals (full series): 0.02 (P-Value = 0.9877)			

In this case, however, the series for the stocks and flows available in the statistical data warehouse do not exhibit great historical depth, starting only from March 2020. In general, this would be sufficient, since seasonal adjustment of monthly time series typically requires a minimum of 36 observations, corresponding to a 3-year time span. However, to improve the quality of the estimates, an effort was made to reconstruct the series backward, reaching as far back as January 2018. To this end, we used data from internal archive reporting not fully harmonized information on loans to non-financial corporations split by branch of economic activity at the provincial level for the period before March 2020. Moreover, in order to ensure that both the stocks and flows of the subseries summed up to the total loans to non-financial corporations, some imputations on data needed to be made. In particular, we distributed the differences in the four subseries based on their proportions with respect to the total. This solution, while not ideal in itself, is justified by the overall small size of the corrections made, which did not alter the dynamics of the series and the seasonal adjustment estimates. As for the definition of the input parameters for the X-13ARIMA-SEATS procedure, the same process of fine-tuning described above in the case of loans to households by purpose of lending was applied for these series as well.

Figure 8. Loans to non-financial corporations by branch of economic activity – stocks

(billions of euros)

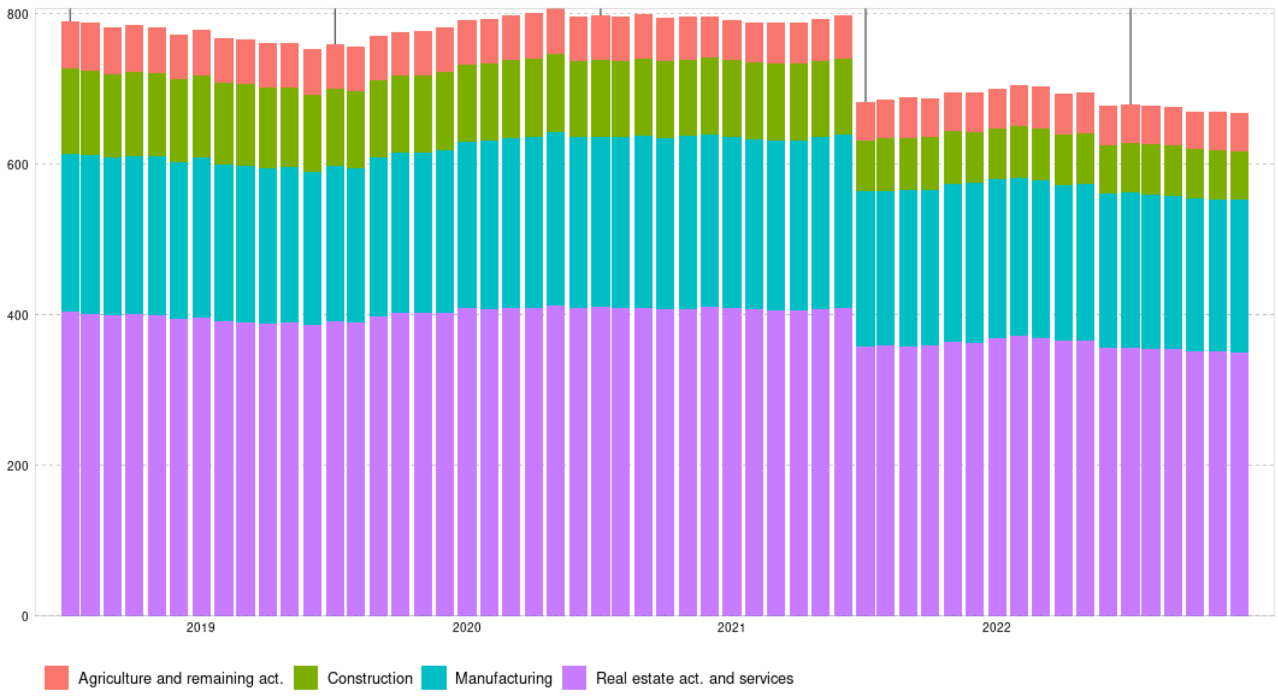
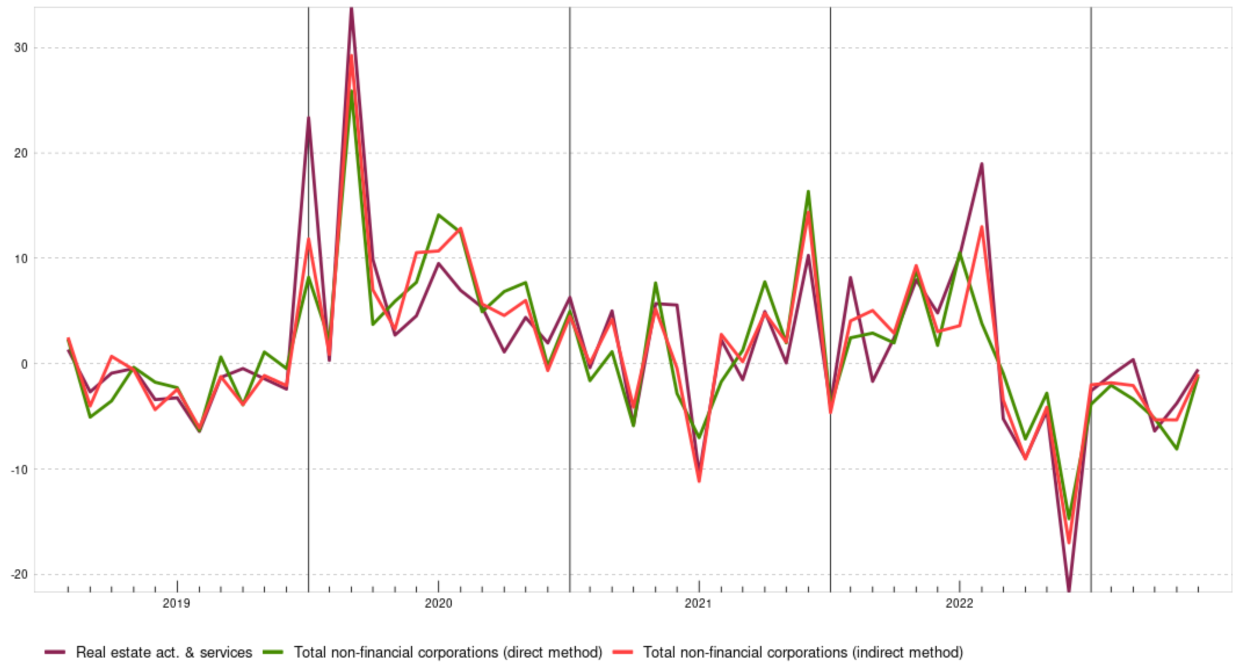


Figure 8 shows the distribution of the stock of loans to non-financial corporations between the four different branches of economic activity taken in consideration, for the period between January 2019 and June 2023. One can notice that the composition remains stable over time, with the largest share held by the sector of real estate activities and services. The sharp decrease in volumes that can be observed in 2022 is due to a change in the ECB Regulation for the MFI Balance Sheet Statistics, whereby, as of January 2022, the stock of securitised and derecognised loans no longer includes those that are managed by financial servicers.

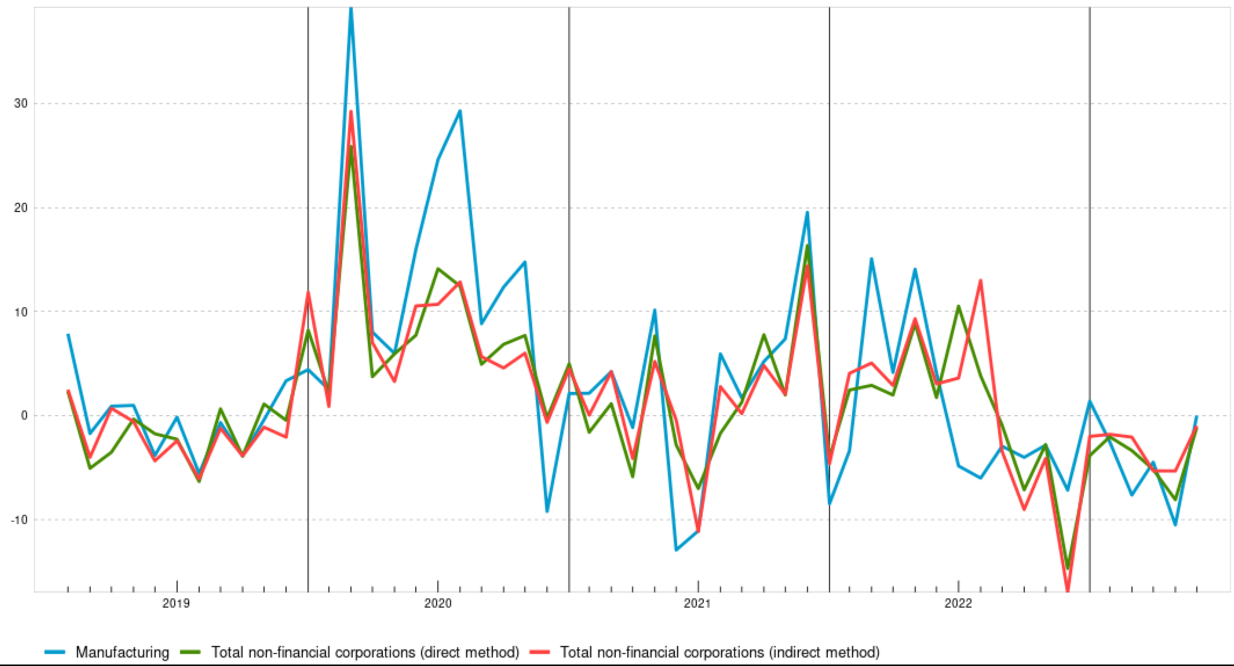
Let us now consider, for example, the series of loans to non-financial corporations in the sector of real estate activities and services. Figure 9 represents the comparison between the seasonally adjusted growth rates of the subseries with respect to the series of total loans to non-financial corporations. The reference period is from January 2019 to June 2023.

Figure 9. One-month percentage changes on an annual basis: loans to real estate activities and services and total loans to non-financial corporations – seasonally adjusted data



Apart from some small differences in the magnitude of the oscillations (particularly in January 2020), the dynamics of the two series are very similar. Greater differences can instead be found when comparing loans to non-financial corporations with those to the manufacturing sector (Figure 10). In this case, the very high values reached by the growth rate for the subseries in the central months of 2020 stand out, particularly in August. In fact, while the activity in the Italian banking sector is typically reduced during summer months, in August 2020, loans to the manufacturing sector recorded high transactions, which, not being ascribable to the seasonal component, translated into the peak that can be observed in the seasonally adjusted growth rate.

Figure 10. One-month percentage changes on an annual basis: loans to manufacturing activities and total loans to non-financial corporations – seasonally adjusted data



4. Conclusions

The seasonal adjustment procedures for macroeconomic and financial time series are essential for accurately interpreting the data. In the context of a monetary union such as the euro area, it is desirable that these procedures for financial time series align with the practices used by the ECB and by most of other central banks.

After reviewing the main seasonal adjustment techniques, we first discuss and compare the output produced by different methodologies applied to the credit time series currently published by the Bank of Italy. We then explain the rationale of the switch from TRAMO-SEATS to X13-ARIMA-SEATS, which represents the latest version of the procedure currently used by the ECB.

Finally, we show the implementation of the selected seasonal adjustment method also in the derivation of new official adjusted time series for banks' loans to non-financial corporations by NACE and to households by purpose of lending.

References

Abeln, B. and J. P. A. M. Jacobs (2022). COVID-19 and Seasonal Adjustment, *Journal of Business Cycle Research*, 18, 159-169.

Akaike, H. (1980). Seasonal Adjustment by a Bayesian Modeling, *Journal of Time Series Analysis*, 1(1): 1-13.

Bell, W., Holan, S. and T. McElroy (2012). *Economic Time Series: Modeling and seasonality.* New York, CRC Press.

Battipaglia, P. and Focarelli, D. (1996). A comparison of indicators for evaluating X-11-ARIMA seasonal adjustment, *Journal of the Italian Statistical Society*, 2, 179-202.

Box, G.E.P. and Jenkins, G.M. (1970). Time Series Analysis: Forecasting and Control, Holden Day, San Francisco.

Buys Ballot, C. H. D. (1847). Les Changements Périodiques de Température, Kemink and Fils, Utrecht.

Cabrero, A. (2000). Seasonal Adjustment In Economic Time Series: The Experience Of The Banco De España (With The Model-Based Method), Documento de Trabajo 0002, Banco De España.

Chang, I., Tiao, G.C., and Chen, C. (1988). Estimation of Time Series Parameters in the Presence of Outliers, Technometrics, 30, 193-204.

Chung, C. and L. Lon-Mu (1993). Joint Estimation of Model Parameters and Outlier Effects in Time Series, *Journal of the American Statistical Association*, 88:(421), 284-297.

Cividini, A. and Cottarelli, C. (1989). Procedure di destagionalizzazione dei depositi bancari mensili in Italia, Temi di Discussione (Economic Working Papers) 112, Bank of Italy, Economic Research and International Relations Area.

Cleveland, R. B. (1990). STL: A seasonal-trend decomposition procedure based on loess, *Journal of Official Statistics*, 6:(1), 3-73.

Dagum E.B. (1980). La méthode de désaisonnalisation X-11-ARMMI, Statistique Canada, Ottawa.

Eurostat (2015). *ESS guidelines on seasonal adjustment*. Eurostat, Manuals and Guidelines.

Eurostat (2017). *Handbook on Methodology of Modern Business Statistics*. Eurostat, Manuals and Guidelines. https://cros-legacy.ec.europa.eu/content/handbook-methodology-modern-business-statistics_en

Eurostat (2018). *Handbook on Seasonal Adjustment*. Eurostat, Manuals and Guidelines. https://ec.europa.eu/eurostat/web/products-manuals-and-guidelines/-/ks-gq-18-001

Eurostat (2020). Guidance On Time Series Treatment In The Context Of The Covid-19 Crisis. Methodological Note, Eurostat, Directotate B.

Ferrara, Laurent, Lardiray, Dominique, and Darné, Olivier. (2018). A Brief History of Seasonal Adjustment Methods and Software Tools.

Findley, D.F., Monsell, B.C., Otto, M.C., Bell, W.R., and Pugh, M.G. (1988). Toward X-12-ARIMA, *Proceeding of the Fourth Annual Research Conference*, US Bureau of the Census, Washington DC.

Findley, D.F., Monsell, B.C., Bell, W.R., Otto, M.C., and Chen, B.C. (1998). New capabilities and methods of the X-12-ARIMA seasonal adjustment program, *Journal of Business and Economic Statistics*, 16, 127-152.

Findley, D. F. and Monsell, B. C. (2009). Modelling Stock Trading Day Effects Under Flow Day-of-Week Effect Constraints. *Journal of Official Statistics.* 25:(3), 415-430

Fonteny, É. (2006). La désaisonnalisation des séries d'agrégats monétaires et de crédit à la Banque de France: théoriques et mise en oeuvre. Notes d'études de recherche 147, Banque de France, direction générale des études et des relations internationales direction de la rcherche.

Fox, A. J. (1972). Outliers in Time Series, *Journal of the Royal Statistical Society*, 34 Series B (Methodological), 350-363.

Gómez, V., and Maravall, A. (1992). Time series regression with ARIMA noise and missing observations: Program TRAMO, Working Paper ECO No 92-81, Department of Economics, European University Institute.

Gómez, V., and Maravall, A. (1997). Programs TRAMO and SEATS: Instructions for the user (beta version: June 1997), Working Paper No 97001, Dirección General de Anélisis y Programación Presupuestaria, Ministerio de Economía y Hacienda, Madrid.

Grether, D. M. and M. Nerlove (1970). Some properties of "optimal" seasonal adjustment, *Econometrica*, 38(5): 682-703.

Harvey, A. C. (1990). Forecasting, Structural Time Series Models and the Kalman Filter, Cambridge University Press.

Jevons, W.S. (1862). On the study of periodic commercial fluctuations, *Investigations in Currency and Finance*, Macmillan, London.

Kaiser, R. and Maravall, A. (1992). Seasonal Outliers in Time Series, *Estadistica*, 53.

Ladiray, D., and Quenneville, B. (2001). Seasonal Adjustment with the X-11 Method, *Lecture Notes in Statistics*, Springer New York, NY.

OECD (2020). Impact of COVID-19 crisis on official statistics: business continuity of official statistics. Issue Note for the 17th meeting of the Committee on Statistics and Statistical Policy, 24-26 June 2020, OECD.

Maravall, A., and Caporello, G. (2004). Program TSW: Revised reference manual. Technical Report, Research Department, Bank of Spain.

Maravall, A. and Gómez, V. (1992). Signal extraction in ARIMA time series: Program SEATS, Working Paper ECO No 92-65, Department of Economics, European University Institute.

Matas Mir A. and V. Rondonotti (2003). The performance of X-12 in the seasonal adjustment of short time series. In M. Manna and R. Peronaci (Eds) *Seasonal Adjustment*, 149-159, ECB.

Pedersen, M. K. and C. F. Fæste. (2006). Seasonal adjustment of Danish financial time series using the X-12-ARIMA procedure, Danmarks Nationalbank Working Papers 44, Danmarks Nationalbank, Copenhagen.

Poměnková J. and E. Klejmova (2019). An Influence of a Seasonal Adjustment on the Time-Frequency Wavelet Modelling of a Bank Loans Activities, Proceedings of the World Congress on Engineering 2019, 3-5 July 2019, London, U.K.

Silvestrini, A. (2009). Seasonal adjustment of bank deposits and loans, Questioni di Economia e Finanza (Occasional Papers) 42, Bank of Italy, Economic Research and International Relations Area.

Tsay, R.S. (1988). Outliers, Level Shifts, and Variance Changes in Time Series, Journal of Forecasting, 7, 1-20.

US Bureau of the Census (2000). X-12-ARIMA: Reference Manual (version 0.2.6), Time Series Staff Statistical Research Division, US Bureau of the Census, Washington DC.

US Bureau of the Census (2017). X-13ARIMA-SEATS Reference Manual, Version 1.1, Time Series Research Staff, Center for Statistical Research and Methodology, U.S. Census Bureau, Washington, DC.