DANMARKS NATIONALBANK

DANMARKS NATIONALBANK WORKING PAPERS 2006 • 44

Mette Kramer Pedersen

Danmarks Nationalbank

and

Charlotte Franck Fæste

Danmarks Nationalbank

Seasonal adjustment of Danish financial time series using the X-12-ARIMA procedure

December 2006

The Working Papers of Danmarks Nationalbank describe research and development, often still ongoing, as a contribution to the professional debate.

The viewpoints and conclusions stated are the responsibility of the individual contributors, and do not necessarily reflect the views of Danmarks Nationalbank.

As a general rule, Working Papers are not translated, but are available in the original language used by the contributor.

Danmarks Nationalbank's Working Papers are published in PDF format at www.nationalbanken.dk. A free electronic subscription is also available at this Web site.

The subscriber receives an e-mail notification whenever a new Working Paper is published.

Please direct any enquiries to
Danmarks Nationalbank, Information Desk, Havnegade 5, DK-1093 Copenhagen K
Denmark
Tel.: +45 33 63 70 00 (direct) or +45 33 63 63 63
Fax : +45 33 63 71 03
E-mail:info@nationalbanken.dk

Nationalbankens Working Papers beskriver forsknings- og udviklingsarbejde, ofte af foreløbig karakter, med henblik på at bidrage til en faglig debat.

Synspunkter og konklusioner står for forfatternes regning og er derfor ikke nødvendigvis udtryk for Nationalbankens holdninger.

Working Papers vil som regel ikke blive oversat, men vil kun foreligge på det sprog, forfatterne har brugt.

Danmarks Nationalbanks Working Papers er tilgængelige på Internettet www.nationalbanken.dk i pdf-format. På webstedet er det muligt at oprette et gratis elektronisk abonnement, der leverer en e-mail notifikation ved enhver udgivelse af et Working Paper.

Henvendelser kan rettes til :
Danmarks Nationalbank, Informationssektionen, Havnegade 5, 1093 København K.
Telefon: 33 63 70 00 (direkte) eller 33 63 63 63
E-mail: info@nationalbanken.dk

ISSN (trykt/print) 1602-1185

ISSN (online) 1602-1193

Seasonal adjustment of Danish financial time series using the X-12-ARIMA procedure

Mette Kramer Pedersen

Danmarks Nationalbank

Havnegade 5

DK-1093 København K

Phone: +45 3363 6867

E-mail: mkp@nationalbanken.dk

Charlotte Franck Fæste

Danmarks Nationalbank

Havnegade 5

DK-1093 København K

Phone: +45 3363 6862

E-mail: cff@nationalbanken.dk

November 2006

Abstract

Danmarks Nationalbank has in November 2006 started to publish five seasonally adjusted financial time series. The series chosen for seasonal adjustment are monthly series for c*urrency in circulation*, the money stock measures *M1* and *M2* and the *MFI sector's lending to households* and to *non-financial corporations.* The purpose of this working paper is to present and document the applied adjustment procedures. The paper is structured as a detailed guide to theoretical considerations and to the practical implementation.

The models and the seasonally adjusted data will be subject to annual revisions in accordance with the general revision policy as defined by Statistics.

Resumé

Danmarks Nationalbank begyndte i november 2006 at offentliggøre fem sæsonkorrigerede finansielle tidsserier. De valgte serier for sæsonkorrektion er månedlige tidsserier for *seddel- og møntomløbet*, pengemængdemålene *M1* og *M2*, *MFI-sektorens udlån til husholdninger* og til *ikke-finansielle selskaber.* Formålet med dette arbejdspapir er at præsentere og dokumentere de anvendte procedurer i forbindelse med sæsonkorrektion. Papiret er struktureret som en detaljeret guide til de teoretiske overvejelser og den praktiske gennemførelse og implementering af sæsonkorrektion.

Modellerne og de sæsonkorrigerede data vil blive evalueret årligt i overensstemmelse med den generelle revisionspolitik i Statistisk Afdeling.

1. Introduction

The seasonally adjusted financial data now published is intended to supplement financial data already published for the MFI sector[1]. The seasonally adjusted data available are *currency in circulation*, the money stock measures *M1* and *M2*, and the *MFI sector's lending to households* and to *non-financial corporations.* The seasonally adjusted series are monthly series starting in 1995.

It has become standard in Statistics Denmark and in several other central banks to use an approach for seasonal adjustment called X-12-ARIMA. X-12-ARIMA is a software package which combines regression analysis, time series modelling of univariate time series and a moving average approach for seasonal adjustment. It is a package with many possibilities for modelling and testing seasonality. However, the generation of seasonally adjusted series should not be too time-consuming; therefore the process is partially based on automatic procedures built into the X-12-ARIMA package[2]. This working paper goes into a lot of detail in the application of X-12-ARIMA. The five financial series represent a number of challenges to seasonal adjustment: There are breaks in the data caused by changes in compilation method, all observations are end-of-month and there is a potential week-day effect as it may matter whether a month ends on a Monday or a Friday, there is a three year cycle in the lending to households due to a three year cycle in the issuance of mortgage bonds, etc. We shall address these potential problems in the context of the X-12-ARIMA package.

The structure of the working paper is as follows: In section 2 the concept of seasonality and the background for providing seasonally adjusted data is briefly addressed. The methodology and principles behind the X-12-ARIMA package is presented in section 3. In section 4 the choice between indirect and direct adjustment is explained and tested. A graphical inspection of the five unadjusted time series is provided in section 5. Prior to ARIMA modelling and seasonal adjustment, the data is subject to preadjustment where effects of the calendar (day-of-the-week effects) and

---

[1] The MFI sector is short for the monetary and financial institutions (in Denmark) including banks, mortgage-credit institutes and other credit institutes.

[2] X-12-ARIMA uses three different files in the production of seasonally adjusted data: an input file, a specification file and an output file. The input file contains the original data series in question; the specification file contains the programming choices used for generating the seasonally adjusted data including model selection and filter length choice; and the output file prints all information and data generated including diagnostics. See Ladiray, D. and B. Quenneville (1999) for more information on the output provided by X-12-ARIMA . Users can obtain the specification files from Statistics, Danmarks Nationalbank.

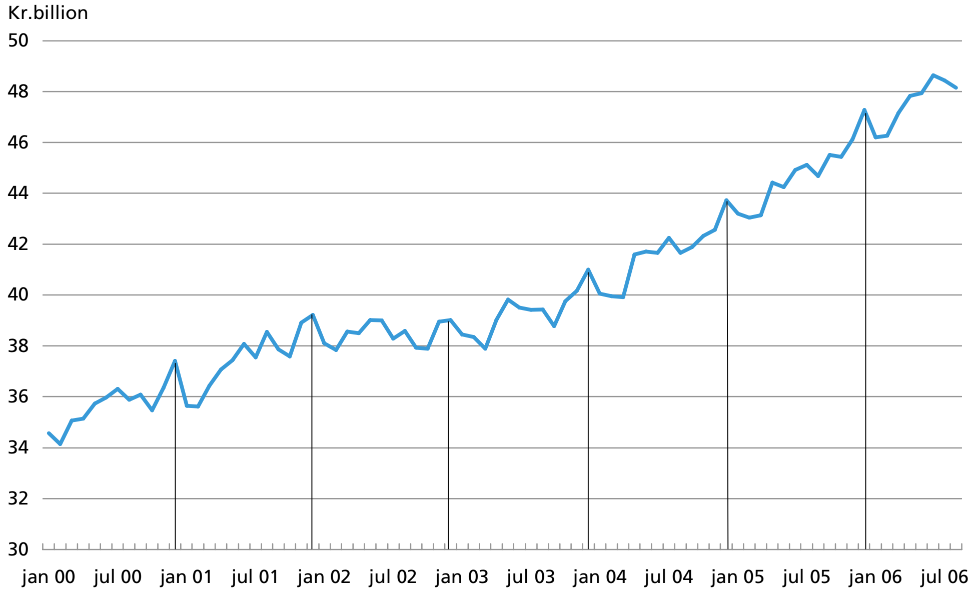
outliers are filtered out. This preadjustment is addressed in section 6. The evaluation of the final seasonal adjustment is discussed in section 7. Finally revision policy is discussed in section 8.

2. What is seasonality?

Seasonality is present in many economic time series and is defined as systematic movements that occur within the year at given time points each year. These effects can either be ascribed to the passing of seasons, e.g. harvest or to more specific calendar effects resulting from Easter, holidays, or specific days of the week. The mechanics of seasonal adjustment of financial time series does not differ from that of real economic time series. Financial time series are, however, to a higher degree than real economic time series affected by institutional changes. An example of institutional change could be changes in the deadlines of e.g. tax payments to the public sector.

Currency in circulation can exemplify the seasonal variation of financial time series. There is an obvious tendency for currency demand to peak in December, cf. Chart 1 – to a large extent caused by Christmas shopping, but also by a change of deposits into cash at the end of the year[3]. Such seasonal variation affects systematically the development from November to December and from December to January.

THE DEVELOPMENT OF *CURRENCY IN CIRCULATION* IN THE PERIOD JANUARY 2000 TO AUGUST 2006 — Chart 1



Note: The vertical lines indicate the months of December each year
Source: Danmarks Nationalbank, the balance and flow statistics for the MFI sector.

[3] Carlsen, M. and J. D. Riishøj, (2006).

Seasonal adjustment is a means for filtering out this systematic variation within the year, and thus easing the analysis of both the short-term and long-term development of the stock of currency.

The basic idea is that a given observed time series, Y_t, can be decomposed into unobserved seasonal and non-seasonal components that will provide useful and interpretable information. The decomposition can be either additive or multiplicative. Often financial and real macroeconomic time series are treated as multiplicative since the seasonal variation is dependent on the observed level. The decomposition can therefore be expressed as follows:

(2.1) $Y_t = T_t \cdot S_t \cdot I_t, \qquad t = 1,2,...T$

where Y_t is the observed time series; T_t is the trend, which not only includes the simple time trend, but also cyclical movements; S_t is the seasonal component, which is assumed to occur annually including calendar effects (such as trading day effects, etc.); I_t is the irregular component, which contains all effects not explained by the trend or the seasonal component (including extreme observations such as outliers)[4]. It is assumed that the seasonal component is exogenous and that its movements are predictable. Under the multiplicative approach the seasonally adjusted figures are obtained by dividing the original series by the estimated seasonal component:

(2.2) $Y_t / S_t = T_t \cdot I_t, \qquad t = 1,2,...T$

Even though the movements of the seasonal component are assumed to be predictable, they can change over time. In fact, it is quite likely that as time goes by, the seasonal pattern will change as a result of institutional changes. Therefore the empirical analysis of the movements of the seasonal component becomes important[5], and should be monitored since it affects the choice of seasonal filters.

3. Methodology and X-12-ARIMA

X-12-ARIMA has been developed from a (pure) moving-average approach (X-11) based on the definition of seasonality as systematic variation within a given year. The package has been further developed by the U.S. Census Bureau and Statistics Canada, with important contributions from e.g. the Deutsche Bundesbank. A main advantage is that it now combines the

---

4 The effects of trading days and outliers will be explained in subsequent sections.

5 The reader is referred to Pedersen, M.K. (2006), for an analysis of the development of the seasonal component in the five financial time series for which Danmarks Nationalbank is publishing seasonally adjusted data.

moving average technique with time series modelling. The Autoregressive-Integrated-Moving-Average (ARIMA) modelling aids the identification of extreme values and of level shifts that could distort the seasonal adjustment. It ensures that forecasted values can be used in the estimation of seasonality at the end of the series, which further improves the quality of seasonal adjustment. Another advantage is that wide ranging statistical diagnostics are provided by the X-12-ARIMA package, including a separate graphics feature available in SAS, which enables a simple procedure for monitoring the nature and quality of the seasonal adjustments.

The steps in the X-12-ARIMA procedure are illustrated in the Appendix, Chart 3.1. The package basically consists of two parts: (i) the regARIMA model consisting of a linear regression and ARIMA modelling of the estimated residuals of the linear regression. The preadjusted time series of ARIMA models and forecasts are provided to calculate symmetric seasonal filters (the seasonal component); and (ii) the X-11 part, where the seasonal component is calculated as a moving average and the final adjustment is performed. We shall now explain these two parts in more detail.

3.1. The regARIMA part

The purpose of regARIMA part of the X-12-ARIMA package is to improve the outcome of the seasonal adjustment in the X-11 part of the package.

3.1.1. The regression model

The linear regression is used in order to remove the day-of-the-week effect – a user-defined version of the trading day effect built into the X-12-ARIMA package, to remove the effects of known (and unknown) outliers in the current data and to improve the forecasts. The forecasts are used to calculate the seasonal component at the end of the series by using symmetric filters and thereby improving the estimation of the seasonal component such that the revisions later on are minimized.

Initially, a model is estimated on the regressors and on the basis of this model, the automatic outlier selection is performed (see also section 6.2).

The linear regression can be written as:

(3.1) $$y_t = \sum_{i=1}^{k} \beta_i x_{it} + y_t^*$$

where y_t is the observed time series in question, x_{it} are regression variables (modelling the day-of-the-week effect and the outliers), β_i are regression parameters, k is the number of regressors, y_t^* the residual from the regression model, $t = \{1,2,.....,n\}$, and $i = \{1,2,...,k\}$.

3.1.2. The ARIMA model

A convenient assumption when modelling economic time series is that the residuals should be independently and identically distributed and uncorrelated. In practice, this is difficult to obtain and therefore ARIMA modelling is introduced to model the estimated residuals from the linear regression in order to achieve uncorrelated error terms. The regARIMA model can be written in a single equation as:

(3.2) $\phi(L)\Phi(L^s)(1-L)^d(1-L^s)^D(y_t - \sum_i \beta_i x_{it}) = \theta(L)\Theta(L^s)e_t$,

where L is the lag operator, L^s is the seasonal lag operator, this means that $(L^s)^1$ is the same month last year such that $(1-L^s)^1 y_t$ is given by $y_t - y_{t-12}$, while $(L^s)^2$ is the month two years ago (i.e. $(1-L^s)^2 y_t = y_t - y_{t-24}$) and so on. As illustrated $(1-L)^d(1-L^s)^D$ implies differencing of the *d*th order and seasonal differencing of the *D*th order. $\phi(L) = (1-\phi_1 L - ... - \phi_p L^p)$ is the non-seasonal AR-operator, $\Phi(L) = (1-\Phi_1 L - ... - \Phi_P L^P)$ the seasonal AR-operator, $\theta(L) = (1-\theta_1 L - ... - \theta_q L^q)$ the non-seasonal MA-operator and $\Theta(L) = (1-\Theta_1 L - ... - \Theta_Q L^Q)$ the seasonal MA-operator. Finally the error terms e_t's are assumed to be white noise (no autocorrelation present). The original time series as well as the regression variables are transformed prior to the ARIMA modelling (more on this in section 5) in order to avoid spurious relationships. The regARIMA model is estimated by exact maximum likelihood[6].

Based on the final regARIMA model, forecasts are made and the seasonal component is calculated without the interference of day-of-the-week effects and effects of outliers.

3.2. The X-11 part

In the X-11 part, the actual seasonal adjustment is performed using the outcome of the ARIMA modelling as input. This X-11 part is based on a moving average approach and the general procedure is outlined in the following steps[7]:

1. First of all X-12-ARIMA calculates a centred moving average over 12 months as a preliminary trend estimate. The value of the moving average in a given month is filtered out from the original series by division $\Rightarrow S_t I_{t,initial} = Y_t / T_{t,initial}$[8], where $T_{t,initial} = M_{12}^C(Y_t)$ is a centred moving average of 12 months[9].

---

6 See U.S. Census Bureau (2006).

7 The assumption is that the model used is multiplicative, $Y_t = T_t \cdot S_t \cdot I_t$.

8 Y_t is the raw series adjusted for day-of-the-week effects and outliers by the regARIMA model.

9 C.f. Chart 3.2 in the Appendix for an example of the calculations.

2. Secondly, the seasonal component is estimated by calculation of a moving average based on l months of the same type (e.g. January 1995, January 1996, January 1997, etc.). The size of l depends on the length of the filter[10]. The estimate is normalized by division of the seasonal component by a centred moving 12 (succeeding) months moving average of the seasonal component. This normalizes the seasonal component to achieve an average of 1 in a given year (1 is equal to no seasonal fluctuations in the series). This is an important part of the process since the theory of seasonal adjustment clearly states that seasonal fluctuations and the trend are orthogonal, i.e. these components have no influence on the evolvement of the series in the long run. The seasonal component can be stated as: $S_t = \left[M_l^C (S_t I_{t,initial}) / M_{12}^C (M_l^C (S_t I_{t,initial})) \right]$, where $(M_l^C (S_t I_t))$ is a centred moving average of l months of the same type based on the initial estimate of the seasonal and irregular component. $M_{12}^C (M_l^C (S_t I_t))$ is a centred moving average of the seasonal component within a given year[11].
3. Thirdly, the trend and the irregular component is calculated as $I_t T_t = Y_t / S_t$, and the trend is estimated by the use of a Henderson filter[12]. Basically, if a Henderson filter is applied to third degree polynomials, the result will be fitting the parabolas. This quality allows the cycles to pass through as a part of the trend. Finally an estimate of the irregular-seasonal component is calculated by division of Y_t with the new trend estimate.
4. At this step the seasonal component is estimated. In step 2 above the initial filter used to calculate the seasonal component (prior to normalization) is a $3x3$ filter[13]. In the following iterations, the filters are chosen by the size of the moving seasonality ratio (MSR). MSR approximately expresses the ratio of the size of fluctuations in the irregular component to the size of the fluctuations in the seasonal component – the larger the ratio, the longer the filter (moving averages) should be. A short filter (depending on fewer lags and leads of the monthly value to calculate the seasonal component) is appropriate when seasonality fluctuates and the irregular component does not dominate the series. Naturally a long filter is appropriate when the

---

[10] See section 7.2.

[11] For more information, see Ladiray, D. and Quenneville, B., (1999), Findley et al. (1998) and http://www.uow.edu.au/~craigmc/henderson.html

[12] For more information on the Henderson filter, see Findley et al. (1998) and http://www.uow.edu.au/~craigmc/henderson.html.

[13] A 3x3 seasonal filter is in fact a weighted moving average. It is a 3-term moving average of a 3-term moving average, c.f. Appendix for an illustration.

opposite is the case[14]. When the seasonal component (S_t) has been estimated, the irregular component is calculated as: $I_t = I_t S_t / S_t = Y_t / T_t S_t$.

5. In the fifth step outlier detection by the X-11 part is performed. Basically a moving standard deviation with a period of five years is calculated for each year. Values of the irregular component (in the given year) which deviate from 1 (the average of a multiplicative model) with more than 2.5 standard deviations are characterized as extreme values and assigned a weight of zero in the following calculations. The moving standard deviation is recalculated and the procedure repeated. Values which are less than 1.5 standard deviations from 1 are assigned the weight 1, and values between 1.5 and 2.5 standard deviations from 1 are assigned weights which depend linearly on their deviation from 1. The whole process (step 1 through 5) is run three times[15] to make sure that estimations of T_t, S_t and I_t are as accurate as possible. The second and third time the input (Y_t) is the output of respectively the first and second run. Finally, the seasonally adjusted series is calculated as[16]: $T_t I_{t,final} = Y_t / S_{t,final}$.

There are several diagnostic checks which can be used to assess the quality of the regARIMA model and the seasonal adjustment. These diagnostics and actual choices are described in section 7.

4. Direct versus indirect adjustment

Seasonal adjustment of time series that are aggregates of other time series can be performed in two ways, either by direct or indirect adjustment. When adjustment is performed on the aggregated time series it is called direct adjustment while indirect adjustment is performed when the subseries are seasonally adjusted individually and then aggregated.

The series treated in this paper can be disaggregated into independent series. For instance *M1* is made up of the two series *currency in circulation* and *overnight deposits*. In the same way *M2* is made by summing *M1* with *deposits with agreed maturity up to two years* and *deposits redeemable at notice up to three months*. This form of aggregation is known as vertical aggregation. Besides, all series are constructed by horizontal aggregation, i.e. by aggregating the individually reported series of different institutions

[14] See U.S. Census Bureau (2006).

[15] There is no outlier detection in the third run.

[16] The seasonally adjusted data do not contain day-of-the-week effects, but the outliers are a part of the series.

(banks, mortgage-credit institutes, etc.). It is not realistic and probably not even preferable to do adjustment of the series at a horizontal level.

The question of whether to perform direct or indirect seasonal adjustment at a vertical level has been subject to an intense debate in the literature[17]. This debate has shown that the choice of method is not unambiguous. Some main points will be covered here.

It is rare that the results produced by the two methods coincide. The conditions for this to occur are very strict: The decomposition of the model has to be purely additive, there should be no preadjustments for outliers, day-of-the-week effects, etc. and the filters used must be the same for all series. These conditions are not easily fulfilled for Danish data. All models chosen are multiplicative and preadjustments are performed (mainly for outliers) in each model. As a consequence the choice between direct and indirect adjustment is not trivial.

4.1. Two main criteria for choosing between direct and indirect adjustment

In choosing between indirect versus direct seasonal adjustment of the Danish data, the main criteria are listed below:

1. *Valid and well-behaved seasonally adjusted series*: A main criterion is that the seasonally adjusted series are valid in the sense that the models used to model the series are well-behaved, the revisions minimized and the seasonally adjusted series smooth.
2. *A smooth monthly production of seasonally adjusted series*: The implementation of the seasonal adjustment of the five series should be as smooth and easy as possible. This criterion is important from a practical point of view, since monthly analysis of the adjusted series is time-constrained. It is important and desirable from the user's point of view that the models are as simple as possible, while still maintaining a high quality of the seasonally adjusted data.

Based on these two criteria, and the results of both directly and indirectly adjusted *M1* and *M2*, a comparison shows that a direct adjustment is preferable. Firstly, the outcome of the models of direct adjustment for *M1* and *M2* are well-behaved. The seasonal component is stable, no seasonality is detected in the residuals and the seasonality is successfully removed from the data. Secondly, the models used for the subseries in the indirect approach have problems with a very noisy and dominating irregular component and with an unstable seasonal component. In fact the ratio of the

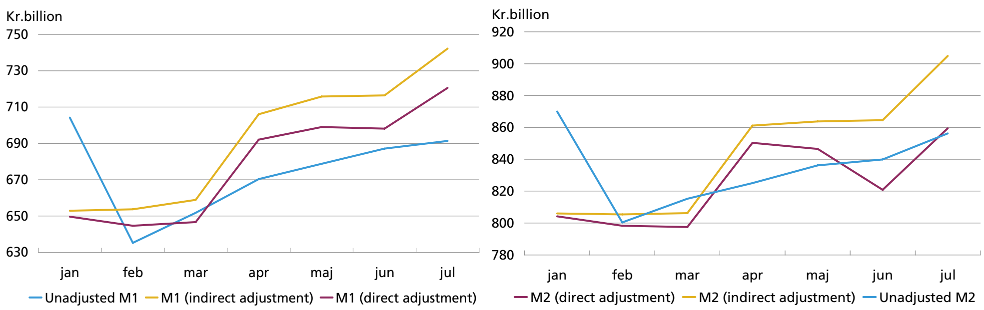
---

[17] See Manna, M. and R. Peronaci (2003).

variance of the irregular component to the total variance increases with the level of disaggregation.

Moreover, the values forecasted by the directly adjusted model are closer to the observed values than the values forecasted by the indirectly adjusted model, c.f. Chart 2. This is an important point because the seasonal filters are based on these forecasted values and the fit of the model is therefore very important.

FORECASTS FOR THE PERIOD FROM JANUARY TO AUGUST 2006 BASED ON DIRECTLY AND INDIRECTLY ADJUSTED SERIES Chart 2



Note: The deviation in January 2006 is caused by an outlier in the unadjusted data and is explained in section 6.3.
Source: Danmarks Nationalbank, the balance and flow statistics for the MFI sector and the authors' own calculations

In general, it was found that the noise in the individual series can be dampened by aggregation. This makes the seasonal component more stable and the irregular component less dominating in the directly adjusted series. Thus, direct adjustment of the five time series seems preferable.

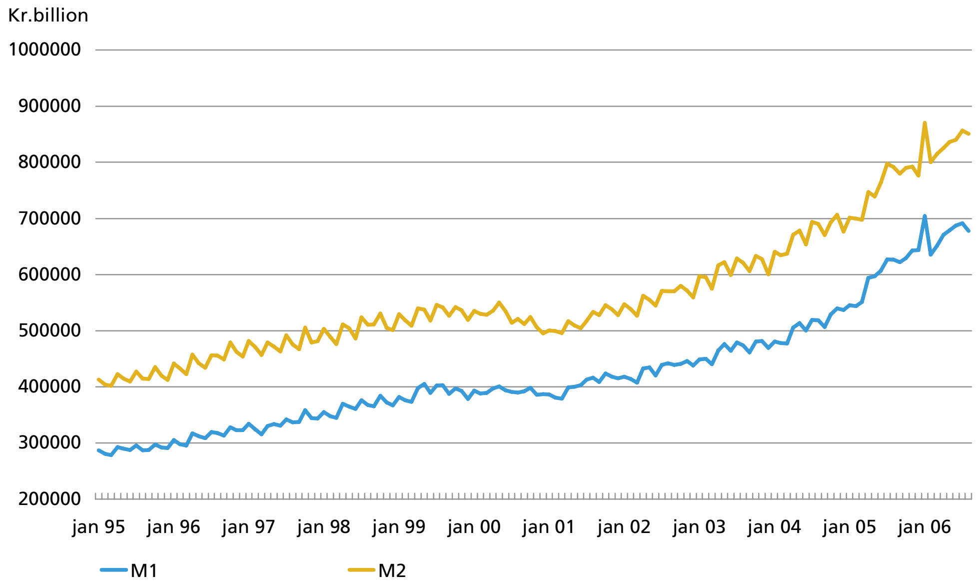
5. Time series properties of the financial data

Danmarks Nationalbank seasonally adjusts five financial time series: *Currency in circulation,* the money stock measures *M1* and *M2* as well as the MFI sector's *lending to households* and *to non-financial corporations*. In this section an inspection of the time series properties of particularly *M1* and *M2* is provided to explain the chosen transformations and preadjustments.

As already mentioned, ARIMA modelling of economic time series is subject to strong assumptions about stationarity and normality of the residuals. Assuming Y_t is a realization of the stochastic process y_t, then in order to fit an appropriate model it is required that the process is stationary, which is true when y_t has a constant variance ($Var(y_t) = \sigma^2$), fluctuating around a constant mean ($E(y_t) = \mu$) and have uncorrelated errors ($Cov(y_{t-k}, y_t) = 0$) for all t's.

Economic behaviour, as expressed in real macroeconomic and financial time series data, is often characterized by non-stationary behaviour; a fact which is further emphasized when inspecting the time series graphs of *M1* and *M2* in Chart 3 below.

THE DEVELOPMENT OF THE MONEY STOCK MEASURES *M1* AND *M2* FOR THE PERIOD JANUARY 1995 TO AUGUST 2006 Chart 3



Source: Danmarks Nationalbank, the balance and flow statistics for the MFI sector.

From Chart 3 it is obvious that neither of the time series show any tendency to return to or fluctuate around a constant mean as time evolves. Both series are clearly time-dependent and upward trending. This explains why transformation of the data is an important step towards meeting the requirements of stationary data, which, first of all, can be met by taking logarithms in order to reduce the time dependence of the variance.

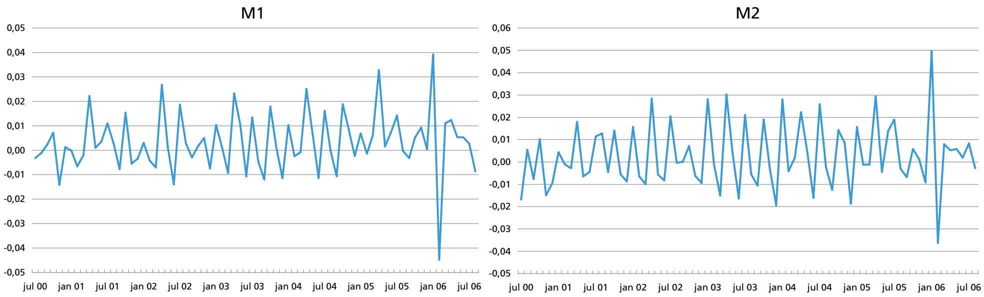
It seems obvious from Chart 3 that there is a break in the year 2000, the money stock measures of *M1* and *M2*, and the break in 2000 can also be detected if inspecting graphs of the three other financial time series investigated, cf. Chart 6.1 in the Appendix. It is caused by a change in the compilation method in mid-2000. Therefore, for the regARIMA modelling part, it has been decided that for all series only data from July 2000 until the present will be subject to ARIMA modelling, while the entire series (from January 1995 until today) are used for the seasonal adjustment part (in X-11)[18]. This decision reflects that the history before the break does not provide useful information for the present nor for forecasting.

---

[18] In Pedersen, M. K. (2006), the approach differs slightly, since the entire sample was initially treated as two separate samples, being subject to two different ARIMA models. This change of approach do not change the conclusions of the article.

Stationarity is not achieved by simple logarithmic transformation, since all of the five series are upward trending. By taking first differences, which are equal to having a series integrated of order 1 ($I(1)$), the series becomes stationary. This is illustrated below by the differenced series of *M1* and *M2*, where it is obvious from the development of the transformed time series for the period from July 2000-August 2006 that there is a mean-reverting, smooth behaviour, c.f. Chart 4. Also an early indication of the possible outliers in the series can be seen from both Chart 3 and 4, for instance in February 2006. Outlier detection will be explained in detail in section 6.2.

THE TRANSFORMED TIME SERIES (BY DIFFERENCING AND LOGARITHMS) OF *M1* AND *M2* FOR THE PERIOD FROM JULY 2000 TO AUGUST 2006 — Chart 4



Source: Danmarks Nationalbank, the balance and flow statistics for the MFI sector and the authors' own calculations

6. Preadjustment

In this section the importance of preadjusting the series prior to seasonal adjustment is more carefully analyzed. The series are preadjusted for events that affect the individual observations in a way that is not consistent with pure seasonal variation as estimated by the X-11 procedure (moving averages). If the data is not preadjusted, the seasonal component will not be valid in the sense that it will be estimated on the basis of observations which represent more than the seasonal pattern in the series. Usually preadjustment of time series is divided into two categories: Permanent and preliminary preadjustment. Preliminary preadjustment takes into account large deviations which can obscure the picture of the seasonal pattern, but which still contains information or news of interest to the analyst. These effects (outliers) are removed from the original series before calculating the seasonal component but re-entered as part of the seasonally adjusted series. Permanent preadjustment implies removing the effects before calculating the seasonal component and not reentering them as part of the seasonally adjusted series. The effects are thus permanently removed. This is done when the effects e.g. originate from institutional factors without representing news.

6.1. Permanent preadjustment

The permanent preadjustments are carried out in the regression part of the regARIMA model. The five time series which have been seasonally adjusted are all observed at monthly frequencies at the last day of banking in each month[19]. For this reason the standard regression variables for end-of-month stock data in X-12-ARIMA are not applicable, because these variables model the last day of the month and not the last day of banking (day-of-the-week effect). Therefore four new dummy variables $(d_{jt}'s)$ have been constructed in order to take account of the day-of-the-week effect:

The last day of banking in a month is defined as w.

$d_{jt}, j \in \{1,2,3,4\}$,
$d_{1t} = (1$ if w is a Monday; -1 if w is a Friday; 0 otherwise)
$d_{2t} = (1$ if w is a Tuesday; -1 if w is a Friday; 0 otherwise)
$d_{3t} = (1$ if w is a Wednesday; -1 if w is a Friday; 0 otherwise)
$d_{4t} = (1$ if w is a Thursday; -1 if w is a Friday; 0 otherwise)

The coefficient of the day-of-the-week regressors in a given month can be calculated as

$$(6.1)\quad a_t = \sum_{j=1}^{4} a_j d_{jt}$$

If the last day of banking is a Friday the effect of the last day of banking is given by:

$$-\sum_{j=1}^{4} a_j = \sum_{j=1}^{4} a_j d_{jt}$$

when w is a Friday.

To make sure the regressors only capture the day-of-the-week effect resulting from the last day of banking, the level and seasonal effects of the regressors are removed from the regression variables beforehand. If the variables have a periodic cycle of 12y months, such that $f_t \approx f_{t+12}$ (y equals the number of years), then the approximate seasonal and level effect is calculated as

$$(6.2)\quad f_t^* = \frac{1}{y}\sum_{j=1}^{y} f_{t+12y} \qquad t = 1,2,\ldots,12$$

equal to the calendar month mean estimating the level and seasonal effect of a given regression variable. The seasonal and level effect of the regressors is removed by subtracting f_t^* from f_t.

---

[19] In Denmark the banks are only open Monday to Friday. This means that if the last day in one month is a Saturday or Sunday then the last day of banking in this month is a Friday.

The distribution of days within a year is periodic with a 28 year cycle. These constructed dummy variables are therefore also periodic with a period of 28 years. However, a few deviations have been found:

- As of 2003 the last day of banking in December can no longer be the 31st, but the 30th. This means that the periodicity of December has shifted.
- Easter can occur at the end of March and therefore the last day of banking in March does not follow the 28 year cycle completely.

These deviations are, however, considered small and the seasonal effects of each regressor (f_t^*) is calculated based on the 28 year cycle. Based on this assumption the long-term monthly means are the same for every month and across the regressors (=-2/7): In 28 years the weekdays, Monday to Thursday, will be observed four times (for each month). This means that the last day of banking will be Monday, Tuesday, Wednesday or Thursday 1/7 of the time in a given month. Friday will be observed 12 times in 28 years (when the last day of the month is a Saturday or a Sunday the last day of banking will be a Friday), which leads to Friday being the stock taking day 3/7 of the time. Since this is true for every month and Friday as stock taking day is represented by a -1 in each dummy the long-term mean of every regressor for every month is as stated equal to -2/7.

All series have been tested to see if they are sensitive to which day of the week is the last day of banking and only *currency in circulation* is significantly sensitive to this effect.

6.2. Preliminary preadjustment

Economic time series data is often subject to unexpected interventions resulting in extraordinary observations or outliers. Outliers in time series can seriously bias the model identification, estimation and inference if these observations are not filtered out prior to modelling and seasonal adjustment. The reason is that the normality condition for the estimated residuals of economic time series is often violated as a result of data being affected by deterministic effects resulting from day-of-the-week effects and outliers. The quality of the final seasonal adjustment is therefore strongly dependent on the outlier detection, since the seasonal filters are directly derived from the ARIMA model. If prior knowledge about a possible outlier due to, for example, a change in the compilation method or to a sudden large change in demand, the effect should be included as a regression variable, rather than being assigned to the calendar, since it does not happen regularly.

There are three main types of outliers: (i) additive outliers (AO), appropriate when only one observation is unusual and the effect disappears immediately, (ii) transitory change (TC), which is comparable to an AO but with a slowly (exponentially) decreasing effect after a couple of observations, or in this case months, and (iii) level shift (LS), included when there is a permanent shift in the time series to a different level from a certain time point onward. Only AO and LS are found in practice in the preadjustment of the Danish data.

A procedure to automatically detect and replace outliers is included in X-12-ARIMA and is applied to the regARIMA modelling part. Outliers are classified according to their specific type and are modelled as regression polynomials expressed as functions of the lag operator *L*, always assigned to the irregular component[20]. The procedure can be demonstrated by considering the time series y_t, which is assumed to follow an ARIMA model of order $(p\ d\ q)^s$(P D Q) as stated in eq. (3.2). The time series being subject to extraordinary events or outliers can be expressed by the following model:

(6.4) $$y_t^* = y_t + \sum_{i=1}^{k} \xi_i(L)\omega_i I_t(t_i)$$

where $\xi_i(L)$ determines the dynamics of the outlier occurring at time t_i, ω_i is the impact of the outlier and $I_t(t_i)$ is an indicator variable taking on the value 1 at t_i and 0 otherwise, thus indicating the occurrence of the outlier. The time series is subject to *k* outliers at time points $t_1, t_2, \ldots, t_k$. In terms of regression polynomials the three different types of outliers detected by X-12-ARIMA can be expressed as follows:

AO: $\xi_i(L) = 1$

TC: $\xi_i(L) = 1/(1-\delta L)$, where δ determines the speed of decay in such a way that close to zero represents a fast decreasing function and close to 1 represents a slowly decreasing function.

LS: $\xi_i(L) = 1/(1-L)$

The procedure selects outliers by comparing regressor t-values to a critical value, which is either user-defined or the program's default value (default critical value=3.3)[21]. The number of outliers should be kept at a minimum if they cannot be assigned any meaningful economic explanation ex-post and need to be quite significant in order to leave them in the regression. In the preadjustment of the Danish data, critical values are set manually to 3.0.

---

20 ECB (2000), p. 19.
21 McDonald-Johnson, K.M. and C.C. Hood (2001) and Findley et al. (1998).

The outlier detection and parameter estimation is performed by an automatic iterative procedure built into the X-12-ARIMA, where AO, TC and LS can be identified[22]. First, the modelling of the original series is performed supposing that there are no outliers. Then the outlier detection step and the parameter estimation steps are undertaken interchangeably until all outliers are replaced. Any outliers detected are included as regression variables and 'final' time series parameters are subsequently estimated in ARIMA models.

In practice, outlier detection is performed in the regARIMA part, but, in addition, X-12-ARIMA also offers the possibility of detecting and replacing outliers in the X-11 part, as described in section 3.2[23].

6.3. Pre-adjustment of the Danish time series – a few examples

In the following a few examples will be given on the specific preadjustment performed prior to seasonal adjustment of the Danish data:

Firstly, additive outliers are detected in the *MFI sector's lending to households* in August 2002 and 2005. This is a regular deviation from trend every third year, and is related to the fact that mortgage-credit lending is still the most common source of lending to Danish households for housing purchases (constituting a large proportion of the total lending to households of approximately 90 per cent). Mortgage-credit institutes' lending is primarily characterized by long maturity lending and thus financed by long maturity bonds. Every third year, by the end of August, latest in 2005, the underlying bonds are refinanced and up to this event the growth rate in lending is characterized by extraordinarily large increases since households' make use of this opportunity to raise a loan in the discontinued series making use of the higher market value (associated with shorter time to maturity) as opposed to the lower market value on the new bond series. Since the effect is occurring every third year in August (1996, 1999, 2002 and 2005), it cannot be assigned as a regular seasonal effect occurring within a year, but is a significant effect that should be filtered out prior to ARIMA modelling.

Secondly, for the money stock measures, *M1* and *M2*, large additive outliers are noticeable and have to be filtered out in January 2006. This is the result of a sudden large increase in the money stock measures caused by one large bank having extraordinarily large overnight deposits in this particular month. This effect was also noted from the time series graphs in section 5, Chart 3 and 4.

---

[22] The outlier detection approach built into X-12-ARIMA is partly based on that of Chang, Tiao and Chen (1988) where a detailed description of the procedure can be found.

[23] The procedure is described in ECB (2000).

Thirdly, the attention is on the implementation of the day-of-the-week effects in the seasonal adjustment. Of the five series adjusted, *currency in circulation* is the only one for which day-of-the-week effects are significant. Actually, the day-of-the-week effects are quite strong in this series[24], since *currency in circulation* is generally highest on Fridays and lowest on Tuesdays and Wednesdays. This reflects that demand for cash increases just before the weekend. *Currency in circulation* is adjusted for this effect to achieve a seasonal component as reliable as possible. Because the regARIMA models used to make forecasts on the *currency in circulation* are only based on the period from July 2000 and onwards, a second regARIMA model for the period from January 1995 and until June 2000 has been made for this particular series. The objective of this second model is to adjust for day-of-the-week effects in the entire sample period. Some outliers have also been detected and replaced in the X-11 part during this period[25]. Based on the preadjusted observations for both periods the seasonal component is estimated.

In Table 7.1 of the Appendix all outliers as well as day-of-the-week regressors included in the models are shown.

7. Model selection criteria

Since the concept of seasonality is not theoretically well-defined, it is important to be explicit about the specific choices made in order to be able to evaluate the implemented seasonal adjustment procedure over time. The selection procedure used by Danmarks Nationalbank for Danish data is based on the following principles:

1. The regARIMA model should be appropriate such that it fulfils the assumptions of well-behaved normally distributed and uncorrelated residuals. This includes a justification of the interpretability of the parameter estimates and of the detected outliers (sometimes it is possible to give a structural interpretation of the outliers, sometimes the outliers simply reflect large random shocks).
2. Seasonality should be present in the adjusted series. This involves testing for the presence and nature of seasonality (F-test).
3. Seasonal filters should be to minimize the fluctuations of the irregular component and to reduce moving seasonality (M-test).

---

[24] See the Appendix and Table 7.1 for *currency in circulation* where the chi-squared test for the group of regressors shows that the day-of-the-week effect is highly significant.

[25] The regARIMA model and diagnostics for the early period of *currency in circulation* is not printed in the Appendix, since no forecasts are made based on this model. For further information please contact Statistics, Danmarks Nationalbank.

4. The quality and stability of the seasonal component should be evaluated by periodograms and sliding spans analysis.

All test statistics mentioned in the above criteria for an appropriate model selection and adjustment procedure as well as their specific use in formally testing the modelled Danish time series will be presented in the following.

7.1. The regARIMA part

The model selection (choosing the most appropriate regARIMA model among different alternatives) is based on various criteria, which have been given equal weight in the evaluation of the estimated models. These include information criteria (Akaike, Hannan-Quinn and BIC), test for no autocorrelation, and for normality (investigating histograms of the residuals and a measure of kurtosis). The test statistics as well as their use in relation to Danish data is presented in Table 7.2 of the Appendix.

First of all the regARIMA model identification is based on an evaluation of autocorrelation functions (ACF's) and partial autocorrelation functions (PACF's) of the regression residuals based on the correct differencing order (including both regular and seasonal differencing). An important basic prerequisite for diagnostic checking of the chosen regARIMA models is whether the residual term is white noise, i.i.d. $N(0,\sigma^2)$, and is performed through various analyses of the residuals. The test for no-autocorrelation can be assessed by use of both ACF's and PACF's and the Ljung-Box portmanteau statistics (Q-statistics)[26]. Basic descriptive statistics, such as histograms of the residuals and kurtosis are provided in the output file in X-12-ARIMA and is consulted in order to decide whether the regARIMA models fulfil the normality condition[27].

The information criteria are used for holding nested models up against each other. The idea is that the loglikelihood value of the estimated models should be minimized, and models with the optimal number of parameters (/lags) are chosen in order to rule out serious autocorrelation of the residuals[28].

Outlier detection – as already mentioned - is also an important issue of diagnostic checking. Only by replacing these outliers iteratively, optimal

---

26 See eq. (A.7.1) in the Appendix.

27 The estimated kurtosis with a mean of three ensures that the residuals follow a normal distribution. Skewness is another measure being equally important for deciding whether the normality condition is fulfilled and is sometimes more important than the kurtosis, but this is a feature not built into the version of the X-12-ARIMA program used at present by Danmarks Nationalbank.

28 In eq. (A.7.2)-(A.7.4) in the Appendix, formulas for the information criteria are presented.

regARIMA models can be obtained. Based on ARIMA models, forecasts are made. These forecasts should, as a final evaluation, be addressed in order to reach the best ARIMA models with respect to out-of-sample forecast performance. The criterion is the mean squared forecast error also included in the appendix, table 7.2.

7.2. Seasonal filters

Whether there is seasonality in the investigated series can be determined both by visual inspection and by formal tests: F-test and M-tests, which will be presented in the following. The tests for choosing appropriate seasonal filters can be divided into two groups: i) introductory tests that give an indication of which filter to choose and ii) evaluating tests assessing the quality of the seasonal adjustment.

7.2.1. Introductory tests

These tests are meant to give an idea of how significant the seasonality in the series is and how the seasonal filter should be formed.

7.2.1.1. Test for presence of seasonality

This parametric test for the presence of seasonality is based on the assumption that seasonality is stable so that the seasonal component can be approximated by the monthly averages of the series. The test is an F-test given by:

(7.5) $$F_S = \frac{S_A^2/(k-1)}{S_R^2/(n-k)} \sim F(k-1, n-k)$$

The null hypothesis, H_0, is that there is no (stable) seasonality present in the series.

$$\frac{1}{n}S_A^2 = \frac{1}{n}\sum_{i=1}^{k} n_i(\overline{X}_i - \overline{X})^2$$

is the variance explained by stable seasonality and S_A^2 is the "inter-month" sum of squares, $(k-1)$ is the degrees of freedom of the variance of the monthly averages,

$$\frac{1}{n}S_R^2 = \frac{1}{n}\sum_{i=1}^{k}\sum_{j=1}^{n_j}(X_{ij} - \overline{X}_i)^2$$

is the residual variance in which S_R^2 is the residual sum of squares and $(n-k)$ is the degrees of freedom of the residual variance. i is the month ($i \in \{1,2,...,12\}$, n is the total number of observations, n_i is the number of observations of month i, and k is the number of observations within the year ($k = 12$ in this case).

The total variance is equal to $\frac{1}{n}(S_A^2 + S_R^2)$. The larger $S_A^2/(k-1)$ is relative to $S_R^2/(n-k)$, the more of the variance is explained by the seasonal component and the more significant is the seasonal pattern. However, the user has to be aware that seasonality is assumed to be stable throughout the entire sample. One should also be aware that the series used for this test is the original series (adjusted for day-of-the-week effects and regression outliers) without the *preliminary* trend estimate. This means that residuals are likely to be correlated and therefore the test statistic is compared to a high critical value to decrease the risk of rejecting a true null hypothesis of no seasonality.

7.2.1.2. Test for moving seasonality

If seasonality is in fact present it is relevant to focus on the nature of the seasonality. The test for moving seasonality is informative in the sense that it indicates the size of the filter needed and whether it is even possible to capture the seasonal variation in a satisfactory way. The test is based on the assumption that the seasonal-irregular component (SI) can be divided into three effects:

(7.6) $SI_{ij} - 1 = X_{ij} = m_i + b_j + e_{ij}$

where m_i is the monthly (seasonal) effect, b_j is an annual effect, where $j = 1,2,...,N$ is the number of years in the time series and e_{ij} is the residual effect which is assumed to be independent and identically distributed with a zero mean. Total sum of squares can be decomposed into:

(7.7) $S^2 = S_A^2 + S_B^2 + S_R^2$

where S_A^2 is "inter-month" sum of squares from above, S_B^2 the "inter-year" sum of squares given by $S_B^2 = k\sum_{j=1}^{N}(\overline{X}_j - \overline{X})^2$
and S_R^2 the residual sum of squares given by [29]

$$S_R^2 = \sum_{i=1}^{k}\sum_{j=1}^{N}(X_{ij} - \overline{X}_i - \overline{X}_j + \overline{X})^2$$

The F-test is given by

(7.8) $F_M = \dfrac{S_B^2/(k-1)}{S_R^2/[(N-1)(k-1)]} \sim F(N-1,(N-1)(k-1))$

under the null hypothesis that the effects b do not change from year to year and hence there is no change in the seasonal pattern. If b is stable, $\overline{X}_j$ is close to $\overline{X}$, and the annual sum of squares close to zero. It is reasonable to

[29] S_R^2 is different from S_R^2 in the test for presence of seasonality in section 7.2.1.1. In the present test in addition to the monthly effect an annual effect is also assumed, which leads to a different expression for S_R^2. See Ladiray, D. and B. Quenneville (1999).

think of b as a measure of change in the seasonal component within the year, but since b is estimated by the yearly average, it should be noticed that the test does not capture changes within the year that are of the same size but in opposite directions.

In general the faster the seasonality moves, the shorter the filters have to be.

7.2.1.3. M7

The M7 test is a test that combines the two F tests described above and is an indication of whether or not the seasonality in the series is identifiable. The test value has been derived as[30]:

$$(7.9)\ M7 = \sqrt{\frac{1}{2}\left(\frac{7}{F_S} + \frac{3F_M}{F_S}\right)}$$

Seasonality is identifiable if the distortion in the final seasonal component is low. High distortion indicates that the seasonality pattern is changing too rapidly and in an unsystematic way to be pinned down. The amount of distortion is positively correlated with the moving seasonality statistics (F_M) and negatively correlated with the statistics for presence of seasonality (F_S). If M7 exceeds 1 the test suggests that it will be difficult to seasonally adjust the series in an appropriate way.

These three tests (the F-test for presence of seasonality in the data, the F-test for moving seasonality and M7) are given most weight in the Danish seasonal adjustment. Not only do the measures contain information on whether or not seasonality is present in the data, but they also give information concerning the nature of the seasonality. Finally the simplicity of the tests is a great strength – the results are easy to interpret and the weaknesses evident. Together with the SI-ratios (below) they serve as evaluation tools prior to estimation of the seasonal component.

7.2.1.4. SI ratios

The SI ratios are the result of removing the trend-cycle component from the analyzed series: $SI_t = Y_t / T_t$ equal to the seasonal-irregular component. If the evolvement of the SI ratios of a given month is not continuous but fluctuates a lot, it indicates that the filter may be too short, and the irregular component dominating in the series. If short filters are used, the danger is that the seasonal component will reflect the irregular component and not just the seasonality. On the other hand, if the SI ratios do not fluctuate more rapidly but continuously, short filters may be very appropriate. All in all the evolvement of the SI ratios should be viewed as a useful indicator to

---

[30] See Lothian J. and M. Morry (1978).

identify the length of filters, but it does not represent the result of a formal test.

7.2.2. Evaluating tests

These tests are used to evaluate whether all of the seasonality has been successfully removed by the seasonal adjustment[31].

7.2.2.1. Test for residual seasonality

This test is the same as the one used for testing for the presence of seasonality in the original series (cf. section 7.2.1.1). In this case the test is performed on the seasonally adjusted series. Again one should remember that the test assumes stable seasonality. Therefore, if there is unstable and moving seasonality in the residuals, the test might not identify it.

7.2.2.2. M8 and M10

The M8 and M10 tests are stability tests used for detection of large random fluctuations in the seasonal component. If the seasonal component exhibits extreme fluctuations, the moving average approach will not be a qualified tool for seasonal adjustment of the particular series. M8 tests for fluctuations throughout the entire sample of the series while M10 tests for fluctuations at the end of the series. M8 is given by:

$$(7.10)\quad M8 = 100\left[\frac{1}{J(N-1)}\sum_{j=1}^{J}\sum_{i=2}^{N}\left|S'_{Ji+j} - S'_{J(i-1)+j}\right|\right]\frac{1}{10}$$

where $J = 12$, N is the number of years in the sample, S'_t is the normalized value of S_t and $\sum_{j=1}^{J}\sum_{i=2}^{N}\left|S'_{Ji+j} - S'_{J(i-1)+j}\right|$ is the sum of the absolute year-to-year changes in the normalized seasonal component.

The fluctuations are too large if the test statistics exceed 1. Intuitively this means that the average change in the seasonal component from one year to the next should not exceed 10 per cent. If it does exceed this value, the model should be reconsidered.

M10 is equivalent to eq. (7.10) except that it only covers the last years of the period for short-term analysis. If M10 exceeds 1 while M8 is acceptable, it could indicate a change in the structure of the series and therefore considerations whether or not the model should be changed in the future are appropriate.

---

[31] This section is for the most part based on formulas explained more in detail in Lothian, J. and M. Morry (1978).

7.2.2.3. M9 and M11

The M9 and M11 tests are used to measure the degree of linear movement in the seasonal component. Compared to the tests of M8 and M10, M9 and M11 measure the average (absolute) change in the seasonal component from year to year. M9 is given by:

$$(7.11)\quad M9 = 100\frac{\sum_{j=1}^{J}\left|S'_{J(N-1)+j} - S'_j\right|}{J(N-1)}\frac{1}{10}$$

equal to the sum of the changes in the seasonal component for a given month over the period (except the last year). If these changes are random then $\lim_{N\to\infty}(S'_{J(N-1)+j} - S'_j) = 0$. A test value larger than 1 indicates that the linear movement in the seasonal component is larger than 10 per cent on average.

M11 is equivalent to eq. (7.11), but again it only concerns the last years of the data analyzed. Initially, before the regARIMA model was used for forecasting, the M11 was important because asymmetric filters used at the end of the time series can only identify a constant seasonal component. With the regARIMA model, the linear movement in the seasonal component can be estimated and forecasted also at the end of the sample which means that a negative result of the test is not as aggravating as earlier on. But both M9 and M11 are still informative tests on the evolvement of the seasonal component.

7.2.2.4. Periodograms of the seasonally adjusted series and the irregular component

The periodograms are graphs printed in the output file of X-12-ARIMA showing the series in the frequency domain instead of in the time domain. This means that the area under the graph in the periodogram represents the total variance of the series and can be decomposed into contributions from different frequencies. The periodogram at the seasonal frequencies gives a visual indication on whether or not the user has been successful in "removing" the seasonality present. The periodograms of the differenced seasonally and outlier-adjusted data as well as the irregular component adjusted for extreme values are investigated in order to check for remaining seasonality in the series[32].

---

[32] The difference between the differenced seasonally adjusted series and the irregular component is that the seasonally adjusted series have been detrended by differencing while the irregular component is the seasonally adjusted series without the trend-cycle. The difference should not be large.

The seasonal frequencies are ϖ =1/12 (yearly frequencies), 1/6 (semi-annual frequencies), 1/4 (every fourth month), 1/3 (every quarter) and 1/2 (every other month). If peaks at these frequencies are present, X-12-ARIMA has done a poor job of seasonally adjusting the data. Significant peaks for ϖ =1/3 are found for the unadjusted series. It is therefore important that these peaks are no longer present in the periodograms of the differenced seasonally and outlier adjusted data nor in the periodogram of the irregular component adjusted for outliers.

7.3. Sliding spans analysis

It has already been shown that most important for seasonal adjustment to be successful is stability of the estimated figures. X-12-ARIMA offers (as a new feature compared to earlier versions) two types of stability diagnostics, *sliding spans* and *revision history* which are used for evaluating whether the seasonal adjustment is significantly changed over the sample. We have only used the sliding spans analysis and this procedure will be explained in the following.

The sliding spans diagnostics are descriptive statistics of how the seasonal adjustment and month-to-month as well as year-to-year changes vary when the span used to calculate them is altered in a systematic way[33]. As such the analysis compares seasonal adjustments of overlapping subspans, checking their stability against each other. The entire sample is divided into four subsamples of the same length. The second sample is obtained by deleting the 12 earliest observations (1 year) in the initial sample, appending the sample one year in the current end. The third sample is obtained by deleting the 12 earliest observations in the second sample, appending the sample one year in the current end and so forth also for the fourth sample, such that this last sample contains the most recent observations. Seasonal adjustment of each span is performed as though it constituted the entire sample and each month in common for more than one span is compared to see if the seasonal adjustment varies significantly from span to span[34]. When comparing the estimated values across spans, large deviations in the estimated seasonal component (S(%)), month-to-month changes (MM(%)) or year-to-year changes (YY(%)) indicate unreliability of the estimates. Excessive variability is defined by X-12-ARIMA and will be flagged by the program if[35]

$$(7.12)\quad \max S_t(\%) = \frac{\max S_t - \min S_t}{\min S_t} > 0.03$$

---

[33] Findley, D. F., B. C. Monsell, H. B. Shulman, and M. G. Pugh (1990)

[34] Findley D. F. and B. C. Monsell (1984).

[35] It should be noted that the given threshold of 3 per cent is a default value included in X-12-ARIMA, which can be changed manually by the user.

where $S_t(\%)$ is the seasonal component estimated from span *k, k=1,2,3,4*, for month *t*. The relation is meant to test whether the maximum percentage difference in the seasonal component for month *t*, $S_t(\%)$, is greater than three per cent. This default value can be changed manually by the user and is in this case changed for all series to 1 per cent, since the seasonal variation is quite stable. A low threshold value also gives somewhat more reliable results for the seasonal adjustment. Another relation measures whether the seasonally adjusted month-to-month or year-to-year changes are unstable, i.e. larger than the same threshold by

(7.13) $\max MM_t(\%) = \max MM_t - \min MM_t > 0.03$[36],

where $MM_t^k = \dfrac{A_t^k - A_{t-1}^k}{A_{t-1}^k}$

and A_t^k denotes the seasonally adjusted value for month *t* when the seasonal adjustment procedure considered is only applied to data in the *k*-th span. MM_t can also be substituted by YY_t to measure the stability of the estimates of the year-to-year changes. The test can be interpreted as evaluating whether the level of the seasonally adjusted series vary substantially from span to span[37].

The criterion for determining whether the seasonal adjustment is unstable and unreliable based on too many observations flagged from the sliding spans analysis is given by the following[38]: (i) If less than 15 per cent of the observations are flagged, the series are reliably estimated; (ii) If more than 25 per cent of the observations are flagged, the seasonal adjustment is not stable and should not be undertaken. In between is the grey area, where the decision is inconclusive.

The use of the sliding spans diagnostic is particularly relevant for determining if a series is being adequately adjusted[39].

7.4. Model selection criteria – Danish data in general

In conclusion, these test statistics confirm the seasonal adjustment of the Danish data. In Table 7.2 of the Appendix some of the various test statistics are presented and by inspection it becomes clear that none of the estimated models suffer from serious autocorrelation or lack of normality.

---

[36] The three per cent default value in (7.13) can also be selected manually by the user and is for all series equally set to 0.01 as in (7.12).

[37] It is, however, worth noting that if a month-to-month change is unreliably estimated, it will not only affect the given month but also surrounding observations and thus more than one month-to-month change will be flagged.

[38] This criterion is based on Findley, D. F., B. C. Monsell, H. B. Shulman, and M. G. Pugh.(1990), pp. 4-5, and Findley D. F. and B. C. Monsell (1984), pp. 7-10.

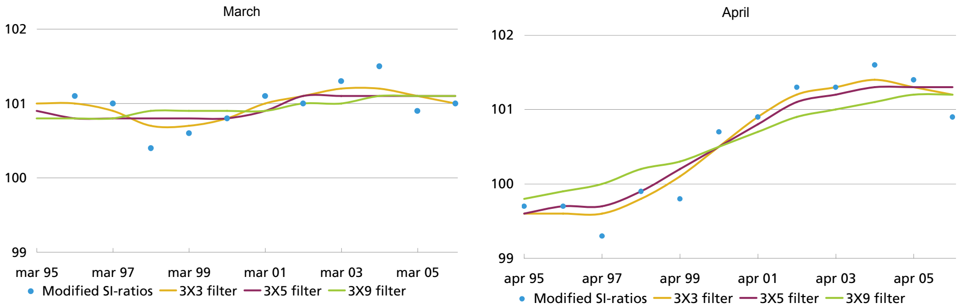
[39] It could also be used for deciding on indirect or direct adjustment or filter length.

Currency in circulation and *lending to households* have moving seasonality as defined by the F_M statistics, but none of the time series violate the assumption of identifiable seasonality. This confirms that seasonal adjustment of the series is manageable.

As an example, in Chart 5, the seasonal-irregular components (SI-ratios) of March and April for *lending to non-financial corporations* are shown in the left and right panel, respectively. It is obvious that the evolvement of the component is quite different in the two graphs. In the April graph, we see a significant change over the time period considered, but the component does not seem to fluctuate much around the trend of the ratio (no dominating irregular component). It is therefore appropriate to use a short filter (3x3) to capture the evolvement of the seasonal component. On the other hand, the SI-ratio for March is quite steady and there is little difference between using a 3x3, 3x5 or 3x9 filter. A 3x5 filter is chosen (see also the Appendix).

MODIFIED SEASONAL-IRREGULAR COMPONENTS (SI-RATIOS) AND FILTERS CALCULATED FOR THE *MFI SECTOR'S LENDING TO NON-FINANCIAL CORPORATIONS* FOR MARCH (IN THE LEFT PANEL) AND FOR APRIL (IN THE RIGHT PANEL)

Chart 5



Note: The reader should be aware that modified SI-ratios are calculated on the basis of the filters chosen which means that the size of the SI-ratios deviate depending on whether a 3X3, 3X5 or 3X9 filter is chosen. The SI-ratios are calculated on the basis of the filters chosen for the monthly component: A 3X3 filter for the April component and a 3X5 filter for the March component.

There is no sign of residual seasonality in either of the adjusted series. There are, however, problems with the M10 and M11 statistics (see the Appendix) for *lending to non-financial corporations* but not with M8 and M9. This could indicate that the seasonality of the series is changing and we should keep a close eye on the series in the future.

In general the seasonal adjustment of all five series investigated seems stable. Even when the threshold value of the sliding spans analysis is set to 1 per cent, the number of observations flagged is less than the 15 per cent limit for all time series except for the month-to-month changes of M2. If the 3 per cent threshold value was kept, no observations would have been flagged. Stability of the estimated figures is important, but it does not mean, however, that revisions are not needed, as will be seen in the following.

8. Revisions

Revisions are unavoidable when dealing with seasonally adjusted data; the uncertainty attached to the figures can either be a result of the seasonal adjustment process itself or of revisions of the unadjusted data. The effect of revisions of the unadjusted data can in general affect observations throughout the entire sample of the time series and can therefore be quite dispersed. However, the uncertainty resulting from the adjustment method itself is normally ascribed to the most recent observations of the time series representing the new information. The inclusion of new information affects the adjustment of the time series as a result of X-12-ARIMA using centred moving averages to forecast observations for trends and the seasonal component. When new information is added, forecasts will be replaced by actual observations, which can cause deviations and thus revisions of data. The size of the revisions resulting from the replacement of forecasts by actual observations depends on how good the forecasts are, and this is expressed by the forecast errors of the regARIMA model. Therefore a minimized forecast error is preferable when deciding on which ARIMA model describes the data generating process most appropriately. The use of regARIMA modelling and forecasted future values makes it possible to use symmetric filters at the ends of the series, and this should, in general, reduce the size of revisions – compared to earlier versions of X-12-ARIMA[40].

In general, revisions resulting from new data will be an important improvement of the adjusted series. However, there will always be a conflicting choice between providing the most accurate seasonally adjusted data as opposed to not revising data too often. This is the choice between implementing revisions as new information becomes available, i.e. each month for Danish data ('concurrent adjustment') or at predetermined longer intervals, e.g. once a year ('adjustment with a forecasted seasonal component'). Concurrent adjustment is more computationally intense since reestimation of seasonal components should be undertaken as each new observation becomes available, and as such is therefore also an important consideration when deciding on the revision policy. The most common practice is to abstain from frequent revisions and to provide seasonally adjusted data based on forecasted seasonal components, especially whenever the seasonal components seem stable. Table 7.2 of the Appendix shows that the expected forecast error of the five adjusted series is small; the

---

[40] In earlier versions filters became more and more asymmetric towards the ends of the series, since symmetric filters were only used in the middle part of the series. By inclusion of each new variable the middle part is enlarged, resulting in more observations earlier being subject to symmetric filters. Thus, the new version of X-12 with the ARIMA modelling is a clear improvement to the quality of seasonal adjustment, cf. Deutsche Bundesbank (1999).

smallest is for the money stock measure *M2*, 1.25 per cent, and the largest is for the money stock measure, *M1*, 3.11 per cent.

8.1. In practice

In order to reduce the frequency of revisions and in accordance with the bank's general revision policy of all financial statistics[41], seasonally adjusted series are constructed based on a forecasted seasonal component. Forecasts are constructed for a 12-month period ahead and should thus be revised once a year (in September). In between the annual revisions, the forecasted seasonal component is combined with new information available in order to produce the seasonally adjusted data each month. At the annual revision, the seasonal component and the seasonal adjustment method in general, i.e. the regARIMA modelling, including the outlier identification, are reconsidered and re-estimated. This procedure should be followed – unless there is specific knowledge about technical changes or other special factors which could significantly change the regARIMA model and the forecasted values for the seasonal component. In such cases it should be possible to update the component outside 'the revisions month'.

During the annual revisions the regARIMA models including the estimation of regression variables (day-of-the-week effects and outliers) should also be revised. A gradually changing sample size implies that the calculated seasonal component changes as well and, in some cases, it could result in the adoption of other filter lengths. If the time series for regARIMA modelling are becoming too long (i.e. longer than 15 years[42]), it is also possible to leave out historical information from the regression[43].

The mentioned revision policy for seasonally adjusted data conforms well to the practice of most other institutions and central banks. Specifically, the procedure undertaken corresponds to the practice of the Deutsche Bundesbank[44]. At the time of the revisions, it is planned to publish a short report explaining the revisions and possible changes to the adjustment procedure. Annual quality reports and revised data are to be published every year in September – starting in 2007.

---

41 See www.nationalbanken.dk/DNUK/Statistics.nsf/side/Revision_policy!OpenDocument

42 ARIMA methods are usually used to model time series of a maximum length of 15 years, since one cannot necessarily assume that the data generating process over a longer time span does not break down. The minimum length is 5 years and the ideal is 8 years for seasonal adjustment.

43 A possible procedure could be to decide on a fixed sample length, e.g. 10 years, and then as new information is added, or more appropriately once a year during the revision period, elimination of the initial year in the sample will take place.

44 ECB (2000), p. 17.

9. Summary

This paper has illustrated that a seasonal adjustment procedure can be set up in more than one way. Consequently, seasonally adjusted data cannot have the same official status as unadjusted data, and it is important that the chosen method for seasonal adjustment is explained and documented. It is also obvious that the specific application of the X-12-ARIMA package may change as experience accumulates.

10. References

Bell, W. R. and S. C. Hillmer (1984). Issues involved with the seasonal adjustment of economic time series. *Journal of Business and Economic Statistics,* 2 (4), pp. 98-127.

Bobbitt, L. and M. Otto (1990). Effects of forecasts on the revisions of concurrent seasonally adjusted data using the X-11 seasonal adjustment procedure. *U.S. Bureau of the Census*.

Burman, J. P. and M. C. Otto (1988). Outliers in time series. *Bureau of the census*, Statistical division report series.

Carlsen, M. and J. D. Riishøj (2006). Brug af kontanter i Danmark. Danmarks Nationalbank, *Working Papers*, WP 41/2006. (Only in Danish).

Chang, I., G. C. Tiao, and C. Chen (1988). Estimation of Time Series Parameters in the Presence of Outliers. *Technometrics*, 30 (2), pp. 193-204.

Chen, C. and L.-M. Liu (1993). Joint estimation of model parameters and outlier effects in time series. *Journal of the American Statistical Association*, 88(421), pp. 284-297.

Danmarks Statistik (2002). *Sæsonkorrigering.* (Only in Danish).

Deutsche Bundesbank (1999). The changeover from the seasonal adjustment method Census X-11 to Census X-12-ARIMA. *Monthly Report,* Deutsche Bundesbank, 51(9), September.

European Central Bank (2000). *Seasonal adjustment of monetary aggregates and HICP for the euro area*. August 2000.

Findley, D. F., B. C. Monsell, W. R. Bell, M. C. Otto, and B.-C. Chen (1998). New Capabilities and Methods of the X-12-ARIMA Seasonal Adjustment Program. *Journal of Business and Economic Statistics,* 16 (2), pp. 127-176.

Findley, D. F., B. C. Monsell, H. B. Shulman and M. G. Pugh (1990). Sliding spans diagnostics for seasonal and related adjustments. *Journal of the American Statistical Association* 75, pp. 345-355.

Ladiray, D. and Quenneville, B. (1999). Understanding the X11 Method: The Various Tables. *Working paper series*, Time series research and analysis centre, Statistics Canada.

Lothian, J. and Morry, M. (1978). A Set of Quality Control Statistics for the X-11-ARIMA Seasonal Adjustment Method. Can be downloaded from www.census.gov/ts/papers/LothianMorry1978.pdf

Manna, M. and R. Peronaci (2003). *Seasonal Adjustment.* European Central Bank, January 2003.

McDonald-Johnson, K. M., and C. C. Hood (2001). Outlier selection for RegARIMA Models. U.S. Census Bureau. Can be downloaded from www.census.gov/pub/ts/papers/asa2001kmm.pdf .

Murdoch, K., V. Wilcox, G. Kuzmicich, and F. Krsinich (2000). The effect of forecasting on revisions to seasonally adjusted time series. *Working paper series*, Statistics New Zealand.

Nielsen, H. B. (1999). En teoretisk og empirisk sammenligning af metoder til sæsonkorrektion. Danmarks Nationalbank, *unpublished working paper.*

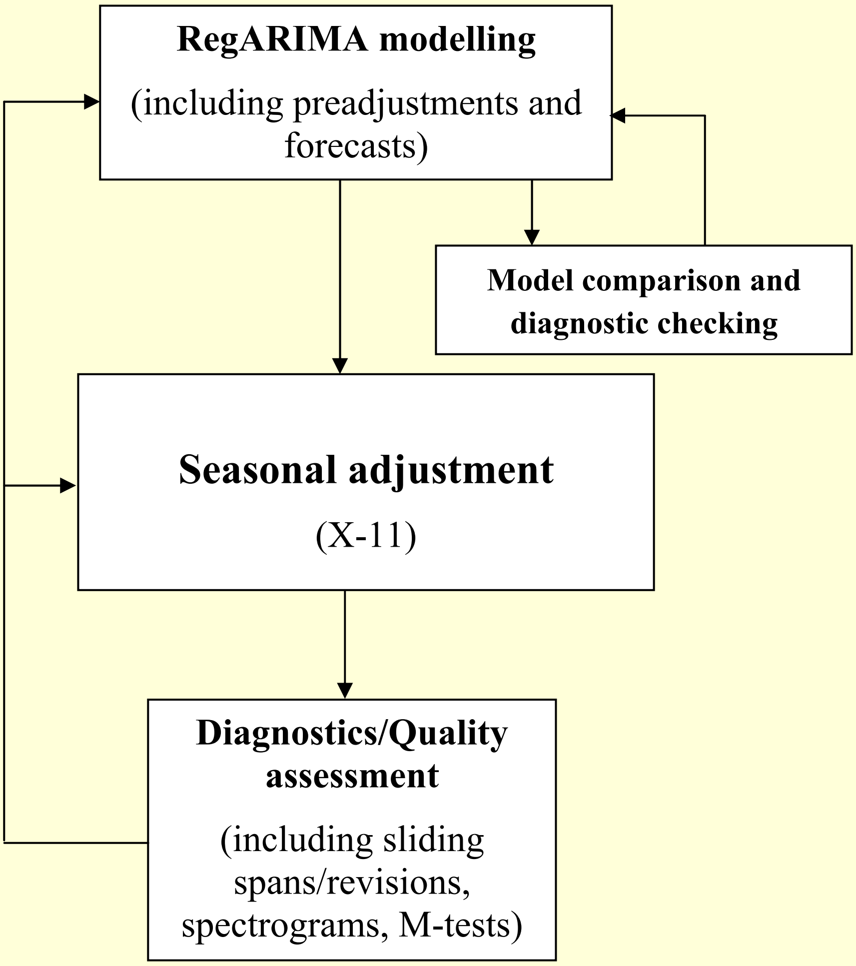
Pedersen, M. K. (2006). Seasonally adjusted financial statistics. Danmarks Nationalbank, *Monetary Review,* 3rd Quarter, pp. 95-103.

Pierce, D. A. and S. K. McKenzie (1985). On concurrent seasonal adjustment. *Bureau of the Census, statistical research division report series,* July 1985, no. CENSUS/SRD/RD-85/09.

U.S. Census Bureau, Time Series Staff (2006). *X-12-ARIMA Reference Manual.* Version 0.3, U.S. Census Bureau. Can be downloaded from http://www.census.gov/srd/www/x12a/.

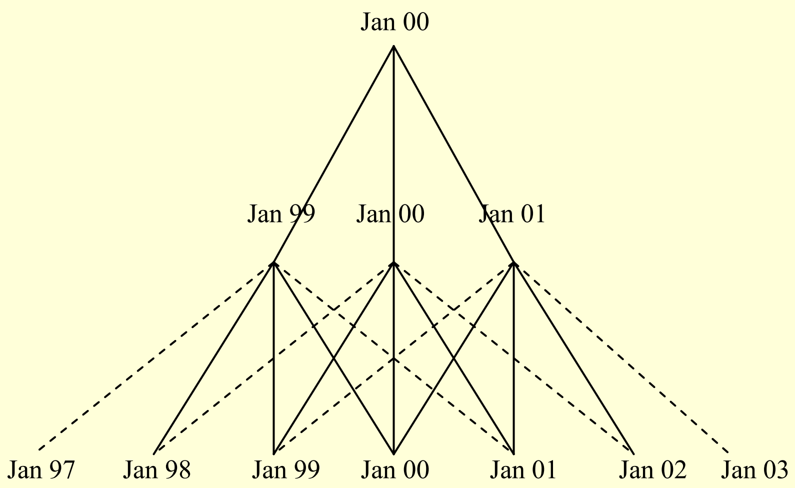
Appendix

THE X-12-ARIMA PROCEDURE Chart 3.1



Note: The figure is meant to illustrate the X-12-ARIMA procedure as presented in sections 3 and 7. The different steps are carried out interchangeably and there is a need for running through the process several times.

ILLUSTRATION OF A 3X3 AND A 3X5 FILTER FOR JANUARY 2000 Chart 3.2



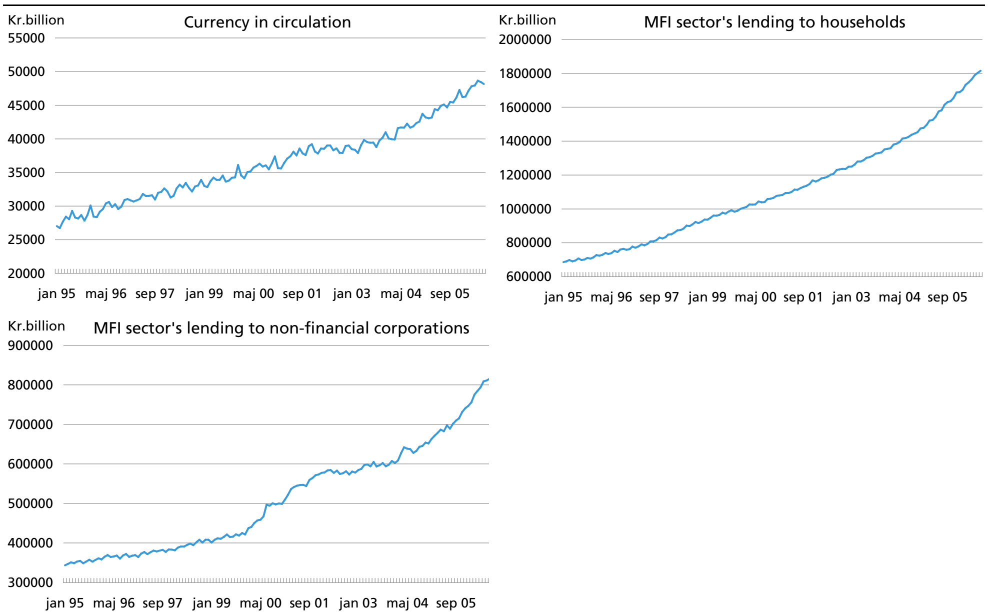
Note: The 3x3 filter is illustrated by the solid lines only, while the 3x5 filter is illustrated by both the solid and broken lines.

Chart 3.2 illustrates a 3x3 filter (the solid lines) which is a 3-term moving average of a 3-term moving average. In this example the final weights will be 1/9 (=1/3 * 1/3) for January 1998 and January 2002, 2/9 (=2 * 1/3 * 1/3) for January 1999 and January 2001 and 1/3 (=3 * 1/3 * 1/3) for January 2000.

The chart also contains an illustration of a 3x5 filter (the solid *and* broken lines), which is a 3-term moving average of a 5-term moving average. The final weights of the 3x5 filter will be 1/15 (= 1/3 * 1/5) for January 1997 and January 2003, 2/15 (= 2* 1/3 * 1/5) for January 1998 and January 2002 and 1/5 (= 3 * 1/3 * 1/5) for January 1999, January 2000 and January 2001.

In step 1 of section 3.2, a 12-terms centred moving average is used. A 12-term centred moving average is really a 2-terms moving average of a 12-term moving average.

THE DEVELOPMENT OF THE *CURRENCY IN CIRCULATION*, *THE MFI SECTOR'S LENDING TO HOUSEHOLDS* AND TO *NON-FINANCIAL COIRPORATIONS* FOR THE PERIOD JANUARY 1995 TO AUGUST 2006 — Chart 6.1



Source: Danmarks Nationalbank, the balance and flow statistics for the MFI sector.

PREADJUSTMENT OF THE DANISH DATA – A DOCUMENTATION OF INCLUDED OUTLIERS AND DAY-OF-THE-WEEK EFFECTS — Table 7.1

	Currency in circulation	**M1**	**M2**	**MFI sector's lending to households**	**MFI sector's lending to non-financial corporations**
Day-of-the-week effect	Monday [*0.01*] Tuesday [*-5.56*] Wednesday [-2.*37*] Thursday [*-0.27*] Chi^2 (df=4) =189.9 [P=0%] *Jointly significant*				
Additive outliers (AO)	2001:8 [*3.95*] 2003:12 [*4.22*]	2006:1 [*8.17*]	2006:1 [*5.66*]	2000:8 [*-3.62*] 2001:12 [*4.77*] 2002:8 [*4.73*] 2005:8 [*4.19*]	2000:7 [*1.61*] 2003:12 [4.87] 2004:4 [*2.83*]
Level shifts (LS)		2005:4 [*2.13*]	2005:6 [*4.12*]		

Note: In the above table the preadjustment of the five investigated time series are presented by the observations affected. Corresponding *t*-values (in brackets) are also presented to give an indication of the significance of the regressors included.

Descriptive statistics related to the regARIMA part in section 7.1
The Ljung-Box portmanteau statistics (Q-statistics) testing for no autocorrelation is given by:

(A.7.1) $Q_{LB} = n(n+2)\sum_{i=1}^{k} \frac{\rho_i^2}{n-i}$, t=1,2,…n,

where ρ_i is the estimated residual autocorrelation for lag *i* given by:

$$\rho_i = \frac{\mathrm{cov}(e_t e_{t+i})}{\sqrt{\mathrm{var}(e_t)\,\mathrm{var}(e_{t+i})}}$$

n is the number of observations and *k* is the number of lags included in the estimation. The test statistics are $\chi^2(n-k)$-distributed with degrees of freedom (*df*) equal to the difference between the number of observations and the lags included.

The formulas for the information criteria are given by:

(A.7.2) $AIC = -2L_N + 2n_p$
(A.7.3) $HannanQuinn_N = -2L_N + 2n_p \ln\ln N$
(A.7.4) $BIC_N = -2L_N + n_p \ln N$

where L_N is the estimated maximum loglikelihood value for a time series with *N* observations after differencing and n_p is the number of estimated parameters in the model. The information criteria should only be used to compare models with the same outliers included as well as the same differencing order.

REGARIMA MODEL STATISTICS, EVALUATING TESTS AND SLIDING SPANS DESCRIPTIVE STATISTICS USED FOR MODEL SELECTION AND QUALITY ASSESSMENT FOR THE FIVE DANISH FINANCIAL TIME SERIES

Table 7.2

	Currency in circulation	**M1**	**M2**	**MFI sector's lending to households**	**MFI sector's lending to non-financial corporations**
Model	$(0\ 1\ 1)_{12}(0\ 1\ 1)$	$(0\ 1\ 3)_{12}(0\ 1\ 1)$	$(0\ 1\ 1)_{12}\ (0\ 1\ 1)$	$(0\ 1\ 3)_{12}(0\ 1\ 1)$	$(0\ 1\ 2)_{12}(0\ 1\ 1\ )$
Number of estimated parameters	6	7	5	9	7
Number of observations	61	61	61	61	61
RegARIMA-related statistics					
Adjusted loglikelihood	-428.8	-629.1	-652.0	-600.1	-609.2
AIC	875.6	1272.1	1314.0	1219.3	1232.4
Hannan-Quinn	883.1	1277.9	1318.1	1222.8	1238.2
BIC	894.6	1286.9	1324.5	1238.3	1247.2
Residual normality and autocorrelations					
Kurtosis	2.319	2.830	3.034	2.571	2.336
Autocorrelation based on Q-statistics	No sign of residual autocorrelation (except at lag 5)	No sign of residual auto-correlation	No sign of residual auto-correlation	No sign of residual auto-correlation	No sign of residual auto-correlation
Forecasting Performance					
Average absolute percentage error in out-of-sample forecasts (last 3 years)	1.68	3.11	1.25	1.37	1.93
Introductory tests					
F_S-test	56.9 [*P<0.1%*]	32.4 [*P<0.1%*]	28.5 [*P<0.1%*]	30.9 [*P<0.1%*]	15.3 [*P<0.1%*]
F_M-test	2.4 [*P<5%*]*	1.8 [*P>5%*]	1.4 [*P>5%*]	6.8 [*P<1%*]*	0.5 [*P>5%*]
M7	0.353	0.437	0.447	0.667	0.524
Evaluating tests					
F_S-test (residual seasonality)	0.17 [*P>10%*]	0.58 [*P>10%*]	0.47 [*P>10%*]	0.10 [*P>10%*]	0.10 [*P>10%*]
M8	0.669	0.715	0.700	0.805	0.928
M9	0.456	0.575	0.633	0.674	0.681
M10	0.811	0.787	0.741	0.945	1.093*
M11	0.746	0.727	0.741	0.926	1.036*
Sliding Spans					
S(%)	0% (0/104)	2.9% (3/104)	6.7% (7/104)	0% (0/104)	0 % (0/104)
MM(%)	1.0% (1/103)	10.7% (11/103)	17.5% (18/103)	0% (0/103)	1.9% (2/103)
YY(%)	0% (0/92)	0% (0/92)	0% (0/92)	0% (0/92)	0 % (0/92)

Note.: For the introductory and evaluating tests, the values in brackets indicate the p-values. A * means that the test has failed. For the sliding spans descriptive statistics, the numbers in brackets are the number of observations flagged by X-12-ARIMA (following the criteria presented in section 7.3) out of total number of observations used for the span in question.

Source: Authors' own calculations – included in the output files generated by X-12-ARIMA.