DEUTSCHE
BUNDESBANK
EUROSYSTEM

Discussion Paper

Deutsche Bundesbank
No 07/2016

A data-driven selection of an appropriate seasonal adjustment approach

Karsten Webel

Deutsche Bundesbank, Wilhelm-Epstein-Straße 14, 60431 Frankfurt am Main,
Postfach 10 06 02, 60006 Frankfurt am Main

Tel +49 69 9566-0

Please address all orders in writing to: Deutsche Bundesbank,
Press and Public Relations Division, at the above address or via fax +49 69 9566-3077

Internet http://www.bundesbank.de

ISBN 978–3–95729–239–1 (Printversion)
ISBN 978–3–95729–240–7 (Internetversion)

Non-technical summary

Research Question

Recent releases of X-13ARIMA-SEATS and JDemetra+ have led to a paradigm shift since both seasonal adjustment programs unify the non-parametric X-11 method and the parametric ARIMA model-based approach under one umbrella. Users can thus choose the better methodological alternative for each time series being studied within a single software package. For that purpose, however, selection criteria – which should be as general as possible – have to be identified and combined appropriately to form a sound decision-making process.

Contribution

We suggest a decision tree that takes into account conceptual differences between the two approaches as well as statistical characteristics of the studied time series. The key theoretical issue is whether the observations can be modelled sufficiently well by an ARIMA process. The key empirical issue is to measure the quality of the seasonally adjusted figures obtained from each approach.

This decision tree is especially suited for situations where a given time series is seasonally adjusted for the first time. For institutionals, it may provide additional support for regular (but less frequent) in-depth revisions of specification files stored in institutional databases.

Results

We illustrate the decision tree using four German macroeconomic time series. Results show the X-11 method tends to be better for longer time series, whereas the ARIMA model-based approach is better for shorter and moderate-length ones. A possible explanation may be that some phenomena which in theory cannot be captured adequately by ARIMA processes, such as seasonal heteroskedasticity, require a large amount of data to appear significantly in the observations.

Nichttechnische Zusammenfassung

Fragestellung

Die jüngsten Veröffentlichungen von X-13ARIMA-SEATS und JDemetra+ führten zu einem Paradigmenwechsel, da beide Saisonbereinigungsprogramme die nicht-parametrische X-11-Methode und den parametrischen ARIMA-modellbasierten Ansatz unter einem Dach vereinen. Ihre Nutzer können somit für jede zu bereinigende Zeitreihe das bessere Verfahren innerhalb eines einzigen Softwarepakets auswählen. Jedoch müssen dafür möglichst allgemeine Vergleichskriterien gefunden und zu einem durchdachten Entscheidungsprozess zusammengefügt werden.

Beitrag

Wir schlagen einen Entscheidungsbaum vor, der sowohl konzeptionelle Unterschiede zwischen den beiden Ansätzen als auch statistische Eigenschaften einer gegebenen Zeitreihe berücksichtigt. Den Kern unserer theoretischen Überlegungen bildet die Frage, ob sich die Beobachtungen hinreichend gut durch einen ARIMA-Prozess modellieren lassen. Im Fokus des darauffolgenden empirischen Teils steht die Beurteilung der Qualität der mit beiden Ansätzen bestimmten saisonbereinigten Angaben.

Dieser Entscheidungsbaum eignet sich vor allem für die erstmalige Saisonbereinigung einer gegebenen Zeitreihe. Institutionelle Nutzer können ihn zusätzlich unterstützend im Rahmen regelmäßig (aber nicht allzu häufig) durchzuführender umfangreicher Revisionen der in ihren Datenbanken gespeicherten Spezifikationsdateien einsetzen.

Ergebnisse

Wir illustrieren den Entscheidungsbaum anhand vier makroökonomischer Zeitreihen für Deutschland. Dabei zeigt sich, dass die X-11-Methode tendenziell etwas vorteilhafter für längere Zeitreihen ist, während der ARIMA-modellbasierte Ansatz eher für vergleichsweise kürzere Zeitreihen bevorzugt wird. Eine mögliche Erklärung dafür könnte sein, dass manche Phänomene, die theoretisch nicht adäquat durch ARIMA-Prozesse abgebildet werden können, etwa saisonale Heteroskedastizität, eine große Anzahl an Daten benötigen, um nachhaltig in den Beobachtungen sichtbar zu werden.

Bundesbank Discussion Paper No 07/2016

A data-driven selection of an appropriate seasonal adjustment approach*

Karsten Webel
Deutsche Bundesbank

Abstract

Recent releases of X-13ARIMA-SEATS and JDemetra+ enable their users to choose between the non-parametric X-11 and the parametric ARIMA model-based approach to seasonal adjustment for any given time series without the necessity of switching between different software packages. To ease the selection process, we develop a decision tree whose branches combine conceptual differences between the two methods with empirical issues. The latter primarily include a thorough inspection of the squared gains of final X-11 and Wiener-Kolmogorov seasonal adjustment filters as well as a comparison of various revision measures. We finally illustrate the decision tree on selected German macroeconomic time series.

Keywords: ARIMA model-based approach, linear filtering, signal extraction, unobserved components, X-11 approach.

JEL classification: C13, C14, C22.

---

*Contact address: Karsten Webel, Deutsche Bundesbank, Central Office, Statistics Department and Research Centre, Wilhelm-Epstein-Strasse 14, 60431 Frankfurt am Main, Germany. Phone: +49 69 9566 2702. E-mail: karsten.webel@bundesbank.de. The author thanks Edgar Brandt, Gary Brown, Robert Kirchner, Andreas Lorenz, Agustín Maravall, Jens Mehrhoff, Jörg Meier, an anonymous referee and seminar participants at the Bundesbank for helpful comments.

1 Introduction

Many macroeconomic monthly or quarterly time series exhibit a considerable amount of seasonal variation, ie predictable fluctuations that recur each year in the same period with similar intensity. Since it is generally agreed that economic developments are best judged from indicators revealing new information, seasonal adjustment, ie the removal of predictable seasonal variation, has become a standard tool in both official statistics and academic research. By osmosis, different adjustment strategies have been developed over the last decades (see Hylleberg (1992), Ghysels and Osborn (2001), Bell, Holan, and McElroy (2012) for respective overviews), with the X-11 and ARIMA model-based (AMB) approaches being utilised most commonly in both fields, and the state-space approach being considered primarily in academic research. Due to their distinct backgrounds, these approaches have been implemented in esoteric software packages: the X-11 method initiated by Shiskin, Young, and Musgrave (1967) has been embedded in the X-11 package and its successors such as X-12-ARIMA (see Findley, Monsell, Bell, Otto, and Chen (1998)), the AMB approach has been integrated into the TRAMO-SEATS package and its Windows companion TSW documented by Gómez and Maravall (1996) and Caporello, Maravall, and Sánchez (2001), and the state-space approach of Harvey (1989) has been implemented in the STAMP package (see Koopman, Harvey, Doornik, and Shephard (1996)).

As a consequence, decisions of many practitioners at statistical agencies on their preferred seasonal adjustment approach have directly depended on the choice of their favourite seasonal adjustment package, see European Central Bank (2000) for an in-depth discussion. In many cases, the choice has been based on pragmatic reasons, such as employees' individual backgrounds, data users' demands or the package's suitability for statistical mass production. Then, all time series, or at least broad subsets thereof, have been seasonally adjusted according to the single method implemented in the favourite software package.

Recent releases of X-13ARIMA-SEATS and JDemetra+ may change decisions as they jointly incorporate both the X-11 and AMB approaches alongside regARIMA/TRAMO pretreatment. Users of either package may thus choose between the two methods for each particular time series under review without the necessity of switching software, bearing in mind differences between the versions of X-11 and SEATS in the two packages and differences between either version of SEATS and TRAMO-SEATS/TSW. This immediately raises the question of which criteria should guide the choice of approach. For that purpose, we propose a decision tree that combines theoretical considerations, regarding conceptual differences between the philosophies underlying the X-11 and AMB approaches, with empirical findings. Elaborating on previous work done by Webel (2013a,b), our idea extends (purely empirical) comparisons between the two methods, see Scott, Tiller, and Chow (2007).

It should be noted, however, that we do not suggest applying this decision tree as part of those regular seasonal adjustment routines which recur monthly or quarterly. Seasonally adjusted figures, in particular those designated for official use, should be calculated consistently over time. Therefore, changing the seasonal adjustment approach every time new unadjusted data become available is not the aim of the decision tree. This is in line with the general recommendation to avoid concurrent adjustment due to several drawbacks, such as the high frequency of revisions and the increased risk of a highly instable seasonal pattern, see Item 4.2 of Eurostat (2015). Similarly, consistency of seasonal adjustment approaches among components of aggregates is recommended. However, we do not discuss this issue, and whether a direct or an indirect approach should be preferred to obtain the seasonally adjusted aggregate, see Items 3.4 and 3.5 of Eurostat (2015) for further information. Therefore, our decision tree should be understood as guidance for seasonally adjusting a given time series for the first time, or to assist regular reviews of existing seasonal adjustment specification files, which is usually done less frequently, for example annually. For aggregate series, one may apply our decision tree to all component series and then choose the approach based on the most important components or the majority of components.

The remainder of this paper is organised as follows. Section 2 provides the notational framework as well as basic ideas of both seasonal adjustment methods considered. The decision tree is presented in Section 3, followed by a detailed empirical illustration in Section 4 which uses selected macroeconomic time series for Germany. Finally, Section 5 draws some conclusions.

2 Background

2.1 Notations

Let $\{x_t\}$ denote the time series under review and assume that it can be decomposed additively, possibly after taking logs, into two orthogonal unobserved components (UCs) according to

$$x_t = s_t + n_t, \tag{1}$$

where $\{s_t\}$ is the signal that captures the non-seasonal variation and $\{n_t\}$ is the noise that contains the seasonal fluctuations. We assume for convenience that $\{x_t\}$ is linearised in the sense that it has been cleaned temporarily for outliers and permanently for calendar effects. Seasonal adjustment is thus interpreted as a signal extraction problem which can be solved by linear filtering. To see this, let $\mathbf{x} = (x_1, \dots, x_T)^\top$ denote a finite realisation of $\{x_t\}$. We may then express any estimator of the seasonally adjusted time series as

$$\hat{\mathbf{s}} = \mathcal{W}\,\mathbf{x}, \tag{2}$$

where $\hat{\mathbf{s}} = (\hat{s}_1, \dots, \hat{s}_T)^\top$ and $\mathcal{W} \in \mathbb{R}^{T\times T}$ is a matrix whose t-th row, $\mathcal{W}^{(t)}$, contains the filter weights employed to estimate s_t. Accordingly, we may rewrite Equation 2 for each $t \in \{1, \dots, T\}$ as

$$\hat{s}_t = \mathcal{W}^{(t)}\,\mathbf{x} = \sum_{j=1}^{T} \mathcal{W}_j^{(t)}\,x_j,$$

which reveals that the weights stored in $\mathcal{W}$ do not only depend on t but also on the number of observations available, T. Assuming for convenience that T is odd, ie $T = 2k+1$ for some $k \in \mathbb{N}$, the (symmetric) central seasonal adjustment filter $\mathcal{W}^{(k+1)}$ and the (asymmetric) concurrent seasonal adjustment filter $\mathcal{W}^{(T)}$ are of the greatest interest.

To judge the performance of any seasonal adjustment filter, we operate in the spectral rather than the time domain. Let ψ denote any linear filter transforming data $\{x_t\}$ into output $\{y_t\}$ via $y_t = \sum_j \psi_j\,x_{t-j}$. Its gain is defined as

$$g_\psi(\lambda) = \left|\sum_j \psi_j\,e^{-ij2\pi\lambda/\tau}\right|, \tag{3}$$

where τ is the seasonal period, ie $\tau = 12$ for monthly data and $\tau = 4$ for quarterly data, and $\lambda \in [0, \tau/2]$ is in units of cycles per year. If $\{x_t\}$ is a stationary time series with spectral density $f_x(\lambda)$, then the spectral density of $\{y_t\}$ is given by $f_y(\lambda) = g_\psi^2(\lambda)\,f_x(\lambda)$. The squared gain thus signals suppression of the input series' variance component over frequency bands where $g_\psi^2(\lambda) < 1$ and amplification where $g_\psi^2(\lambda) > 1$. This principle also applies in a more general way if both input and output are non-stationary, which is the standard case in seasonal adjustment, see Section 2 of Findley and Martin (2006).

To handle non-stationarity, seasonal ARIMA models are usually employed although the motivation is quite different for the X-11 and AMB seasonal adjustment approaches. A time series $\{x_t\}$ is called a seasonal ARIMA process if there exists a white noise $\{\varepsilon_t\}$ such that

$$\phi(B)\Phi(B^\tau)\nabla^d\nabla_\tau^D(\{x_t\}) = \theta(B)\Theta(B^\tau)(\{\varepsilon_t\}),$$

where B is the backshift operator, ie $B^k x_t = x_{t-k}$ for $k \in \mathbb{Z}$, $\nabla = 1 - B$, $\nabla_\tau = 1 - B^\tau$, and $(d, D) \in \mathbb{N}^2$. Furthermore,

$$\phi(B) = 1 - \phi_1 B - \cdots - \phi_p B^p, \quad \phi_p \neq 0,$$

and

$$\Phi(B^\tau) = 1 - \Phi_1 B^\tau - \cdots - \Phi_P B^{\tau P}, \quad \Phi_P \neq 0,$$

denote the non-seasonal and seasonal AR polynomials. Analogously,

$$\theta(B) = 1 - \theta_1 B - \cdots - \theta_q B^q, \quad \theta_q \neq 0,$$

and

$$\Theta(B^\tau) = 1 - \Theta_1 B^\tau - \cdots - \Theta_Q B^{\tau Q}, \quad \Theta_Q \neq 0,$$

are the non-seasonal and seasonal MA polynomials. Thereby, $k \in \mathbb{N} \cup \{\infty\}$ for each $k \in \{p, q, P, Q\}$. We additionally assume that all roots of these polynomials lie outside the unit circle, ie both AR polynomials are stationary and both MA polynomials are invertible. To indicate a seasonal ARIMA model, we use the standard notation $f\,(pdq)(PDQ)_\tau$, where f is any particular Box-Cox transformation that may be applied to the values of $\{x_t\}$ prior to model fitting.

2.2 X-11 approach

Within the X-11 framework, the signal of Equation 1 is composed of a trend-cyclical component $\{t_t\}$ and an irregular component $\{i_t\}$. The former conceptually captures long-term movements and periodic fluctuations whose cycles last longer than one year, while the latter captures non-seasonal short-term random shocks. Equation 1 can thus be rewritten as

$$x_t = t_t + n_t + i_t. \tag{4}$$

Apart from the prior beliefs about the movements each UC should capture, no further assumptions, possibly expressed through a parametric model, are made about their stochastic structure. Hence, the X-11 approach is referred to as non-parametric.

Estimation of all UCs in Equation 4 is achieved through an iterative application of predefined linear filters. Each iteration comprises several subroutines which follow a set pattern. The trend-cyclical component is estimated first by smoothing the series with either a simple moving average or a Henderson filter, depending on the particular subroutine. The latter type of filter can be chosen manually by the user or automatically according to the I/C ratio, which is the ratio between the average absolute growth rates of a temporarily estimated irregular component and a temporarily estimated trend-cyclical component. An estimate of the seasonal component is then obtained from smoothing the detrended series, also called seasonal-irregular component, for each period (month/quarter) with a weighted $3 \times k$ seasonal moving average, where $k \in \{1, 3, 5, 9, 15\}$. As for Henderson filters, $3 \times k$ seasonal filters can be chosen manually or automatically based on the (global) I/S ratio, which is the ratio between the average absolute annual growth rates of a temporarily estimated irregular component and a temporarily estimated seasonal component. In the former case, different seasonal filters may be selected for different periods. To reduce the size of distortions caused by extreme values in the seasonal-irregular component, the X-11 routine is also equipped with an automatic detection and down-weighting procedure for such extremes. Eventually, removal of estimated seasonal fluctuations from the series yields a seasonally adjusted version. Throughout each subroutine, symmetric filters are used whenever possible. Otherwise, asymmetric versions are automatically applied. To reduce their (usually undesired) side-effects, for example on the revisions of the most recent seasonally adjusted figures, regARIMA forecasts of the unadjusted figures can be used as they enable application of less asymmetric or even totally symmetric filters, especially at the current end of the time series under review.

Final X-11 seasonal adjustment filters, whose weights are stored in the matrix $\mathcal{W}$ of Equation 2, thus result from the convolution of all symmetric and asymmetric trend and seasonal filters chosen during each iteration step. Further details on the X-11 method are provided by Findley et al. (1998), Section 4 of Ghysels and Osborn (2001) and Ladiray and Quenneville (2001), amongst others.

2.3 AMB approach

Within the AMB framework, the non-seasonal intra-year variability in the unadjusted figures is further specified by the irregular component of Equation 4 being decomposed as

$$i_t = r_t + w_t,$$

where $\{r_t\}$ is the transitory component, which is introduced to account for fluctuations not persistent enough to be considered trend-cyclical, but still too persistent to reflect behaviour like white noise, and $\{w_t\}$ is white noise, which is primarily assumed to facilitate testing and statistical interpreting. Equation 1 thus turns into

$$x_t = t_t + n_t + r_t + w_t. \tag{5}$$

The key theoretical assumption of the AMB approach states that each UC of Equation 5 can be represented as an individual ARIMA process. Hence, this approach is referred to as parametric. Excellent discussions of the stochastic structure of the trend-cyclical, seasonal and white noise components are given by Maravall (1987, 1989, 1993). Sticking to Equation 1, we assume that the ARIMA models for signal and noise are given by

$$\phi^{(k)}(B)(\{k_t\}) = \theta^{(k)}(B)\left(\left\{\varepsilon_t^{(k)}\right\}\right),$$

where $k \in \{s, n\}$ and $\phi^{(k)}$ and $\theta^{(k)}$ are AR (including differencing) and MA polynomials whose roots are assumed to lie on or outside the unit circle without sharing common roots for each UC. Also, both the two AR and the two MA polynomials do not have common (unit) roots. The innovation sequences are uncorrelated Gaussian white noise processes with finite variances $\sigma^2_{\varepsilon^{(s)}}$ and $\sigma^2_{\varepsilon^{(n)}}$. By construction, $\{x_t\}$ can then be represented as

$$\phi^{(x)}(B)(\{x_t\}) = \theta^{(x)}(B)(\{\varepsilon_t\}), \tag{6}$$

where $\phi^{(x)}(B) = \phi^{(s)}(B)\phi^{(n)}(B)$ and

$$\theta^{(x)}(B)(\{\varepsilon_t\}) = \phi^{(n)}(B)\theta^{(s)}(B)\left(\left\{\varepsilon_t^{(s)}\right\}\right) + \phi^{(s)}(B)\theta^{(n)}(B)\left(\left\{\varepsilon_t^{(n)}\right\}\right). \tag{7}$$

If the amount of data available, $\mathbf{x}$, is infinite and both signal and noise are stationary, the minimum mean squared error (MMSE) estimator of the signal, ie the estimator $\{\hat{s}_t\}$ that minimises $\mathbb{E}\left[(s_t - \hat{s}_t)^2 \middle| \mathbf{x}\right]$, is given by $\{\hat{s}_t\} = \nu(B,F)(\{x_t\})$, where ν is the Wiener-Kolmogorov (WK) filter defined as

$$\nu(B,F) = \frac{\sigma^2_{\varepsilon^{(s)}}}{\sigma^2_{\varepsilon}} \frac{\theta^{(s)}(B)\phi^{(n)}(B)}{\theta^{(x)}(B)} \frac{\theta^{(s)}(F)\phi^{(n)}(F)}{\theta^{(x)}(F)}, \tag{8}$$

where $F = B^{-1}$, see Whittle (1963). Several authors, including Bell (1984) and Maravall (1988), demonstrate that this also holds true if $\mathbf{x}$ is finite, as in Equation 2, and both signal and noise are non-stationary.

In practice, however, the quantities forming the numerator of the WK filter are unknown and have to be derived from the ARIMA model fitted to the observed time series. The decomposition algorithm developed originally by Burman (1980), and improved by Hillmer and Tiao (1982) and Bell and Hillmer (1984) is used for this purpose. In the first step, the estimated AR polynomial $\hat{\phi}^{(\mathbf{x})}(B)$ is factorised, and its (unit) roots are assigned to either signal or noise according to their associated frequencies, which yields $\hat{\phi}^{(s)}(B)$ and $\hat{\phi}^{(n)}(B)$. In a second step, the MA polynomials and innovation variances of both signal and noise are derived from a partial fraction decomposition of $\hat{\theta}^{(\mathbf{x})}(B)\,\hat{\theta}^{(\mathbf{x})}(F)$. To achieve a unique decomposition, the canonical assumption is made that the variance of the white noise component is maximised, coinciding with the minimisation of the variances of all other UCs. If necessary, several approximations to the ARIMA model fitted to $\{x_t\}$ can be considered.

Eventually, all polynomials and innovation variances constituting the WK filter are replaced with their estimators to yield the estimated WK filter. Application of this filter to the observed time series, possibly extended with ARIMA back- and forecasts, finally provides an estimate of the signal, ie the seasonally adjusted time series. Any WK filter thus coincides with a particular final SEATS seasonal adjustment filter and, accordingly, its weights directly constitute a particular row of the weighting matrix $\mathcal{W}$ occurring in Equation 2. Further details on the AMB approach are provided by the excellent overviews of Maravall (1995) and Gómez and Maravall (2001).

3 Decision tree

To choose an appropriate seasonal adjustment approach for any observed time series, we propose the three-step procedure illustrated in Figure 1. We thereby consider the time series' length as given exogenously. The reason is that different data generating processes (DGPs), and hence different seasonal patterns, may be identified for two time series of different lengths even if their DGP was the same. In addition, we assume that the amount of observations available is sufficiently large to enable reliable estimation of both regARIMA model parameters and the seasonally adjusted figures.[1] This should be kept in mind throughout this section since some phenomena studied later on, for example presence of seasonal heteroskedasticity, relate (more or less directly) to the length of the observed time series.

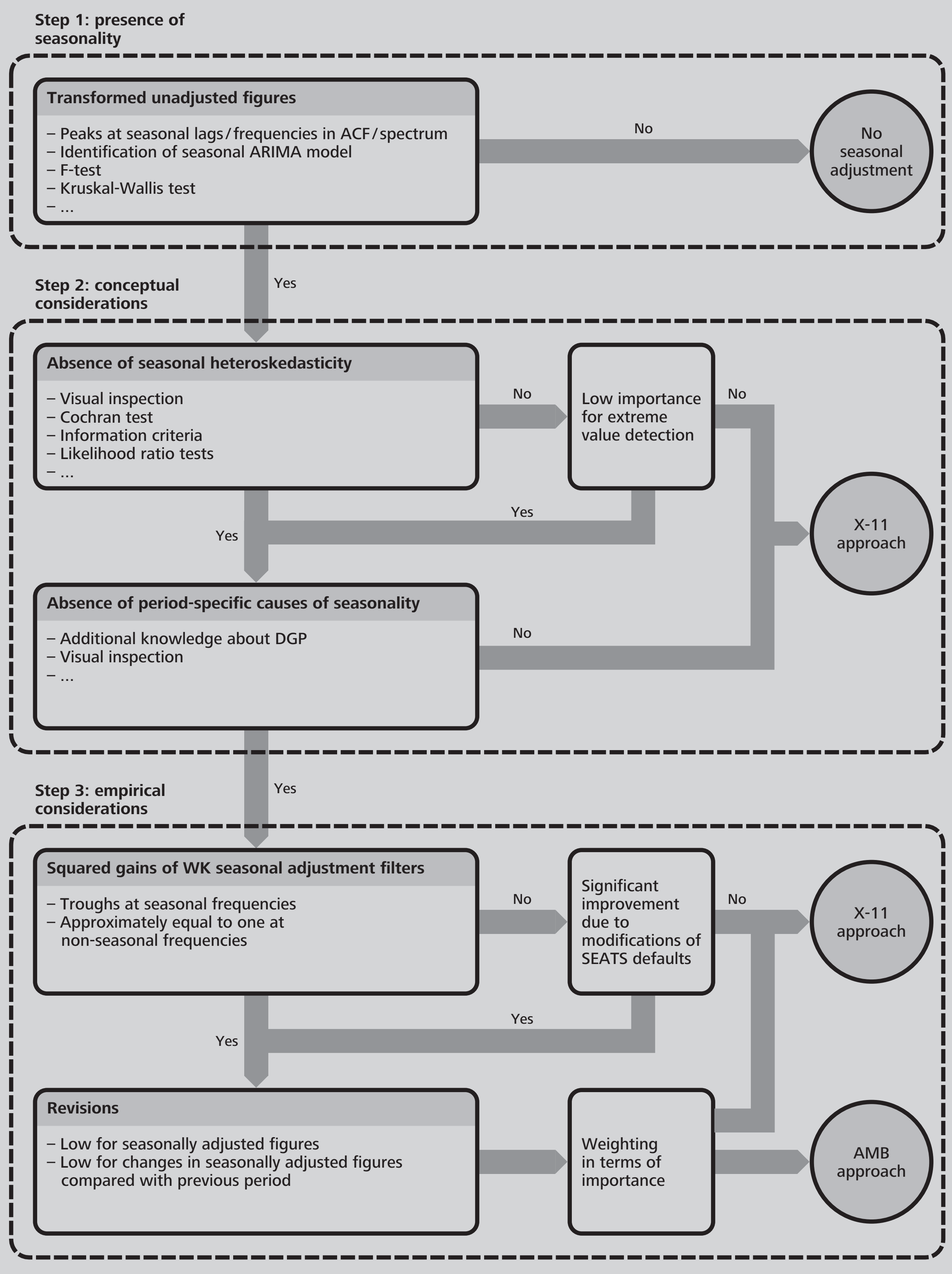
3.1 Step 1: presence of seasonality

To ensure that the time series is actually in need of seasonal adjustment, it is first examined for presence of stable seasonality. To do this, both descriptive statistics and formal tests can be calculated. Regarding the first category, the autocovariance function and the periodogram of the transformed and/or differenced time series should be checked for (positive) peaks at seasonal lags and frequencies, respectively. The automatic identification of a seasonal ARIMA model is another good indication of seasonality in the time series. Regarding the second category, the parametric F-test and the related non-parametric Kruskal-Wallis-test can be used to detect stable seasonal pattern. Both tests perform ANOVA-like comparisons of the (detrended) linearised time series' variability between

[1] On the one hand, instability issues may arise in regARIMA model estimation for short and very short series, see Item 6.1 of Eurostat (2015). On the other hand, identification of a "correct" model with time-constant parameters may be problematic for long series. To exclude both cases, we restrict our analyses to series whose lengths range between five and 25 years. In addition, we assume that they do not exhibit structural breaks since this phenomenon is usually tackled by separate seasonal adjustments of the (shorter) subseries created by these breaks.

The choice between the X-11 and AMB approaches

Figure 1



Deutsche Bundesbank

and within different periods, and do not require the specification of an underlying time series model. Complementary, several tests on both deterministic and stochastic seasonality have been suggested under additional modelling assumptions, including, for example, time series regression models with seasonal dummies, seasonal unit root processes, and periodic time series models, see Sections 2, 3 and 6 of Ghysels and Osborn (2001) for an overview. Specific problems within these frameworks have been tackled by Busetti (2006), Busetti and Harvey (2003), Busetti and Taylor (2003) and Franses (1992), amongst others. It should be noted, however, that tests for the presence of deterministic seasonality are typically less informative as they cannot appropriately handle moving seasonality.

3.2 Step 2: conceptual considerations

Assuming the time series exhibits a fair amount of stable seasonality, the second step covers conceptual differences between the X-11 and AMB approaches. The basic idea is that X-11 has advantages whenever the ARIMA representation given by Equation 6, which is derived straightforwardly from the key assumptions of the AMB approach, is likely to be too restrictive to adequately account for all dynamics of the observed time series. This might be the case for various phenomena, but we only focus on those two of them here which, according to our experience, are most relevant to practitioners.[2]

The first phenomenon is seasonal heteroskedasticity which is a particular form of periodic behaviour. In theory, ARIMA processes with time-constant coefficients cannot model explicitly periodic movements and, thus, need proper augmentation to do so. For example, Osborn (1991) and Tiao and Grupe (1980) propose models with time-varying coefficients, whereas Bell (2004) introduces a seasonal heteroskedastic noisy component to the Airline model, see also Proietti (1998) for a more general discussion of possible model extensions. By construction, the AMB approach does not make use of any of these augmentations: consequently, seasonal heteroskedasticity has to be captured as well as possible by seasonal ARIMA models with time-constant coefficients. In contrast, in X-11, GARCH-type effects may be taken into account, at least up to some degree, within the built-in extreme value detection procedure which allows for consideration of period-specific variances of the irregular component.[3] However, if the importance of such effects is sufficiently low, application of either the X-11 or the AMB approach usually does not

[2] An example not discussed here is given by integer-valued time series, such as some business climate indices and sentiment indicators. From a theoretical point of view, real-valued ARIMA processes might not immediately qualify as a first choice of DGP model for those series, even though they still seem to perform quite well empirically.

[3] To check the period-specific variances for equality, the one-sided test suggested by Cochran (1941), two-sided extensions thereof, and model-based tests can be employed, see Trimbur and Bell (2012) for a recent discussion.

make a huge difference as far as seasonal heteroskedasticity is concerned, at least according to our experience. In addition, as mentioned above, it should be borne in mind that this phenomenon is related to sample size, as it is less likely to be statistically significant in short and moderate-length time series. Hence, the AMB approach may be more viable a priori for those lengths of series.

The second phenomenon emerges when major sources of (moving) seasonality in observed unadjusted figures are substantially different across periods. Usually, detection and proper treatment of such effects require a thorough knowledge of both economic circumstances and data collection methods. A prime example is given by German retail sales in stores (see Box 1). As for seasonal heteroskedasticity, the X-11 approach can deal with this phenomenon to a certain extent, as it allows for application of period-specific seasonal filters. In contrast, the AMB approach is intentionally designed not to model such scenarios explicitly.

Box 1: The decreasing December peak of German retail sales in stores.

In Germany, and probably many other countries, retail sales in stores exhibit a significant peak in December which is commonly attributed to the Christmas business. In the recent past, however, the usual December peak has been decreasing steadily. Moreover, the speed of the decrease has been faster than gradual (if any) changes of seasonality in all other months. There are three widely accepted explanations: first, the number of firms that ceased their Christmas bonus payments has increased constantly, and those firms which continued paying Christmas bonuses have cut them steadily on average compared to regular salaries; second, the market share of retail sales via mail order houses or via Internet has soared despite some recent coverage issues; third, cash has become an increasingly popular Christmas present. Overall, unadjusted retail sales in stores for December should be treated somewhat individually by, for example, application of X-11 seasonal filters that are shorter compared to all other months as these are unlikely to be affected significantly by this particular December effect.

3.3 Step 3: empirical considerations

In practice, many time series behave rather less exceptionally and, hence, the phenomena described in Step 2 do not usually suffice to select an appropriate seasonal adjustment approach. For that reason, the decision tree analyses further criteria, which are rather empirical, in a third step.

We first suggest looking at squared gains of selected final X-11 and WK seasonal adjustment filters. The reason is that central and concurrent filters are of paramount importance in virtually all practical applications. Their quality should thus be assessed carefully as a matter of routine, and squared gains provide an easy yet effective tool to judge the overall performance of any linear filter. Regarding seasonal adjustment, the key idea is that any "acceptable" seasonal adjustment filter should completely eliminate seasonal fluctuations without significantly altering movements associated with non-seasonal frequencies. Assuming deterministic seasonality, the squared gain of an ideal seasonal adjustment filter would thus be equal to one at non-seasonal frequencies and discontinuously drop to zero at seasonal frequencies. However, as mentioned earlier, this model of how a seasonal component should evolve is often too restrictive in practice and should thus be relaxed by assuming stable stochastic seasonality, which allows the seasonal pattern to change gradually over time. The minimum requirement any "acceptable" seasonal adjustment filter should meet under this relaxed assumption is that its squared gain stays close to one over the range of non-seasonal frequencies and shows dips at all seasonal frequencies.[4] The widths of these dips can then be seen as an indication of how stable (ie "close-to-deterministic") the estimated seasonal component actually is. To mitigate over- and under-adjustment issues, they should match, at least approximately, the widths of the seasonal peaks in the spectrum of the (transformed and/or differenced) unadjusted figures, see the discussions given in Section 8.2 of Gómez and Maravall (2001) and Section 3 of Maravall (1995), including some related criticism of the X-11 approach.

By construction, final X-11 seasonal adjustment filters satisfy this minimum requirement, see Bell and Monsell (1992). The only exception is the 3×1 seasonal moving average which is likely to amplify significantly some intra-seasonal frequencies, regardless of the trend filter. This issue, however, is less severe given the rare use of this seasonal filter in practice. More importantly, some attention should be paid to the unwanted amplification and/or suppression of (mostly intra-year) non-seasonal frequencies introduced unavoidably by the finite length of real-world data. Since this issue is typical of finite filters, it should be checked not only for the X-11 approach, but also for the AMB approach. In general, final WK seasonal adjustment filters do not fulfil the minimum requirement by construction. Accordingly, their squared gains should be inspected more carefully for a "strange" curvature, which, if present, does not immediately indicate inferiority of the AMB method. To illustrate, transitory effects falsely assigned to the seasonal component are likely to lead to squared gains that are too low for a non-ignorable range of non-seasonal frequencies, which may happen especially when SEATS is run with default

[4] Alternative benchmarks may be established if either further assumptions are made about the DGP and its components, as in structural time series models, see Harvey (1989), or additional restrictions are imposed on the spectral density of the signal's estimator, as in digital signal processing, see the appendix of Findley and Martin (2006).

options. Noticeable deviations from the ideal squared gain should thus be interpreted rather as an invitation to thoroughly study the outcome of the SEATS decomposition algorithm and to make appropriate modifications, if possible.

If both approaches perform equally well after this spectral analysis, we suggest comparing them with respect to the revisions they cause in the seasonally adjusted figures and the period-on-period changes. The reason is that, in general terms, revisions reflect the price to be paid for up-to-date figures, and, thus, are of interest for virtually all users. This revision analysis, however, is not intended to be done in real-time. We do not aim to take into account all possible sources of revisions, see Deutsche Bundesbank (2011) for an overview thereof, but focus only on those due to the process of seasonal adjustment. A prime example is the incorporation of new observations as this has a direct impact on which (asymmetric) seasonal adjustment filters can be applied to the unadjusted figures preceding the new data.[5] Eventually, the seasonal adjustment approach which yields lower average revisions is chosen, which according to our experience should be in line with demands of users. In this regard, we also recommend weighting the revisions in terms of importance for users. For example, more weight could be put a priori on revisions of period-on-period changes since statements on economic developments typically focus on them rather than the seasonally adjusted figures.

3.4 Caveats

We close this section with four caveats. First, both approaches offer far more possibilities for conducting inference than considered here. This is especially true for the AMB method whose parametric framework, among many other things, allows for derivation of ARIMA models for both UC estimators and revisions of these estimators. Their theoretical properties, such as standard errors, auto- and cross-covariances, can thus be compared with the empirical properties of UC estimates and revisions, respectively, to check for model inadequacies. Within the non-parametric X-11 framework, such comparisons cannot be achieved since the theoretical properties of UC estimators cannot be derived in an analogous way. Nevertheless, various variance measures have been constructed to quantify different sources of error in X-11 seasonal adjustments, see Pfeffermann (1994) and Scott, Pfeffermann, and Sverchkov (2012).

Second, even though our decision tree is largely based on differences between the philosophies underlying the X-11 and AMB approaches, it should be noted that the former has a model-based interpretation in the sense that we can find an ARIMA signal extraction

[5] This revision analysis is fundamentally different from studying "optimal" (instead of "minimal") revisions that are obtainable in a model-based framework, see the discussion in Section 8 of Maravall and Pérez (2012).

whose final WK seasonal adjustment filter closely mimics the respective X-11 filter. Early contributions, for example Burridge and Wallis (1984) and Cleveland and Tiao (1976), demonstrate this for the X-11 default filter but later studies show that this statement holds in a more general sense for all combinations of X-11 trend and seasonal filters. In most cases, however, ARIMA signal extraction is based on decomposing Airline models, see Depoutot and Planas (1998) and Planas and Depoutot (2002). Hence, the AMB approach provides more flexibility in terms of an infinite number of potential final seasonal adjustment filters.

Third, depending on the methodology finally chosen by the decision tree, an "approach-specific" quality assessment should be carried out to further improve the estimation of the seasonally adjusted figures.

Fourth, the decision tree is not meant to be definitive but rather a general guide, and its character is still prototypical. While the current version will work most of the time, there will always be exceptions which may encourage further improvements, bearing in mind the impossibility of covering all possible situations.

4 Illustration

To prove the concept of the decision tree, we run X-13ARIMA-SEATS using its Windows companion Win X-13 (version 1.0 build 150)[6] with the following four German macroeconomic time series: turnover of industry originating from non-euro-area countries (TO), output of main construction industry (OUT), orders received from abroad by producers of non-durable consumer goods (OR) and gross domestic product (GDP). Table 1 provides basic information on these series, which have already been adjusted for calendar effects using regARIMA models. The underlying calculations are described in Deutsche Bundesbank (2012), while further details on the theory of regARIMA models can be found in Findley et al. (1998) and Section 4 of Ghysels and Osborn (2001).

For each series, the specification file closely mimics the setting used within the production process of official seasonally adjusted figures in Germany.[7] More specifically, for regARIMA modelling we adopt the UC decomposition, user-defined outliers (for example, level shift sequences to account for the economic downturn starting in 2008) and critical values for automatic outlier detection. ARIMA models are selected according to the

[6] During the finalisation of this paper, version 1.1 of Win X-13 was released. It contains a slightly different SEATS core and, thus, a replication of our analyses with this updated version could lead to results that differ marginally from the outcome reported here.

[7] Note, however, that macroeconomic aggregates as well as their major components are usually seasonally adjusted according to an indirect approach in official statistics. In this case, seasonally adjusted figures obtained from the direct approach are considered for comparative analyses.

Table 1: Selected metadata on turnover of industry (TO), output of main construction industry (OUT), orders received from abroad (OR) and gross domestic product (GDP).

Series	Frequency	Span	Reference year	Unit
TO	Monthly	01.2003 – 12.2013	2010 = 100	At current prices
OUT	Monthly	01.1991 – 12.2013	2010 = 100	At constant prices
OR	Monthly	01.1991 – 12.2013	2010 = 100	At current prices
GDP	Quarterly	01.1991 – 04.2013	2005 = 100	At previous-year prices

automatic model identification routines. Regarding X-11, we adopt trend and (period-specific) seasonal filters as well as the way extreme values in the irregular component are down-weighted.[8] In contrast, SEATS is first run under its default setting. We thus take the position of a practitioner who already has a thorough knowledge of X-11 (plus regARIMA modelling) but is less experienced in AMB seasonal adjustment.

To check the series for presence of stable seasonality, we consider the F-tests for the linearised and detrended linearised series (hereafter F_l and F_{dl}, respectively), the Kruskal-Wallis-test for the detrended linearised series (hereafter KW_{dl}) and the automatic ARIMA model identification. In addition, we consider the periodogram for the three monthly series.

Regarding seasonal heteroskedasticity, we use the (one-sided) Cochran-test for equality of period-specific variances (hereafter C), which is applied to the preliminary irregular component (as given by output table B 13).

To calculate the squared gains of the final X-11 seasonal adjustment filters, which are not provided by any Win X-13 output table, we use R (version 3.0.1) developed by the R Core Team (2013). The computations are based on Equation 3 using the weights stored in the matrix $\mathcal{W}$ of Equation 2, which is derived according to the impulse response method described in Section 3.4 of Ladiray and Quenneville (2001).

To analyse revisions, if necessary, we seasonally adjust truncated versions of the series under review and, hence, exclude corrections of unadjusted figures from our study. During this procedure, the regARIMA model is re-estimated for each truncated version. For two reasons, early observations are omitted to establish a "burn-in" period. First and foremost, a minimum of 60 observations before the starting date of the revision analysis is required by X-13ARIMA-SEATS for regARIMA model re-estimation. Second, depending on the lengths of the seasonal filters applied to a particular series, a further expansion

[8] If the decision tree is applied to a time series that should be seasonally adjusted for the first time, such adoptions are of course not possible. One may then start with default values (for automatic outlier detection) and automatic built-in procedures. The latter may be particulary helpful for the selection of trend and seasonal filters, see Webel (2013a,b). Alternatively, both types of filters can be chosen according to any criterion that defines "best" trend and seasonal filters, see Chu, Tiao, and Bell (2012) who derive a MSE-based criterion.

of the "burn-in" period is recommended to mitigate those effects at the beginning of the period that are caused by the application of asymmetric filters, which may lead to atypical revisions during the early span of the period actually analysed. The revision measures considered then are the mean revision, mean absolute revision and standard deviation of the revisions. Also, as mentioned exemplarily at the end of Section 3.3, we put more weight on the revisions of period-on-period changes.

For each series, the entire decision-making process is finally summarised in a copy of Figure 1 where the respective realised path through the decision tree is highlighted with dark grey arrows.

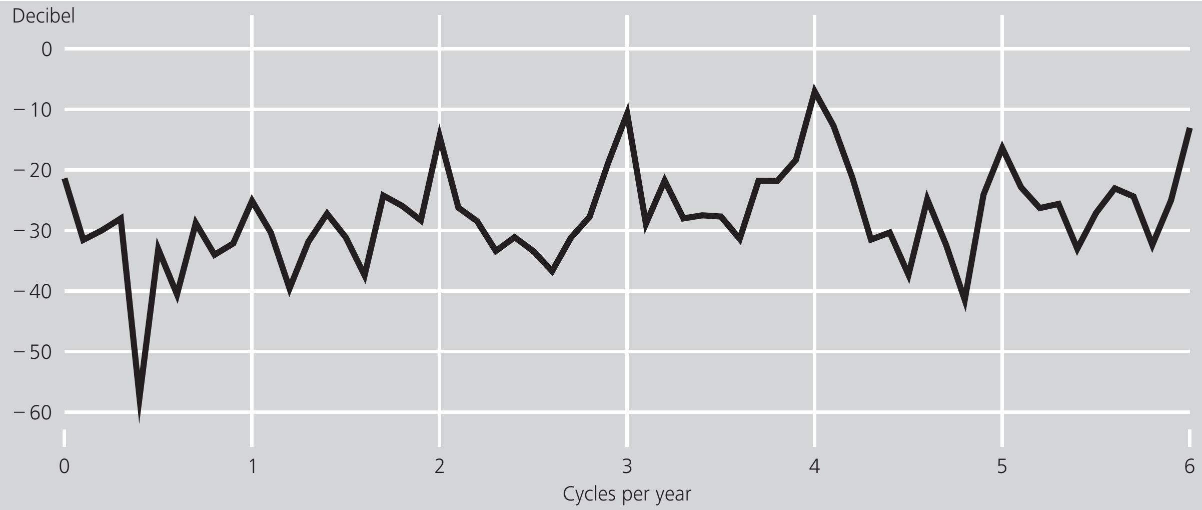
4.1 Turnover of industry

We first inspect the TO series for presence of stable seasonality. The three respective test statistics are given by $F_l = 85.436$, $F_{dl} = 134.528$ and $KW_{dl} = 121.197$. Thus, they are larger than respective critical values at any conventional level of significance, indicating presence of stable seasonality. A similar conclusion can be drawn from spectral analysis. Figure 2 shows the periodogram of the first differences of the logged linearised TO series, which is estimated from January 2006. It exhibits visible peaks at all seasonal frequencies, although the peak at the first seasonal frequency is somewhat less pronounced. Finally, the automatic ARIMA model selection procedure identifies a seasonal model for the linearised TO series, which is the Airline model given (to two decimal places) by

Periodogram of turnover of industry

Figure 2

Differenced logged linearised series



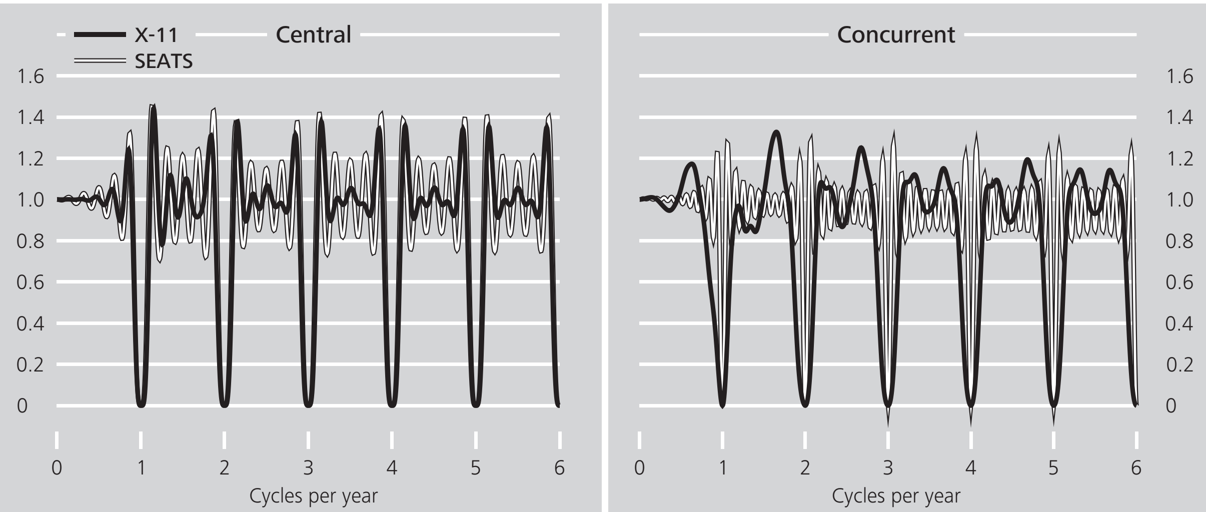
Deutsche Bundesbank

$$\nabla\nabla_{12}(\{\log x_t\}) = (1 - \underset{(0.08)}{0.33}B)\left(1 - \underset{(0.06)}{0.88}B^{12}\right)(\{\varepsilon_t\}),$$

where the figures reported in parentheses underneath the parameter estimates are the respective standard errors. Note that according to the model equation a multiplicative decomposition is used. Overall, the TO series can be assumed to contain a fair amount of stable seasonality and, therefore, we advance to the conceptual considerations.

Inspecting the TO series for seasonal heteroskedasticity first, we observe that $C = 0.22$, and the critical value at a 5% level of significance, which depends on the frequency and the length of the series, is given by 0.20. Thus, presence of seasonal heteroskedasticity is evident, albeit weakly. To underline the fragility of this result, we recalculated the Cochran-test with the last year being omitted. The C-statistic and critical value were then 0.18 and 0.21, respectively, indicating absence of seasonal heteroskedasticity. Since the presence/absence of seasonal heteroskedasticity plays a crucial role in the automatic X-11 extreme value detection and down-weighting routine and, thus, in the estimation of the seasonally adjusted figures, we calculated the mean absolute difference (MAD) between those months' values of the seasonally adjusted TO series which this procedure treats differently under consideration and ignorance of seasonal heteroskedasticity. Differences occur in 11 out of 132 months and the MAD is given by 0.33 index points. Thus, the importance of seasonal heteroskedasticity for the X-11 extreme value detection routine is rather low, and we should not rule the AMB approach on these grounds.

Squared gains of seasonal adjustment filters for turnover of industry Figure 3



Deutsche Bundesbank

Regarding presence of month-specific causes of seasonality, our thorough knowledge about the DGP of the TO series, including detailed information on data collection, does not contain any sufficient hint of the need to apply different month-specific seasonal filters. Hence, we still cannot decide on an appropriate seasonal adjustment approach and, accordingly, proceed to the empirical considerations.

We first compare the squared gains of the central and concurrent seasonal adjustment filters obtained from the X-11 and AMB approaches, see Figure 3. To choose between the two approaches, it is necessary to introduce further preferences where both approaches meet the minimum requirements outlined in Section 3.3. To exemplify, we could favour the approach that yields the more stable estimated seasonal component.

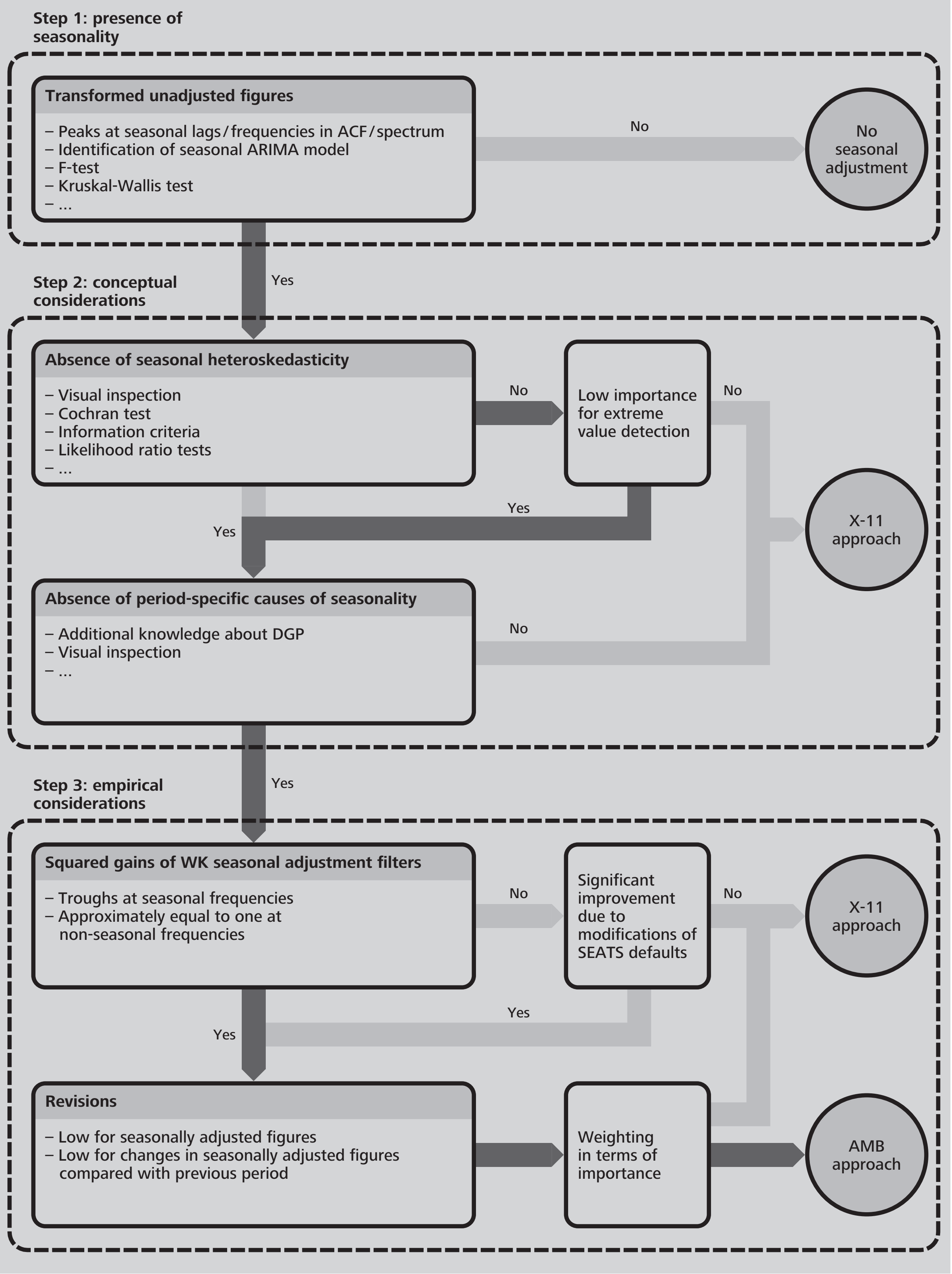
In X-11, the 17-term Henderson filter and 3×9 seasonal moving averages are used. The squared gains of the central and concurrent X-11 seasonal adjustment filters stay close to one at virtually all non-seasonal frequencies and decrease rapidly to zero in the vicinity of the seasonal frequencies. However, the squared gain of the central X-11 filter displays oscillations between the seasonal frequencies and thus signals amplification or dampening of respective intra-year fluctuations, which is most apparent at "close-to-seasonal" frequencies. This issue is also reflected in the squared gain of the concurrent X-11 filter which tends to stay above one over a broad range of non-seasonal frequencies. The WK filters exhibit squared gains whose curvatures are more or less similar to those of X-11. The peak amplification of "close-to-seasonal" frequencies is basically the same for both central filters, and the squared gains differ only in the amplitudes of their oscillations at non-seasonal frequencies, which are larger for the WK filter regardless of the frequency range considered. The squared gain of the concurrent WK filter stays closer to one over the range of non-seasonal frequencies than the central WK filter, despite the fact that its oscillatory behaviour is far more pronounced there compared to the squared gain of the concurrent X-11 filter, and the dips at seasonal frequencies are noticeably narrower. In sum, the advantages and disadvantages of both approaches basically cancel, so that they seem to perform equally well for the TO series in terms of the spectral analysis.

Table 2: Mean revision (MR), mean absolute revision (MAR) and standard deviation of the revisions (STD) for turnover of industry.

	Seasonally adjusted figures (as a percentage)			Changes in seasonally adjusted figures compared with previous month (in percentage points)		
Core	MR	MAR	STD	MR	MAR	STD
X-11	0.1383	0.5600	0.7063	0.0800	0.6248	0.8306
SEATS	0.1266	0.4847	0.6464	0.0537	0.5821	0.8243

The choice between the X-11 and AMB approaches for turnover of industry

Figure 4



Deutsche Bundesbank

Therefore, we now compare the two approaches with respect to the revisions they generate in the seasonally adjusted TO series and its month-on-month changes. The revisions are computed from January 2009 to November 2013, and Table 2 reports the three revision measures considered. Since the revisions of both the seasonally adjusted TO series and the month-on-month changes are consistently lower for AMB, this approach is preferred in line with the a priori weighting. Finally, Figure 4 shows the realised path through the decision tree for the TO series.

4.2 Output of main construction industry

Checking the OUT series for stable seasonality first, we observe that, as in the previous example, the three measures considered consistently indicate presence of stable seasonality. More precisely, $F_l = 386.861$, $F_{dl} = 483.862$ and $KW_{dl} = 248.943$. Also, the periodogram of the differenced linearised series, which is estimated from January 2000, exhibits visible peaks at all seasonal frequencies, see Figure 5. Finally, the automatic ARIMA model identification yields the following Airline model:

$$\nabla\nabla_{12}(\{x_t\}) = (1 - \underset{(0.05)}{0.56}B)\left(1 - \underset{(0.04)}{0.78}B^{12}\right)(\{\varepsilon_t\}),$$

indicating an additive decomposition. Overall, the OUT series should be seasonally adjusted, and we advance to the conceptual considerations.

Periodogram of output of main construction industry

Figure 5

Differenced linearised series

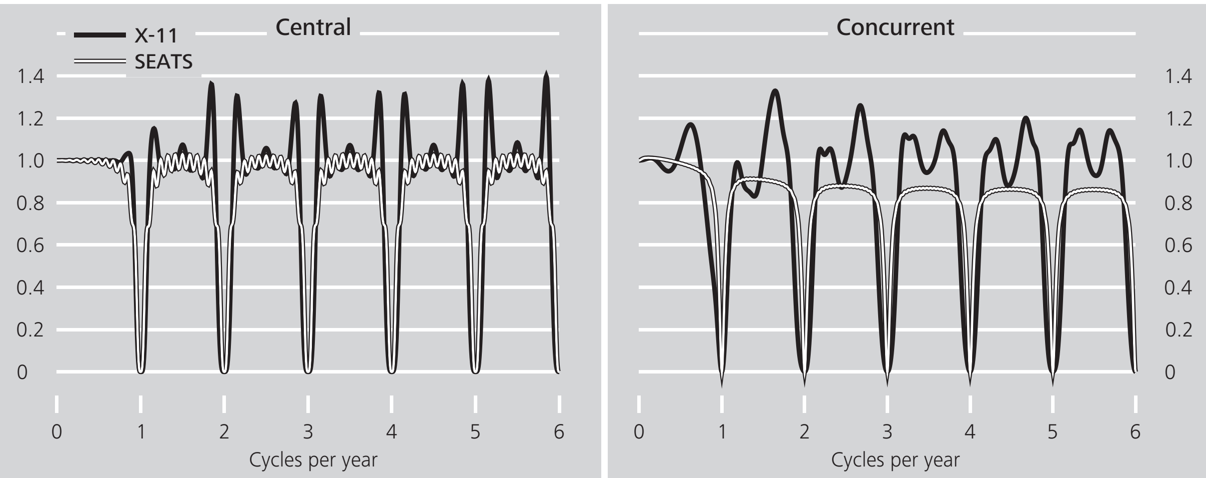
Deutsche Bundesbank

The Cochran-test shows strong evidence of seasonal heteroskedasticity as $C = 0.25$, which is significantly larger than the critical value at a 5% level of significance, given by 0.15. To assess the importance of this issue, as for TO, we calculated the MAD between the seasonally adjusted OUT series for those months which the X-11 extreme value detection routine treats differently under consideration and ignorance of seasonal heteroskedasticity. This occurs in 32 out of 276 months and the MAD is given by 1.10 index points. The seasonal heteroskedasticity issue has thus a strong influence, and it is much more severe compared to the TO series. Therefore, the X-11 approach appears preferable. For illustration, however, we do not make a final decision yet and keep the OUT series under study instead.

Since we do not have any information on month-specific causes of seasonality, we directly proceed to the empirical considerations and compare the squared gains of the final seasonal adjustment filters, see Figure 6. Regarding X-11, the same Henderson filter and, with rare exceptions, basically the same month-specific seasonal moving averages are applied as to the TO series. Accordingly, the squared gains of both central and concurrent X-11 filters barely differ from this series, and a similar comment applies. Regarding AMB, this approach also yields acceptable results, as for the TO series. Despite some oscillatory behaviour, the central WK filter's squared gain stays below but close to one at virtually all non-seasonal frequencies. Thus, amplification, especially at "close-to-seasonal" frequencies, is not an issue for this filter. Moreover, the troughs at seasonal

Squared gains of seasonal adjustment filters for output of main construction industry

Figure 6



Deutsche Bundesbank

Table 3: Mean revision (MR), mean absolute revision (MAR) and standard deviation of the revisions (STD) for output of main construction industry.

Core	Seasonally adjusted figures (as a percentage) MR	MAR	STD	Changes in seasonally adjusted figures compared with previous month (in percentage points) MR	MAR	STD
X-11	−0.0015	1.0866	1.4713	−0.0141	1.0534	1.4852
SEATS	0.0510	1.0530	1.4540	0.0424	1.1224	1.6540

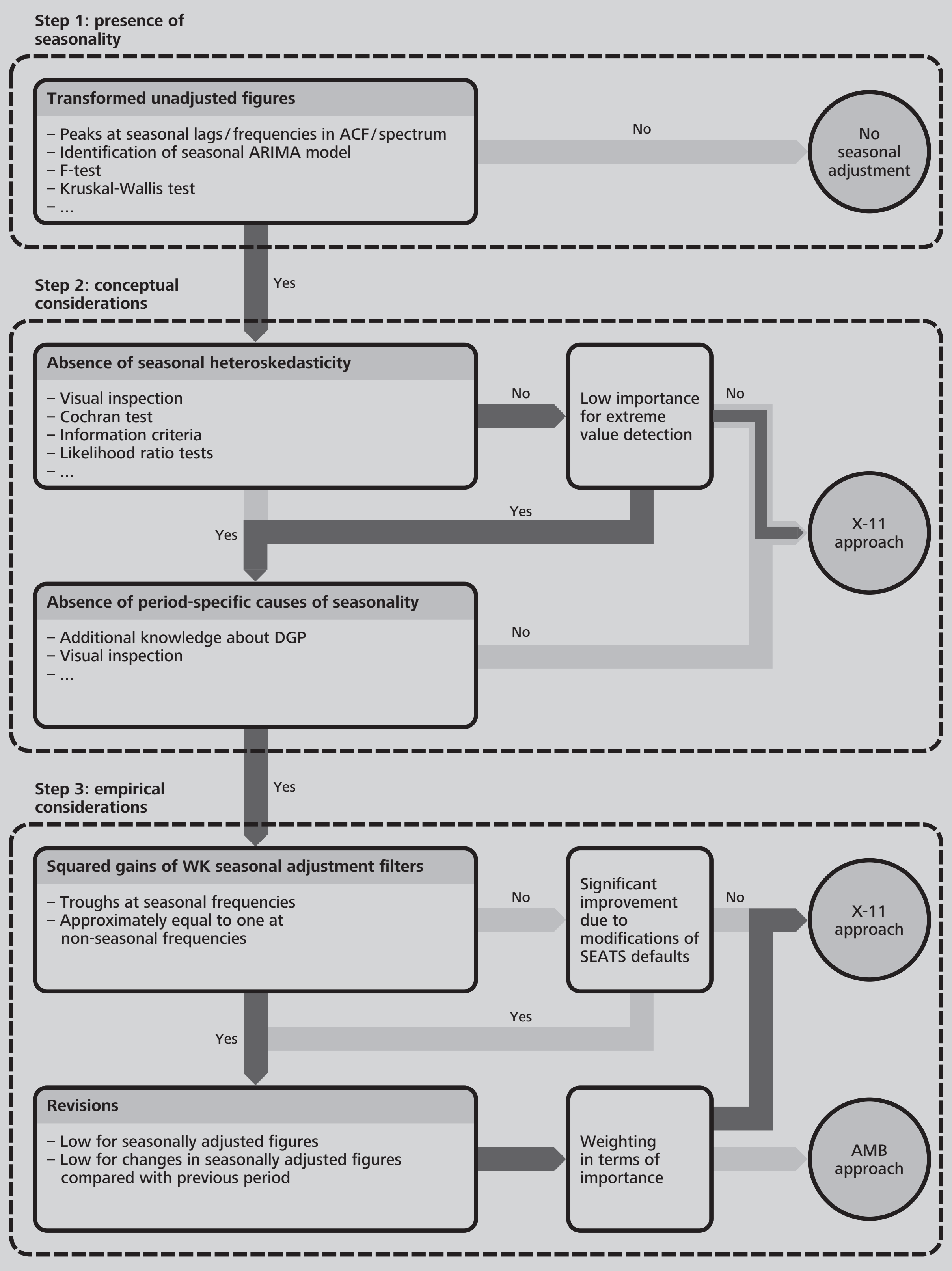
frequencies are even slightly narrower compared to the squared gain of the central X-11 filter. The latter statement also holds for the concurrent WK filter which, however, reveals a minor shortcoming since its squared gain stays between 0.8 and roughly 0.9 at all non-seasonal intra-year frequencies. Associated periodic movements are thus suppressed in a general way. Nevertheless, the AMB approach appears to gain a small (but still indecisive) advantage from this spectral inspection, especially when more weight is put on central seasonal adjustment filters.

Therefore, we finally analyse the revisions, which are computed from January 2000. Table 3 shows that both the seasonally adjusted figures and their month-on-month changes are revised downwards by X-11 and upwards by AMB. Leaving aside mean revisions, the other two measures are marginally lower for AMB for the seasonally adjusted OUT series and significantly smaller for X-11 for the month-on-month changes. According to the a priori weighting, X-11 is slightly preferable. Bearing in mind that the seasonal heteroskedasticity issue clearly favoured X-11, and the spectral inspection was slightly advantageous for AMB, we finally recommend X-11 for seasonal adjustment of the OUT series. The decided path through the tree, which is shown in Figure 7, is thus the same as for the TO series, but the final decision is different. However, note that under a stricter interpretation of the tree's criteria a decision could be made immediately after the "importance of seasonal heteroskedasticity" knot, which was the same as before. This is symbolised in Figure 7 by the thin dark grey arrow pointing from this knot to the X-11 approach.

4.3 Orders received from abroad

The three test statistics considered to check for presence of stable seasonality amount to $F_l = 50.144$, $F_{dl} = 52.876$ and $KW_{dl} = 200.163$, respectively. Again, they are larger than respective critical values at any conventional level of significance, but the evidence for stable seasonality is somewhat weaker compared to the previous two examples. This result is in line with the periodogram of the first differences of the logged linearised OR

The choice between the X-11 and AMB approaches for output of main construction industry — Figure 7



Deutsche Bundesbank

Periodogram of orders received from abroad

Differenced logged linearised series

Figure 8

Deutsche Bundesbank

series, which is calculated from January 2000, as it exhibits visible peaks only at three out of the six seasonal frequencies, see Figure 8. Still, a seasonal ARIMA model is automatically identified for the OR series:

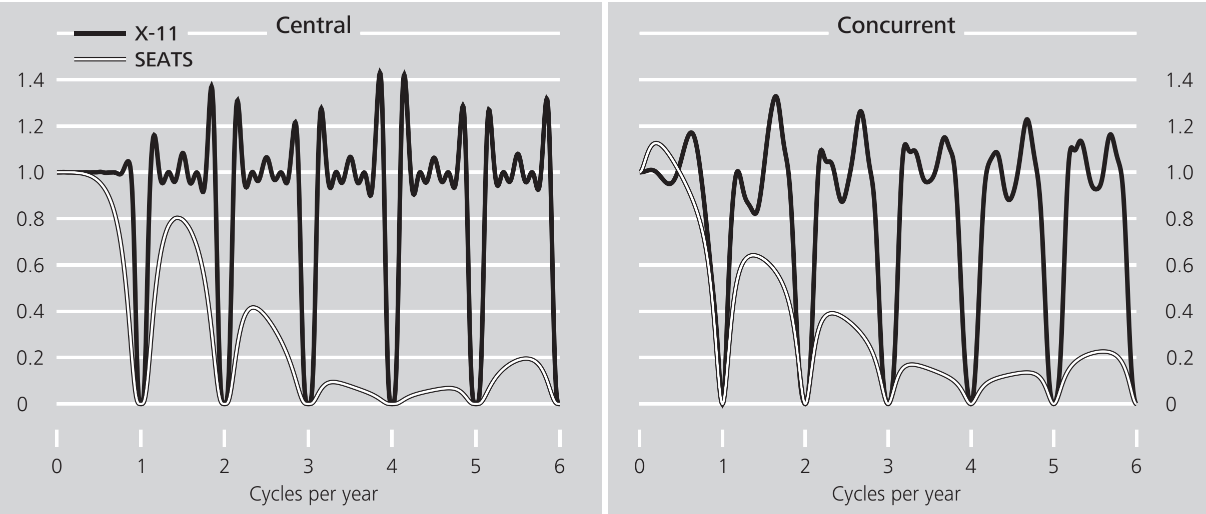
$$\left(1+\underset{(0.18)}{0.11}B+\underset{(0.15)}{0.05}B^2-\underset{(0.11)}{0.19}B^3\right)(\{\tilde{x}_t\})=(1-\underset{(0.16)}{0.71}B)\left(1-\underset{(0.05)}{0.47}B^{12}\right)(\{\varepsilon_t\}), \quad (9)$$

where $\{\tilde{x}_t\} = \nabla\nabla_{12}(\{\log x_t\})$. According to the model equation, a multiplicative decomposition is used. Overall, the series can be assumed to contain a fair amount of stable seasonality, and we advance to the conceptual considerations.

Checking the series for seasonal heteroskedasticity first, we observe that $C = 0.15$, which coincides with the critical value at a 5% level of significance. Thus, the Cochran-test does not provide evidence of seasonal heteroskedasticity. Also, we do not have sufficient hints of the need to apply different period-specific seasonal filters and, hence, immediately proceed to the spectral analysis of the empirical considerations.

Regarding X-11, the same Henderson filter and basically the same month-specific seasonal filters are applied as to the TO and OUT series. Thus, as shown in Figure 9, the squared gains of both the central and concurrent X-11 filters do not differ significantly from these series, and the same comment applies as for TO. Regarding AMB, the squared gain of the two WK filters exhibit an awkward curvature that signals noticeable suppression of non-seasonal movements whose periods are shorter than six months. More

Squared gains of seasonal adjustment filters for orders received from abroad Figure 9



Deutsche Bundesbank

specifically, they barely manage to exceed 0.2 beyond the third seasonal frequency and, thus, look rather like squared gains of poor trend extraction filters. Apparently, the decomposition of the ARIMA model fitted to the OR series does not work the way it should under default options. In particular, one may conjecture that transitory effects have been falsely assigned to the seasonal component. Hence, we should search for appropriate modifications to remedy this situation. For that purpose, we now examine the stationary AR polynomial of Equation 9 and notice that it factorises as

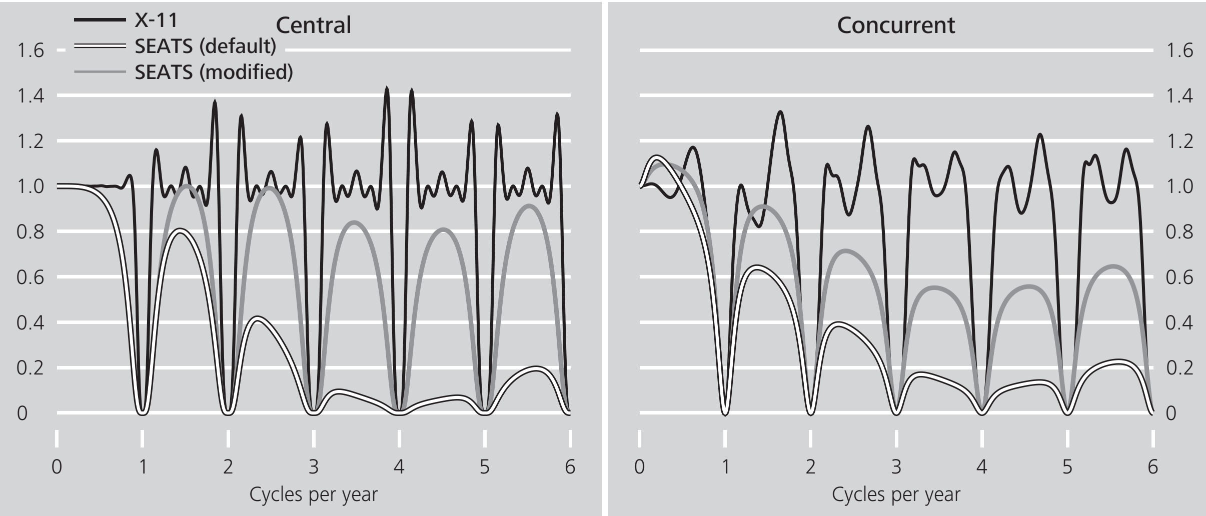
$$(1 - 0.51B)\left(1 + 0.62B + 0.37B^2\right), \tag{10}$$

posing two problems. First, although the real positive root of Equation 10 is correctly assigned to the trend-cyclical component since its (reciprocal) modulus is larger than the default threshold 0.5, it is likely to cancel with the non-seasonal MA polynomial of Equation 9 despite the fact that both roots do not share exactly the same modulus.[9] This potential cancellation issue is the reason why Equation 9 reports relatively high estimated standard errors for all non-seasonal parameter estimates. Second, the complex root of Equation 10 is in fact assigned to the seasonal component since its associated frequency is approximately $0.67\,\pi$ which, under default settings, is sufficiently close to the

[9] Apart from that, one may also argue that due to its small (reciprocal) modulus the real positive root of Equation 10 generates movements that are not persistent enough to be considered trend-cyclical behaviour, and should be assigned to the transitory component instead.

Squared gains of seasonal adjustment filters for orders received from abroad

Figure 10



Deutsche Bundesbank

fourth seasonal frequency, $2\pi/3$.

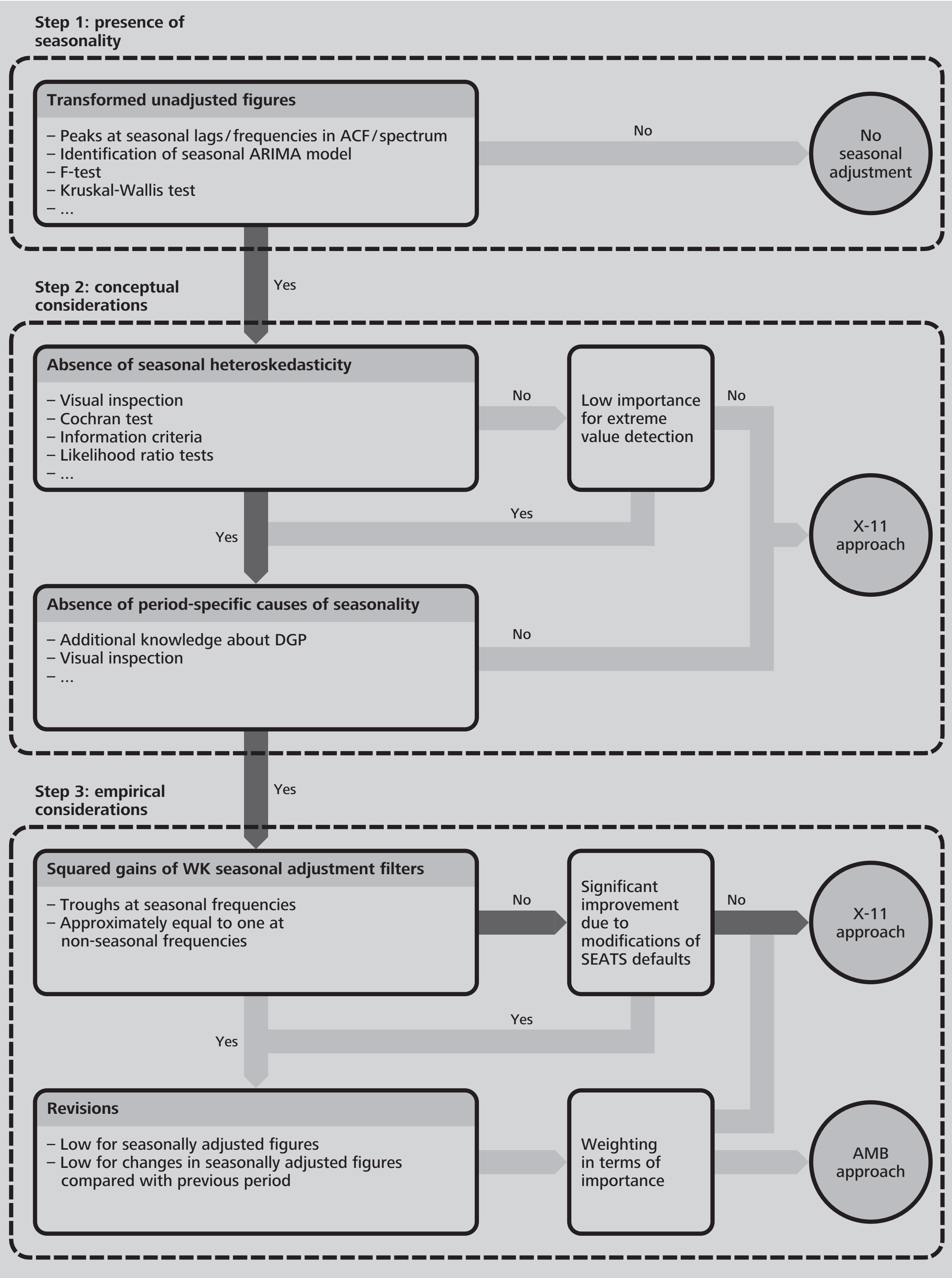
To tackle both problems, we drop the (almost) common AR and MA root from the model. Re-estimation of the reduced model of type $\log(210)(011)_{12}$ yields

$$\left(1 + \underset{(0.05)}{0.73}B + \underset{(0.05)}{0.46}B^2\right) \nabla\nabla_{12}(\{\log x_t\}) = \left(1 - \underset{(0.05)}{0.45}B^{12}\right)(\{\varepsilon_t\}). \tag{11}$$

As expected, all remaining parameter estimates are basically left unchanged compared to the $\log(311)(011)_{12}$ model. What is more, the standard errors of the two non-seasonal parameter estimates are now greatly reduced. However, the complex AR root of Equation 11 is still associated with the seasonal component. To enforce its allocation to the transitory component, we additionally reduce the range of “near-seasonal” frequencies to a sufficiently small interval. As a result, the seasonal component captures only “true” seasonal fluctuations as its assigned AR polynomial now coincides with the annual aggregation operator that is part of the seasonal differencing operator in Equation 11. As shown in Figure 10, the two WK filters perform visibly better under these modifications. Although retaining their overall curvature, both squared gains stay much closer to one at virtually all non-seasonal frequencies, especially for the central WK filter. Despite these substantial improvements, however, suppression of some non-seasonal intra-year fluctuations is still an issue for both filters since, for example, the squared gain of the concurrent filter drops below 0.7 beyond the second seasonal frequency. To sum up, the X-11 approach is prefe-

The choice between the X-11 and AMB approaches for orders received from abroad

Figure 11



Deutsche Bundesbank

rable for the OR series even after customisation of the AMB approach. Thus, we implicitly put less weight on the potential risk of amplifying a small set of "close-to-seasonal" frequencies compared to dampening a broader range of non-seasonal frequencies. As a consequence, there is no need to compare the two approaches with respect to revisions to arrive at a final decision, see Figure 11 for the path through the tree chosen for the OR series.

However, this example should not be misinterpreted as evidence that changing SEATS defaults is generally a hopeless task. For example, Webel (2013a) provides two cases of successful customisations of the AMB approach.

4.4 Gross domestic product

As for the three monthly series, the three tests provide compelling evidence of stable seasonality as their test statistics, which are given by $F_l = 165.610$, $F_{dl} = 239.236$ and $KW_{dl} = 82.483$, respectively, are larger than respective critical values at any conventional level of significance. Also, the automatic ARIMA model identification routine yields a seasonal model, which is given by

$$\nabla\nabla_4(\{\log x_t\}) = \left(1 - \underset{(0.09)}{0.46}B^4\right)(\{\varepsilon_t\}),$$

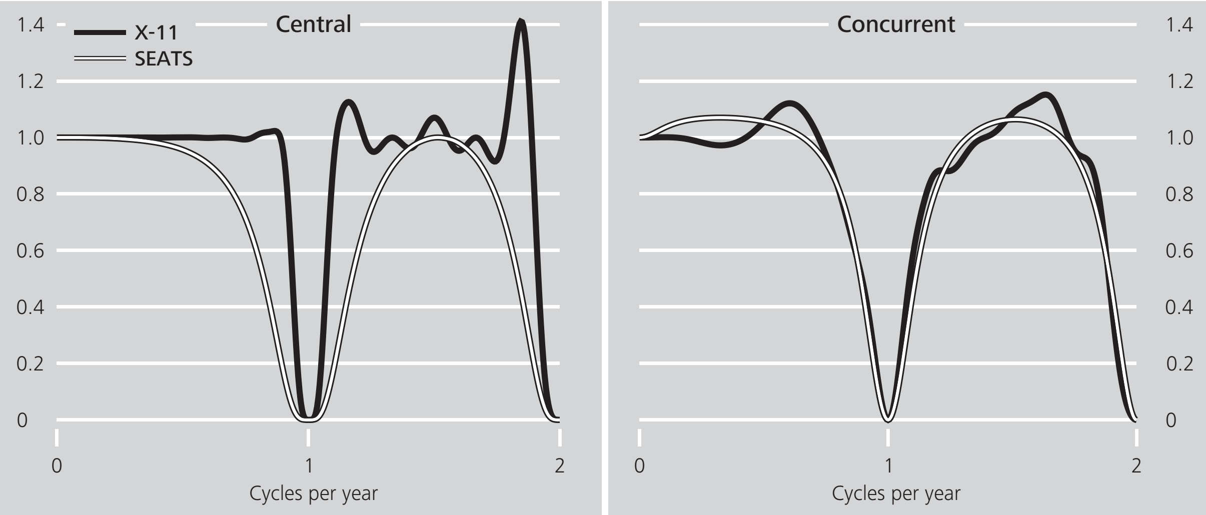
indicating a multiplicative decomposition. Overall, the quarterly GDP series is in need of seasonal adjustment, and we advance to the conceptual considerations.

We first observe that, as for the OR series, the Cochran-test does not show evidence of seasonal heteroskedasticity since $C = 0.42$, which is equal to the critical value at a 5% level of significance. Regarding quarter-specific causes of seasonality, Deutsche Bundesbank (2014) provides empirical evidence of how the GDP series has been affected recently by exceptional weather conditions. However, the official seasonally adjusted GDP series, whose calculation we attempt to mimic as closely as possible, is not adjusted for weather-induced effects, in accordance with Item 2.6 of Eurostat (2015). Overall, the application of quarter-specific seasonal filters is not justified clearly, and, therefore, we proceed to the spectral analysis of the empirical considerations.

The squared gains of the X-11 and AMB seasonal adjustment filters are shown in Figure 12. In general, it is immediately recognised that the squared gains of the WK filters are far smoother than those of the X-11 filters. However, both approaches have advantages and disadvantages. On the one hand, the squared gain of the central X-11 filter stays closer to one at most non-seasonal frequencies, despite some oscillatory behaviour at the intra-seasonal frequencies, and has narrower dips at both seasonal frequencies. On the other hand, a major weakness of the central X-11 filter is an amplification of periodic

Squared gains of seasonal adjustment filters for gross domestic product

Figure 12



Deutsche Bundesbank

movements whose cycle duration is slightly larger than two quarters, an issue not shared by the central WK filter. Finally, the squared gains of the concurrent filters do not exhibit any significant difference. Hence, both approaches perform reasonably well from this spectral point of view and a final choice between them cannot be made.

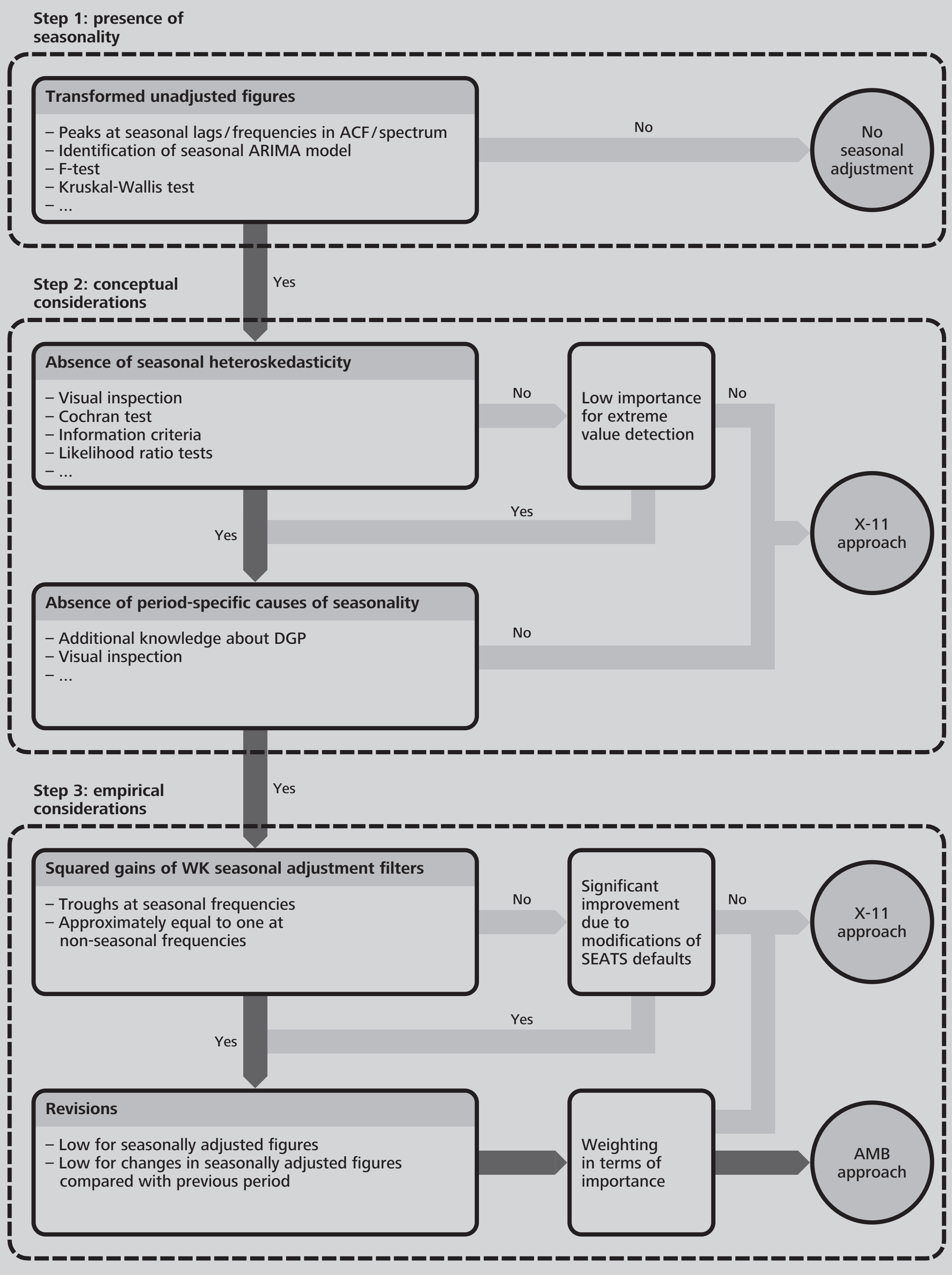
Therefore, we now compare both approaches with respect to the revisions they generate in the seasonally adjusted GDP series and its quarter-on-quarter changes. The revisions reported in Table 4 are calculated from Q1 2006 to Q3 2013. Regarding seasonally adjusted figures, they tend to be marginally lower as well as less volatile with X-11, whereas the converse is true for revisions of quarter-on-quarter changes. According to the a priori weighting, the AMB approach is thus preferred. Finally, Figure 13 shows the realised path through the decision tree for the GDP series.

Table 4: Mean revision (MR), mean absolute revision (MAR) and standard deviation of the revisions (STD) for gross domestic product.

	Seasonally adjusted figures (as a percentage)			Changes in seasonally adjusted figures compared with previous quarter (in percentage points)		
Core	MR	MAR	STD	MR	MAR	STD
X-11	−0.0075	0.1399	0.1747	−0.0065	0.2085	0.2697
SEATS	−0.0064	0.1434	0.1764	−0.0058	0.1826	0.2331

The choice between the X-11 and AMB approaches for gross domestic product

Figure 13



Deutsche Bundesbank

5 Summary

Recent releases of X-13ARIMA-SEATS and JDemetra+ provide easy access to both the X-11 and AMB approaches to seasonal adjustment. To assist users to choose the better approach for any given time series, we suggest a decision tree whose branches consider both conceptual differences and empirical issues. For the purpose of illustration, we use four macroeconomic time series that reflect distinct fields of the German economy. According to these examples, the X-11 approach tends to be recommended for longer time series, such as output of main construction industry and orders received from abroad, while the AMB approach is preferred for moderate-length and short series, such as turnover of industry and gross domestic product. This result may be explained by the fact that some phenomena which in theory cannot be modelled adequately by ARIMA processes, such as seasonal heteroskedasticity, require a sufficient amount of data to become visible in the observations.

However, the decision tree is still prototypical, and further criteria should be added to make it production-ready. Regarding the conceptual considerations, the length of the observed time series should be incorporated explicitly due to its being of key importance for the two seasonal heteroskedasticity knots, amongst other things. In this regard, future research should pay more attention to short and very short series. In addition, the two approaches could be compared with respect to the capacity to leave the seasonally adjusted series unchanged when the seasonal adjustment approach that has produced this series is applied to it again (idempotency), the capacity to avoid introduction of spurious seasonality to a non-seasonal series, or the capacity to deal with cointegrating relationships and/or survey errors, see Maravall (1998) and Tiller (2012) for discussions of some of these issues.

Regarding the empirical considerations, the spectral analysis should be enhanced. For example, comparisons between time shifts of final concurrent seasonal adjustment filters (with special emphasis on the range of cyclical frequencies) and periodograms of seasonally adjusted figures could complement the visual inspection of squared gains, which itself could be improved by using distance measures to assess the gains' closeness to the preferred benchmark. Also, further quality criteria for the seasonally adjusted series could be incorporated, including, for example, checks for absence of residual seasonality and stability of the estimated seasonal component.

Notwithstanding these possible extensions of Steps 2 and 3, the idea of resorting to a pragmatic solution may be developed further to become the fourth step of the decision tree, which relates rather to the generic culture of practitioners and the expectations of users. Respective branches could compare the two approaches with respect to compatibility with production systems, practitioners' expertise in time series theory, and user understanding.

References

Bell, W. R. (1984, June). Signal Extraction for Nonstationary Time Series. *The Annals of Statistics 12*(2), 646–664.

Bell, W. R. (2004). On RegComponent Time Series Models and Their Applications. In A. Harvey, S. J. Koopman, and N. Shephard (Eds.), *State Space and Unobserved Component Models: Theory and Applications*, pp. 248–283. Cambridge: Cambridge University Press.

Bell, W. R. and S. C. Hillmer (1984, October). Issues Involved With the Seasonal Adjustment of Economic Time Series. *Journal of Business & Economic Statistics 2*(4), 291–320.

Bell, W. R., S. H. Holan, and T. S. McElroy (Eds.) (2012). *Economic Time Series – Modeling and Seasonality*. Boca Raton: CRC Press.

Bell, W. R. and B. C. Monsell (1992). X-11 Symmetric Linear Filters and their Transfer Functions. Research Report No 1992-15, Statistical Research Division, U. S. Census Bureau, Washington, D. C.

Burman, J. P. (1980). Seasonal Adjustment by Signal Extraction. *Journal of the Royal Statistical Society A 143*(3), 321–337.

Burridge, P. and K. F. Wallis (1984, October). Unobserved-Components Models for Seasonal Adjustment Filters. *Journal of Business & Economic Statistics 2*(4), 350–359.

Busetti, F. (2006, May-June). Tests of Seasonal Integration and Cointegration in Multivariate Unobserved Component Models. *Journal of Applied Econometrics 21*(4), 419–438.

Busetti, F. and A. Harvey (2003, July). Seasonality Tests. *Journal of Business & Economic Statistics 21*(3), 420–436.

Busetti, F. and A. M. R. Taylor (2003, November). Testing against Stochastic Trend and Seasonality in the Presence of Unattended Breaks and Unit Roots. *Journal of Econometrics 117*(1), 21–53.

Caporello, G., A. Maravall, and F. J. Sánchez (2001). Program TSW Reference Manual. Working Paper No 0112, Servicio de Estudios, Banco de España, Madrid.

Chu, Y.-J., G. C. Tiao, and W. R. Bell (2012). A Mean Squared Error Criterion for Comparing X-12-ARIMA and Model-Based Seasonal Adjustment Filters. *Taiwan Economic Forecast and Policy 43*(1), 1–32.

Cleveland, W. P. and G. C. Tiao (1976, September). Decomposition of Seasonal Time Series: A Model for the Census X-11 Program. *Journal of the American Statistical Association 71*(355), 581–587.

Cochran, W. G. (1941). The Distribution of the Largest of a Set of Estimated Variances as a Fraction of Their Total. *Annals of Human Genetics 11*, 47–52.

Depoutot, R. and C. Planas (1998). Comparing Seasonal Adjustment and Trend Extraction Filters with Application to a Model-Based Selection of X11 Linear Filters. Eurostat Working Paper No 9/1998/A/9, Eurostat, Luxembourg.

Deutsche Bundesbank (2011, July). Reliability and Revision Profile of Selected German Economic Indicators. *Monthly Report 63*(7), 49–62.

Deutsche Bundesbank (2012, December). Calendar Effects on Economic Activity. *Monthly Report 64*(12), 51–60.

Deutsche Bundesbank (2014, May). The Impact of Weather Conditions on Gross Domestic Product in the Latter Part of 2013 and Early Part of 2014. *Monthly Report 66*(5), 54–55.

European Central Bank (2000). Seasonal Adjustment of Monetary Aggregates and HICP for the Euro Area. Report, ISBN 92-9181-86-X.

Eurostat (2015). ESS Guidelines on Seasonal Adjustment. Office for Official Publications of the European Commission, Luxembourg, ISBN 978-92-79-45176-8.

Findley, D. F. and D. E. K. Martin (2006, March). Frequency Domain Analyses of SEATS and X-11/12-ARIMA Seasonal Adjustment Filters for Short and Moderate-Length Time Series. *Journal of Official Statistics 22*(1), 1–34.

Findley, D. F., B. C. Monsell, W. R. Bell, M. C. Otto, and B.-C. Chen (1998, April). New Capabilities and Methods of the X-12-ARIMA Seasonal-Adjustment Program. *Journal of Business & Economic Statistics 16*(2), 127–152.

Franses, P. H. (1992, March). Testing for Seasonality. *Economics Letters 38*(3), 259–262.

Ghysels, E. and D. R. Osborn (2001). *The Econometric Analysis of Seasonal Time Series.* Cambridge: Cambridge University Press.

Gómez, V. and A. Maravall (1996). Programs TRAMO and SEATS – Instructions for the User. Working Paper No 9628, Servicio de Estudios, Banco de España, Madrid.

Gómez, V. and A. Maravall (2001). Seasonal Adjustment and Signal Extraction in Economic Time Series. In D. Peña, G. C. Tiao, and R. S. Tsay (Eds.), *A Course in Time Series Analysis*, pp. 202–247. New York: Wiley.

Harvey, A. C. (1989). *Forecasting, Structural Time Series Models and the Kalman Filter.* Cambridge: Cambridge University Press.

Hillmer, S. C. and G. C. Tiao (1982, March). An ARIMA Model-Based Approach to Seasonal Adjustment. *Journal of the American Statistical Association 77*(377), 63–70.

Hylleberg, S. (Ed.) (1992). *Modelling Seasonality.* Oxford: Oxford University Press.

Koopman, S. J., A. C. Harvey, J. A. Doornik, and N. Shephard (1996). *STAMP: Structural Time Series Analyser, Modeller and Predictor.* London: Chapman and Hall.

Ladiray, D. and B. Quenneville (2001). *Seasonal Adjustment with the X-11 Method*, Volume 158 of *Lecture Notes in Statistics.* New York: Springer.

Maravall, A. (1987, January). Minimum Mean Squared Error Estimation of the Noise in Unobserved Component Models. *Journal of Business & Economic Statistics 5*(1), 115–120.

Maravall, A. (1988, June). A Note on Minimum Mean Squared Error Estimation of Signals with Unit Roots. *Journal of Economic Dynamics & Control 12*(2-3), 589–593.

Maravall, A. (1989, January). On the Dynamic Structure of a Seasonal Component. *Journal of Economic Dynamics & Control 13*(1), 81–91.

Maravall, A. (1993, March). Stochastic Linear Trends – Models and Estimators. *Journal of Econometrics 56*(1-2), 5–37.

Maravall, A. (1995). Unobserved Components in Economic Time Series. In M. H. Pesaran and M. R. Wickens (Eds.), *The Handbook of Applied Econometrics*, Volume 1, pp. 12–72. Oxford: Blackwell Publishing.

Maravall, A. (1998, April). Comment on New Capabilities and Methods of the X-12-ARIMA Seasonal-Adjustment Program. *Journal of Business & Economic Statistics 16*(2), 155–160.

Maravall, A. and D. Pérez (2012). Applying and Interpreting Model-Based Seasonal Adjustment – The Euro-Area Industrial Production Series. In W. R. Bell, S. H. Holan, and T. S. McElroy (Eds.), *Economic Time Series – Modeling and Seasonality*, pp. 281–313. Boca Raton: CRC Press.

Osborn, D. R. (1991, June). The Implications of Periodically Varying Coefficients for Seasonal Time-Series Processes. *Journal of Econometrics 48*(3), 373–384.

Pfeffermann, D. (1994, January). A General Method for Estimating the Variances of X-11 Seasonally Adjusted Estimators. *Journal of Time Series Analysis 15*(1), 85–116.

Planas, C. and R. Depoutot (2002, March). Controlling Revisions in ARIMA-Model-Based Seasonal Adjustment. *Journal of Time Series Analysis 23*(2), 193–213.

Proietti, T. (1998, January). Seasonal Heteroscedasticity and Trends. *Journal of Forecasting 17*(1), 1–17.

R Core Team (2013). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. ISBN 3-900051-07-0.

Scott, S., D. Pfeffermann, and M. Sverchkov (2012). Estimating Variance in X-11 Seasonal Adjustment. In W. R. Bell, S. H. Holan, and T. S. McElroy (Eds.), *Economic Time Series – Modeling and Seasonality*, pp. 185–210. Boca Raton: CRC Press.

Scott, S., R. B. Tiller, and D. Chow (2007). Empirical Evaluation of X-11 and Model-based Seasonal Adjustment Methods. JSM Proceedings, Business and Economic Statistics Section, Alexandria, VA: American Statistical Association, pp. 967–978.

Shiskin, J., A. H. Young, and J. C. Musgrave (1967). The X-11 Variant of the Census Method II Seasonal Adjustment Program. Technical Paper No 15, U. S. Department of Commerce, Bureau of the Census, Washington, D. C.

Tiao, G. C. and M. R. Grupe (1980). Hidden Periodic Autoregressive-Moving Average Models in Time Series Data. *Biometrika 67*(2), 365–373.

Tiller, R. B. (2012). Frequency Domain Analysis of Seasonal Adjustment Filters Applied to Periodic Labor Force Survey Series. In W. R. Bell, S. H. Holan, and T. S. McElroy (Eds.), *Economic Time Series – Modeling and Seasonality*, pp. 135–158. Boca Raton: CRC Press.

Trimbur, T. M. and W. R. Bell (2012). Seasonal Heteroskedasticity in Time Series Data: Modeling, Estimation, and Testing. In W. R. Bell, S. H. Holan, and T. S. McElroy

(Eds.), *Economic Time Series – Modeling and Seasonality*, pp. 37–62. Boca Raton: CRC Press.

Webel, K. (2013a). Data-driven selection criteria for X-13ARIMA-SEATS seasonal adjustment algorithms. JSM Proceedings, Business and Economic Statistics Section, Alexandria, VA: American Statistical Association, pp. 1972–1986.

Webel, K. (2013b). Time series-dependent selection of an appropriate seasonal adjustment approach. Proceedings of the 59th ISI World Statistics Congress, Hong Kong, pp. 2962–2967.

Whittle, P. (1963). *Prediction and Regulation by Linear Least-Square Methods*. London: English University Press.