*Terje Skjerpen*

# Seasonal Adjustment of First Time Registered New Passenger Cars in Norway by Structural Time Series Analysis

*Terje Skjerpen*

# Seasonal Adjustment of First Time Registered New Passenger Cars in Norway by Structural Time Series Analysis

**Statistisk sentralbyrå • Statistics Norway**
Oslo–Kongsvinger 1995

| Standardtegn i tabeller | Symbols in tables | Symbol |
|---|---|---|
| Tall kan ikke forekomme | Category not applicable | . |
| Oppgave mangler | Data not available | .. |
| Oppgave mangler foreløpig | Data not yet available | ... |
| Tall kan ikke offentliggjøres | Not for publication | : |
| Null | Nil | - |
| Mindre enn 0,5 av den brukte enheten | Less than 0.5 of unit employed | 0 |
| Mindre enn 0,05 av den brukte enheten | Less than 0.5 of unit employed | 0,0 |
| Foreløpige tall | Provisional or preliminary figure | * |
| Brudd i den loddrette serien | Break in the homogeneity of a vertical series | — |
| Brudd i den vannrette serien | Break in the homogeneity of a horizontal series | \| |
| Rettet siden forrige utgave | Revised since the previous issue | r |

ISBN 82-537-4200-2
ISSN 0806-2056

**Emnegruppe**
59: Andre samfunnsøkonomiske emner
Ny emnegruppe 1995: 00.90 metoder, modeller, dokumentasjon

**Emneord**
Sesongjustering
Tilstandsmodeller
X11-ARIMA

Design: Enzo Finger Design

# Abstract

*Terje Skjerpen*

**Seasonal Adjustment of First Time Registered New Passenger Cars in Norway by Structural Time Series Analysis**

**Reports 95/30 • Statistics Norway 1995**

Within the framework of structural time series models it is shown how monthly unadjusted data can be seasonally adjusted. It is assumed that a time series or the log of it is the sum of three unobserved components corresponding to trend, seasonality and irregularity. For each of these components we assume explicit stochastic processes. Utilizing the State Space Form and Kalman filter techniques the unobserved components can be estimated and the observed time series can be corrected for seasonality. With respect to the seasonal component we comment on two different stochastic specifications which coincide in the generating deterministic case.

To illustrate the application of the structural time series approach to seasonal adjustment, we utilize time series for first time registered passenger cars in Norway from 1973 to 1994. We are also seasonally adjusting this time series using X11-ARIMA, and we apply some practical criteria in order to compare the decompositions obtained from the structural time series models and X11-ARIMA.

**Keywords**: Unobserved components, State Space Form, seasonal adjustment, X11-ARIMA.

# Innhold

# 1. Introduction[1]

Statistics Norway seasonally adjusts a large number of both monthly and quarterly time series on a regular basis. The seasonal adjustment takes place in different departments of the institution , but they all use the same computer program, namely X11-ARIMA (cf Dagum(1988)). X11-ARIMA is well suited for a situation where a lot of time series have to be seasonally adjusted (and smoothed) fastly. This is due to the high level of automatization in the program.

However, in the time series literature competing methods for seasonal adjustment have been available for a long time. In the influential book of Box and Jenkins (1970) the class of seasonal ARIMA-models was introduced in order to model univariate seasonally unadjusted data. Ordinary and seasonal differencing was undertaken in order to model stationary time series data. On the basis of the work of Box and Jenkins two different but strongly related traditions have evolved. In the first of this which one can label a reduced form approach, one identifies an adequate ARIMA-model and for the chosen model one uses frequency domain methods to identify the different time series components. From the results in the frequency domain one can make optimal filters in the time domain which mirror the results obtained from the spectral decomposition. The resulting filters can be interpreted as pass-filters. For instance the trend filter is a low-pass filter which allows only cycles with long periodicity to pass. This methodological approach is pursued by Maravall (1988) and is also the basis for the computer program SEATS (cf Maravall and Gómez (1994)) which has been launched as a competitor to X11-ARIMA.

The other tradition is the structural time series model approach which was orginated by the work of Engle (1978). In the structural the time series model one assumes explicit ARIMA-processes for the unobserved components. From the underlying assumptions it is possible to deduce a reduced form ARIMA-representation in which however the parameters are constrained since they are function of a lower dimensional vector of the parameters in the structural model. After having estimated the structural parameters the seasonal component can be extracted from the observed time series and a resulting seasonally adjusted time series are obtained.

The advantage of all these time series methods is that the assumed data generating process for a time series is written down explicitly. This facilitates the interpretation of the decomposition of the time series since one in some way or another is forced to define for instance what is meant by the seasonal component. Besides they emphasize that there will always be some arbitrariness inherent in the seasonally adjusted data since several distinctive models can be set up for the purpose of seasonal adjustment. Thus one should not forget that secondary data has another status than original data.

X11-ARIMA does not produce any uncertainty measures connected to the estimates of unobserved components such as the seasonally adjusted values. Even if Pfefferman (1993) and Pfefferman, Morry and Wong (1995) have

[1] The author wishes to thank Olav Bjerkholt, Ådne Cappelen and Anders Rygh Swensen for comments and Tone Veiby for the final typing of the report.

given formulas for the variances of estimated components from X11 and X11-ARIMA under very weak conditions, it may still be more straighforward to estimate the uncertainty from parametric models.

In this paper we will be concerned with the structural time series models approach to seasonal adjustment. The computer program STAMP[2], which methodological basis is the book by Harvey (1989), will be used for the numerical calculations. Since STAMP is a menu-driven program it should not be viewed as an alternative to X11-ARIMA when it comes to production of seasonally adjusted data. However, it will frequently be necessary to evaluate the options choosen for seasonal adjustment of specific time series by X11-ARIMA, and as natural part of this process it seems worthwhile to utilize the conclusions which can be drawn from univariate time series modelling.

In order to illustrate the structural time series approach to seasonal adjustment the monthly time series for the number of first time registered new passenger cars in Norway is used. There are at least two reasons for choosing just this economic indicator. First this is an economic variable which receives a lot of attention by economic agents. It is by many viewed as a leading indicator. An upswing in this variable is often believed to be followed by better prospects for the economy in general. However, the time series contain seasonal and irregular components which have to be removed in order to isolate the trend. Second this economic indicator has been subject to time series analysis in which intervention techniques have been combined with X11-ARIMA. Intervention analysis can be used to correct a time series for abrupt changes which may or may not have a substantial explanation. Within a seasonal adjustment context abrupt changes may contaminate the seasonal factors. Pham (1995) preadjusted the number of first time registered new personal cars by using the intervention technique and applied X11-ARIMA to the corrected time series. The intervention technique was first applied to economic data by Box and Tiao (1975) and different variants of it have also been implemented in several computer programs. Also within a structural time series approach one can correct for intervention effects. This has not been done in the current report, but may be a topic for further reasearch.

The rest of the report is organized as follows: In chapter 1 we describe the different structural time series model which are used in this study. Some fundamental properties of these models are deduced in appendices A and B. In section 2 we show how these models can be presented in the State Space Form, whereas we in section 3 dwell upon estimation, diagnostic testing and smoothing. Section 4 is devoted to the empirical results based on the structural time series models. In section 5 we give a very short description of the X11-ARIMA method, and in section 6 we compare the seasonal adjustment of the number of first time registered passenger cars by use of structural time series models and X11-ARIMA. The conclusions are drawn in section 7. The data are given in appendix C.

[2] A new version of STAMP called STAMP 5.0, cf Koopman, Harvey, Doornik and Shephard (1995), is now available. The calculations in this paper are based on STAMP 3.0.

# 2. Modelling framework

In this paper we focus on the following decomposition of a monthly time series, $y_t$:

(1.1) $y_t = \mu_t + \gamma_t + \varepsilon_t ,$

where $\mu_t$ is taken to be an unobserved trend component which in the general case follows a random walk with stochastic drift:

(1.2) $\mu_t = \mu_{t-1} + \beta_{t-1} + u_t$, where

(1.3) $\beta_t = \beta_{t-1} + \varepsilon_t .$

The variable $\gamma_t$ is an unobserved component which represents seasonality. In this paper we explore two stochastic seasonal processes which coincide in the degenerating deterministic case. A detailed treatment of these processes will be given in the next section. Finally, the variable $\varepsilon_t$ is an irregular component which is assumed to have white noise properties.

## 2.1 The seasonal component

For the seasonal component we distinguish between two stochastic specifications which under a degenerating deterministic case coincide. These two seasonal models are termed the dummy seasonal and the harmonic seasonal specification, respectively. In the dummy seasonal case the seasonal component is assumed to develop according to the following stochastic process:

(1.4) $\sum_{i=0}^{11} \gamma_{t-i} = w_t .$

For the measurement error in (1.1) and the disturbances in (1.2)-(1.4) we make the following stochastic assumptions:

(1.5) $(\varepsilon_t, \mu_t, v_t, w_t)' \sim \text{NID}(0_{4\times1}, \text{DIAG}[\sigma^2_{\varepsilon\varepsilon}, \sigma^2_{uu}, \sigma^2_{vv}, \sigma^2_{ww}]),$

where NID stands for normally and independently distributed.

Both the specification of the trend and the seasonal component implies non-stationarity of the observed variable. However a stationary specification can be easily obtained after some appropriate differencing. In appendix A we deduce the socalled stationary form of the basic structural model with the dummy seasonal specification (cf equation (A.10)). This implies that some transformed (by differencing) value of $y_t$ is a linear function of current and lagged values of the measurement error and of the disturbances in the processes of the unobserved components.

Besides the above specification of the seasonal component, we also explore the consequences of using an alternative specification based on trigonometric functions. In this model the seasonal component, $\gamma_t$, is given as a sum of stochastic seasonal cycles corresponding to the fundamental seasonal frequency and the companion harmonics. (For the properties of trigonometric functions see table 1.) Thus, formally we have the following model for the seasonal component:

$$(1.6) \qquad \gamma_t = \sum_{i=1}^{6} \gamma_{i,t}.$$

For $1 \leq i \leq 5$, $\gamma_{i,t}$ is defined implicitly by the following two equations:

$$(1.7a) \qquad \gamma_{i,t} = \cos(\lambda_i)\gamma_{i,t-1} + \sin(\lambda_i)\gamma^*_{i,t-1} + \omega_{i,t}, \text{ and}$$

$$(1.7b) \qquad \gamma^*_{i,t} = -\sin(\lambda_i)\gamma_{i,t-1} + \cos(\lambda_i)\gamma^*_{i,t-1} + \omega^*_{i,t}, \text{ where } \lambda_i = \frac{2\pi i}{12}.$$

The $\gamma^*_{i,t}$'s are auxilliary variables facilitating a state space formulation of the structural time series model. For i=6 we have, since $\sin(\lambda_6) = \sin(\pi) = 0$, a simpler specification given by:

$$(1.8) \qquad \gamma_{6,t} = -\gamma_{6,t-1} + \omega_{6,t}.$$

Let $\omega_t = [\omega_{1,t}, \omega^*_{1,t}, \omega_{2,t}, \omega^*_{2,t}, \ldots, \omega_{5,t}, \omega^*_{5,t}, \omega_{6,t}]'$ be an 11×1 vector containing the disturbance terms at period t. Then the following stochastic assumptions are made:

$$(1.9) \qquad \omega_t \sim N(0_{11\times 1}, \sigma^2_{\omega\omega} I_{11\times 11}), \text{ and}$$

$$(1.10) \qquad E\omega_t\omega_s' = 0_{11\times 11}.$$

Equation (1.9) imposes the same variance for all the disturbances. This is not required for the identification of this structural time series model, but it is an assumption which to a large extent simplifies the numerical calculations associated with the estimation of the hyperparameters. In appendix B we elaborate more on this structural time series model. It is shown that we can apply the same filter on $y_t$ as in the dummy seasonal case in order to obtain stationarity. As for the dummy seasonal case the autocovariance function of the transformed variable only depends on 4 hyperparameters. The difference between the two seasonal specifications comes from the different ways they structure the autocovariance function.

**Table 1. Properties of trigonometric functions**

| Frequency (λ) | Cos (λ) | Sin (λ) |
|---|---|---|
| $\pi/6$ | $1/2\sqrt{3}$ | 1/2 |
| $\pi/3$ | 1/2 | $1/2\sqrt{3}$ |
| $\pi/2$ | 0 | 1 |
| $2\pi/3$ | -1/2 | $1/2\sqrt{3}$ |
| $5\pi/6$ | $-1/2\sqrt{3}$ | 1/2 |
| $\pi$ | -1 | 0 |

# 3. Structural time series model in the State Space Form

We want to estimate the hyperparameters of the structural time series models and we also want to extract the components using all the sample information. Besides we want to assess how successful the decomposition has been. It turns out that the State Space Form is useful for all three purposes. We start up this section with some general remarks about the State Space form and turn later to how our structural time series models can be formulated within this context. It should be pointed out that we do not write out the State Space Form in its most general form since this is not necessary facing our concrete application. In the State Space Form with a single measurement one distinguishes between the measurement equation (2.1) and the transition equations (2.2) respectively:

(2.1) $y_t = A\Gamma_t + \phi_t$ and

(2.2) $\Gamma_t = T\Gamma_{t-1} + R\zeta_t$.

In (2.1) $y_t$ (a scalar) is the value of the observed variable or some simple transformation of it in period t. The observed variable in period t is a linear function of the unobserved state vector in the same period, $\Gamma_t$. Let this vector have the dimension $m \times 1$. The system vector A is of dimension $1 \times m$. In our application this matrix is known. Finally the scalar $\phi_t$ represents measurement noise in period t. The development of the state vector is governed by the transition equations in equation (2.2). The $m \times m$ matrix T which is termed the transition matrix is like the vector A known in our application. R is a time invariant and known m x q matrix (where q is less than or equal to m) and $\zeta_t$ is an q x 1 vector of disturbance terms in the transition equations. The following stochastic assumptions are made:

(2.3) $\phi_t \sim NID(0, \sigma^2_{\phi\phi})$; t = 1,..., T,

(2.4) $\zeta_t \sim NID(0_{q\times 1}, \Omega)$; t = 1,...,T,

(2.5) $E(\zeta_t\phi_s) = 0_{q\times 1}$; t ,s = 1,...,T,

(2.6) $\Gamma_0 \sim N(E(\Gamma_0), E(\Gamma_0\Gamma_0'))$,

(2.7) $E(\Gamma_0\phi_t) = 0$; t=1,...,T, and

(2.8) $E(\Gamma_0\zeta_t') = 0_{m\times q}$; t = 1,..., T.

In (2.3) $\sigma^2_{\phi\phi}$ is the variance of the noise in the measurement equation. The matrix $\Omega$, which is of dimension q x q, is the covariance matrix of the disturbance terms in the transition equations. In our empirical application this will be a diagonal matrix. The assumption in (2.5) means that the noise in the measurement error is stochastic independent of the errors in the transition equations. In (2.6) - (2.8) we state our assumptions with regard to the initial conditions. The initial state vector is assumed to be distributed independently of the noise in the measurement equation and of the disturbances in the transition equations. Since we only have nonstationary elements in the state vector some care must be taken in defining the initial conditions.

Let $h_{t-1}$ denote the optimal estimator of $\Gamma_{t-1}$ using observed information up to and including period t-1, and let $P_{t-1}$ be the covariance matrix of the estimation error:

(2.9) $$P_{t-1} = E[(h_{t-1} - \Gamma_{t-1})(h_{t-1} - \Gamma_{t-1})'].$$

Given $h_{t-1}$ and $P_{t-1}$ the the optimal estimator of $\Gamma_t$ is

(2.10) $$h_{t|t-1} = Th_{t-1}, \text{with covariance matrix}$$

(2.11) $$P_{t|t-1} = TP_{t-1}T' + R\Omega R'.$$

Equations (2.10) and (2.11) are known as the prediction equations. When a new observation of y becomes available the estimator of $h_t$ and its covariance matrix $P_t$ will be updated according to:

(2.12) $$h_t = h_{t|t-1} + P_{t|t-1}A'\left(\frac{1}{f_t}\right)(y_t - Ah_{t|t-1}) \text{ and}$$

(2.13) $$P_t = P_{t|t-1} - P_{t|t-1}A'\left(\frac{1}{f_t}\right)AP_{t|t-1},$$

where the scalar $f_t$ is defined as:

(2.14) $$f_t = AP_{t|t-1}A' + \sigma^2_{\phi\phi}.$$

We are now prepared to recast the two structural time series models in the State Space Form. In the dummy seasonal case the state vector is given as:

(2.15) $$\Gamma_t = (\mu_t, \beta_t, \gamma_t, \gamma_{t-1}, \ldots, \gamma_{t-10})'.$$

The transition matrix T is given by the following 13 × 13 matrix:

$$(2.16)\quad T=\begin{bmatrix}1&1&0&0&0&0&0&0&0&0&0&0&0\\0&1&0&0&0&0&0&0&0&0&0&0&0\\0&0&-1&-1&-1&-1&-1&-1&-1&-1&-1&-1&-1\\0&0&1&0&0&0&0&0&0&0&0&0&0\\0&0&0&1&0&0&0&0&0&0&0&0&0\\0&0&0&0&1&0&0&0&0&0&0&0&0\\0&0&0&0&0&1&0&0&0&0&0&0&0\\0&0&0&0&0&0&1&0&0&0&0&0&0\\0&0&0&0&0&0&0&1&0&0&0&0&0\\0&0&0&0&0&0&0&0&1&0&0&0&0\\0&0&0&0&0&0&0&0&0&1&0&0&0\\0&0&0&0&0&0&0&0&0&0&1&0&0\\0&0&0&0&0&0&0&0&0&0&0&1&0\end{bmatrix}.$$

The system matrices A and R are given by respectively:

$$(2.17)\quad A=\begin{bmatrix}1&0&1&0&0&0&0&0&0&0&0&0&0\end{bmatrix}, \text{and}$$

$$(2.18)\quad R=\begin{bmatrix}I_{3\times 3}\\0_{10\times 3}\end{bmatrix}.$$

Furthermore we have that $\sigma^2_{\phi\phi}=\sigma^2_{\varepsilon\varepsilon}$ and $\zeta_t=(u_t, v_t, w_t)'$ with covariance matrix

$$(2.19)\quad \Omega=\text{DIAG}\left[\sigma^2_{uu},\sigma^2_{vv},\sigma^2_{ww}\right]$$

This completes the specification of the dummy seasonal model in the state space form. In the model with harmonic seasonality the state vector is given by:

$$(2.20)\quad \Gamma_t=\begin{bmatrix}\mu_t & \beta_t & \gamma_{1,t} & \gamma^*_{1,t} & \gamma_{2,t} & \gamma^*_{2,t} & \gamma_{3,t} & \gamma^*_{3,t} & \gamma_{4,t} & \gamma^*_{4,t} & \gamma_{5,t} & \gamma^*_{5,t} & \gamma_{6,t}\end{bmatrix}'.$$

The transition matrix T is given by:

$$(2.21)\quad T=\begin{bmatrix}
1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & \frac{1}{2}\sqrt{3} & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & -\frac{1}{2} & \frac{1}{2}\sqrt{3} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2}\sqrt{3} & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & -\frac{1}{2\sqrt{3}} & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\frac{1}{2} & \frac{1}{2}\sqrt{3} & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\frac{1}{2\sqrt{3}} & -\frac{1}{2} & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\frac{1}{2}\sqrt{3} & \frac{1}{2} & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\frac{1}{2} & -\frac{1}{2}\sqrt{3} & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1
\end{bmatrix}.$$

In the structural time series model with harmonic seasonality the matrix R is the 13 × 13 identity matrix, whereas the system vector A is given by

$$(2.22)\quad A=\begin{bmatrix}1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1\end{bmatrix}.$$

For the variance of the measurement equation we have that $\sigma^2_{\phi\phi}=\sigma^2_{vv}$. The vector of the disturbances in the transition equations is given by:

$$(2.23)\quad \zeta_t=\begin{bmatrix}u_t & v_t & \omega_{1,t} & \omega^*_{1,t} & \omega_{2,t} & \omega^*_{2,t} & \omega_{3,t} & \omega^*_{3,t} & \omega_{4,t} & \omega^*_{4,t} & \omega_{5,t} & \omega^*_{5,t} & \omega_{6,t}\end{bmatrix}'.$$

The covariance matrix of this vector is:

$$(2.24)\quad \Omega=\text{DIAG}\left[\sigma^2_{uu},\sigma^2_{vv},\sigma^2_{\omega\omega},\sigma^2_{\omega\omega},\sigma^2_{\omega\omega},\sigma^2_{\omega\omega},\sigma^2_{\omega\omega},\sigma^2_{\omega\omega},\sigma^2_{\omega\omega},\sigma^2_{\omega\omega},\sigma^2_{\omega\omega},\sigma^2_{\omega\omega},\sigma^2_{\omega\omega}\right]$$

# 4. Estimation, testing and smoothing

Estimation of the hyperparameters can be done both in the time and the the frequency domain. The log-likelihood formulated in the frequency domain can be viewed as an approximation to the true log-likelihood which works well when none of the hyperparameters are close to zero. However in practical applications borderline solutions is rather the rule than the exception. Because of this we use the algorithm in STAMP which is based on the time domain formulation of the log-likelihood.

Let the (theoretical) one-step-ahead prediction errors be defined as:

(3.1) $\varsigma_t = y_t - y_{t|t-1}$.

In equation (3.1) $y_{t|t-1}$ denotes the optimal predictor of $y_t$ given past information. This predictor may also be written as:

(3.2) $y_{t|t-1} = ATh_{t-1}$.

The variance of the one-step-ahead prediction errors is given by:

(3.3) $f_t = ATP_{t-1}T'A' + AR\Omega R'A' + \sigma^2_{\phi\phi}$.

Note that $\varsigma_t$ and $f_t$ is a function of the hyperparameters of the structural time series model. We are now able to express the log likelihood function in the the prediction decomposition form. We assume that the initial state vector has been initialized by a diffuse prior. This means that the initial state vector can be taken to have expectation zero and a scalar covariance matrix in which the scalar is large but finite. In accordance with this assumption maximum likelihood estimation is done conditional on the first d observations. The number d equals the number of non-stationary components in the state vector, which in our application turns out to be 13. The log-likelihood function may now be written as:

(3.4) $\ln(L) = -\frac{(T-d)}{2}\ln(2\pi) - \frac{1}{2}\sum_{t=d+1}^{T}\ln(f_t) - \frac{1}{2}\sum_{t=d+1}^{T}\frac{\varsigma_t^2}{f_t}$.

However, it is possible to simplify the numerical problem further by reparameterizing all the variances in the set of hyperparameters in such a way that they share a common scalar. The ususal normalization is to let this scalar coincide with the variance of the measurement noise. This means that this scalar can be concentrated out of the log-likelihood functions and that maximimum likelihood estimation can proceed by maximizing a lower dimensional concentrated log-likehood. For this procedure to work it is however essential that the variance of the measurement error is positive. If it is not, the common scalar must coincide with one of the remaining variances which is positive.

Let a symbol with an $\wedge$ indicate a realized value based on maximum-likelihood estimation of the hyperparameters. The standardized innovations which can be utilized in the definitions of several test diagnostics are defined as:

$$(3.5) \qquad I_t = \frac{\hat{\varsigma}_t}{\hat{f}_t^{0.5}}.$$

The standardized innovations should be free from autocorrelations. This can be tested by a Box-Ljung portmanteau statistic suited for structural time series models. Let the autocorrelation at distance $\tau$ of the standardized innovations be defined as:

$$(3.6) \qquad r_I(\tau) = \frac{\sum_{t=d+1+\tau}^{T} (I_t - \bar{I})(I_{t-\tau} - \bar{I})}{\sum_{t=d+1}^{T} (I_t - \bar{I})^2}.$$

The Box-Ljung statistic of the significance of the first P autocorrelations is then given by:

$$(3.7) \qquad Q = (T-d)(T-d+2)\sum_{\tau=1}^{P} (T-d-\tau)^{-1} r_I^2(\tau).$$

Under the null of no autocorrelation Q is $\chi^2$-distributed with (P-(n-1)) degrees of freedom, where n is the number of parameters in the concentrated log-likelihood function. In the empirical application P will be set to one third of the (effective) sample size.

A simple test of heteroskedasticity is based on the following statistic:

$$(3.8) \qquad H(h) = \frac{\sum_{t=T-h+1}^{T} I_t^2}{\sum_{t=d+1}^{d+1+h} I_t^2}.$$

Under the null of no heteroscedasticity hH(h) is $\chi^2$-distributed with h degrees of freedom. In the empirical application h will be taken as one third of the (effective) sample size.

To test for normality in the standardized innovation we employ a Jarque-Bera type statistic. It can be shown that the maximum likelihood of the variance of the measurement noise, which is concentrated out of the log-likelihood function, is given by the following formula:

$$(3.9) \qquad \hat{\sigma}_{vv}^2 = \frac{1}{(T-d)} \sum_{t=d+1}^{T} I_t^2.$$

To set up the test statistic for normality let us first define the following expressions:

$$(3.10)\quad \sqrt{b_1} = \hat{\sigma}_{vv}^{-3} \sum_{t=d+1}^{T} \frac{(I_t - \bar{I})^3}{(T-d)}, \text{ and}$$

$$(3.11)\quad b_2 = \hat{\sigma}_{vv}^{-4} \sum_{t=d+1}^{T} \frac{(I_t - \bar{I})^4}{(T-d)}.$$

The test statistic for normality is now given by:

$$(3.12)\quad N = \frac{(T-d)}{6} b_1 + \frac{(T-d)}{24} (b_2 - 3)^2.$$

The first term on the right hand side of (3.12) takes care of skewness whereas the latter takes care of kurtosis. Under the null of normality N is $\chi^2$-distributed with two degrees of freedom.

Since our aim is to estimate the components of the state vector, it will be optimal to utilize all the sample information. This is known as smoothing and is conducted by running the Kalman-filter recursions backward. Several algorithms are available, but below we will only mention the fixed interval smoothing algorithm. Let generally $x_{t|T}$ mean that we are estimating a component at period t utilizing the sample information of all T periods and let us for short write $x_{T|T}$ as $x_T$. The fixed interval smoothing equations are given by:

$$(3.13)\quad h_{t|T} = h_t + P_t^*(h_{t+1|T} - Th_t), \text{ and}$$

$$(3.14)\quad P_{t|T} = P_t + P_t^*(P_{t+1|T} - P_{t+1|t}), \text{ where}$$

$$(3.15)\quad P_t^* = P_t T' P_{t+1|t}^{-1}, t = T-1, \ldots, 1.$$

For t=T-1 we have:

$$(3.13a)\quad h_{T-1|T} = h_{T-1} + P_{T-1}^*(h_T - Th_{T-1}),$$

$$(3.14a)\quad P_{T-1|T} = P_{T-1} + P_{T-1}^*(P_T - P_{T|T-1}), \text{ and}$$

$$(3.15a)\quad P_{T-1}^* = P_{T-1} T' P_{T|T-1}^{-1}.$$

Note that all terms on the right hand side can be obtained from the Kalman filter. Having calculated the left hand side of these equations one can proceed to calculate $h_{T-2|T}$, $P_{T-2|T}$ and $P_{T-2}^*$. The process continues until one has calculated $h_{1|T}$, $P_{1|T}$ and $P_1^*$.

# 5. Empirical results

Table 2 contains some estimation and test results for the structural time series models. The upper part of the table corresponds to the dummy seasonal specification, whereas the lower corresponds to the harmonic seasonal specification. The entire sample length is January 1973 to December 1994. Since we in the next section will be occupied by revision of the seasonal component table 2 also contains some estimation results based on subsamples. Our first subsample does only cover the period up to december of 1990. Then in the next subsample 12 new observations are added. This process continues until we reach the entire sample. In the right part of the table we have given som test diagnostics connected to the standardized innovations.

The most striking feature is that the drift of the trend and the seasonal component seem to be close to deterministic. In the last three subsamples the variance of the slope of the trend component $\sigma^2_{vv}$ is zero, i. e. at the border of the admissible parameter set. In the two first subsamples we have a small positive value. For the disturbance variable of the seasonal component it is the other way around. The estimation result for the first three subsamples supports a fixed seasonal pattern. For the two last subsamples there is some evidence of a changing seasonal pattern. These results are common for both representations of the seasonal component. The highest variance is obtained for the disturbance term of the trend component, which is somewhat higher than the estimated variance of the irregular component. Of course it would be more satisfactory to test for fixed seasonality on a formal basis and simultaneously assuming a deterministic drift in the trend component. This is however not straigthforward since it involves non-standard inference. It should be noted that if no evidence of stochastic seasonality is found in any of the the models, the models in the upper and lower panel degenerates to the same. There is no evidence of heteroscedasticity or non-normality in any of the models or subsamples. However there is some sign of serial correlation in the standardized innovations. In the full sample the significance probability of the Box-Ljung statistic is 2 per cent in the model with the dummy seasonal specification and 2.8 per cent in the model with the harmonic seasonal specification. These relative low significance probabilities are mainly due to a rather high autocorrelation of the standarized innovations at lag 12.

Since both the drift in the trend component and the seasonal process seems to be fixed it is of great interest to test the joint hypotheses that both the accompanion variances are zero. It is however not straightforward to use standard testing principles such as LR, WALD or LM. However, Franzini and Harvey (1983) have constructed a point optimal test which can be used in this case, which they term the partially deterministic case. Using a point optimal test both the null and the alternative hypotheses are fully specified. Harvey shows that the test statistic, which he denotes the $b^*$ - statistic, can be calculated as the ratio between the sum of squares of standardized one-step-ahead prediction errors corresponding to the partially deterministic case divided by an other sum of squares of standardized prediction errors corresponding to the model under the alternative hypoteses. The seasonal specification of the seasonal process under the alternative hypothesis is based on the dummy seasonal specification. The numerator is obtained by running the Kalman filter with $\sigma^2_{\varepsilon\varepsilon} = 0, \sigma^2_{uu} = 1$ and $\sigma^2_{vv} = \sigma^2_{ww} = q$ (where q is a fixed number which depends on the frequency of the data as well as the sample size) whereas the denominator is obtained by running the Kalman filter with $\sigma^2_{\varepsilon\varepsilon} = 0, \sigma^2_{uu} = 1$ and $\sigma^2_{vv} = \sigma^2_{ww} = 0$. The partially deterministic model is rejected if the ratio become to small. Critical values (corresponding to an optimal value of q) have been tabulated by Franzini and Hendry (1983) for quarterly data. However, in principle a corresponding test can also be constructed when facing monthly data. This implies fixing q to a new value and establishing new critical values.

**Table 2. Results from estimation of the structural time series models**[a]

| | | | | | Test diagnostics | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Last month | | | | | Serial correlation | | | Heteroscedasticity | | | | Non-normality | |
| in sample | $\sigma^2_{uu}$ | $\sigma^2_{vv}$ | $\sigma^2_{ww}$ | $\sigma^2_{\varepsilon\varepsilon}$ | Q | DF | P | hH(h) | DF | P | N | P | |
| *I. Dummy seasonal specification* | | | | | | | | | | | | | |
| 1990.12 | 6.1699 | 0.0002 | 0 | 4.6014 | 18.15 | 11 | 0.078 | 0.8587 | (67,67) | 0.733 | 2.5181 | 0.284 | |
| | (1.3894) | (0.0003) | | (1.0296) | | | | | | | | | |
| 1991.12 | 5.9365 | 0.0002 | 0 | 4.5092 | 18.92 | 11 | 0.063 | 0.8199 | (71,71) | 0.798 | 2.7144 | 0.257 | |
| | (1.3028) | (0.0004) | | (0.9724) | | | | | | | | | |
| 1992.12 | 5.6345 | 0 | 0 | 4.6750 | 23.53 | 12 | 0.024 | 0.7366 | (75,75) | 0.906 | 2.6062 | 0.272 | |
| | (1.2010) | | | (0.9386) | | | | | | | | | |
| 1993.12 | 5.7988 | 0 | 0.0002 | 4.6328 | 23.46 | 12 | 0.024 | 0.8158 | (79,79) | 0.816 | 1.7597 | 0.415 | |
| | (1.1977) | | (0.0123) | (0.9533) | | | | | | | | | |
| 1994.12 | 5.7130 | 0 | 0.0145 | 4.3586 | 24.02 | 12 | 0.020 | 0.7818 | (83,83) | 0.868 | 2.4040 | 0.301 | |
| | (1.1513) | | (0.0163) | (0.9044) | | | | | | | | | |
| *II. Harmonic seasonal specification* | | | | | | | | | | | | | |
| 1990.12 | 6.1697 | 0.0002 | 0 | 4.6015 | 18.15 | 11 | 0.078 | 0.8587 | (67,67) | 0.733 | 2.5184 | 0.284 | |
| | (1.3894) | (0.0003) | | (1.0296) | | | | | | | | | |
| 1991.12 | 5.9368 | 0.0002 | 0 | 4.5091 | 18.92 | 11 | 0.063 | 0.8199 | (71,71) | 0.798 | 2.7137 | 0.257 | |
| | (1.3029) | (0.0003) | | (0.9724) | | | | | | | | | |
| 1992.12 | 5.6304 | 0 | 0 | 4.6782 | 23.52 | 12 | 0.024 | 0.7365 | (75,75) | 0.906 | 2.6107 | 0.271 | |
| | (1.2004) | | | (0.9834) | | | | | | | | | |
| 1993.12 | 5.4872 | 0 | 0.0015 | 4.4797 | 22.20 | 12 | 0.035 | 0.7728 | (79,79) | 0.873 | 1.9671 | 0.374 | |
| | (1.1602) | | (0.0008) | (0.9364) | | | | | | | | | |
| 1994.12 | 5.3867 | 0 | 0.0018 | 4.2489 | 23.00 | 12 | 0.028 | 0.7440 | (83,83) | 0.910 | 2.8433 | 0.241 | |
| | (1.1071) | | (0.0009) | (0.8857) | | | | | | | | | |

[a] Standard deviations in parentheses. Parameters and standard deviations are multiplied by 1000.

# 6. Seasonal adjustment using X11- ARIMA

The number of registered new personal cars has also been seasonally adjusted using X11-ARIMA. The seasonal adjustment by X11-ARIMA has mainly been conducted in agreement with the official way Statistics Norway seasonally adjusts and publishes this time series. This means running the programme mainly with default options. Calendar effects, i. e. trading day effects and the effect of the position of easter, are allowed for to the extent they come out significantly. In principle it is rather easy to take account of such calender effects also within a structural time series framework (cf for instance Dagum, Quennville and Sudrathar (1992)). However, the estimation of the augmented models taking care of such effects could not be done because of too little workspace. In X11-ARIMA five ARIMA-models are automatically fitted to the time serie in question and properties of the respectively models are decisive of whether they are accepted for forecasting. One always starts out with the most parsimonious model, i. e. the so-called airline model. If this model is in accordance with the requirements it is chosen for forecasting, if not one proceeds to the next model. However, a situation in which no of the five models are accepted is not unusual. For the time series in this paper it is the case for some of the subsamples. The question is how one should meet this result. For an explanation of the underlying algorithm confer Dagum (1988). Like in the structural model we have assumed a multiplicative model for the unobserved components.

Besides the trend, seasonal and irregular component allowance is also made for extreme values and calendar effects. The decomposition has been done on different subsamples, with observations ending in different months of 1994. Except for the case when the observations ends in December 1994, the simple airline model is found to be adequate as a forecast model. Also when the sample ends in December 1994 one is very close to the acceptance region using the airline model, and the airline model has also been applied here. The X11-ARIMA programme supports some qualitative statistics in order to assess how successful the decomposition has been. These statistics are all constructed in such a way that they vary in the interval from 0 to 3, where values below 1 are indicative of a good decomposition model. For a further discussion of all measures confer Scott (1992) and Lothian and Morry (1978). All the subsample results show that all statistics but the first one is acceptable. The first statistic, which measure the relative contribution of the irregular component to the variance of the percentage change in the original series over 3 months span, is about 1.30. Figure 1 supports the seasonal adjusted values based on the structural time series model with the dummy seasonal specification and the seasonal adjusted values from XII-ARIMA. These values are based on the full sample. Since the seasonally adjusted values in the two structural time series are rather equal, the seasonally adjusted values from the structural time series model with harmonic seasonality has been omitted from figure 1. A comparison between the two structural models is given in figure 2.

**Figure 1. Seasonal adjustment based on a structural time series model and X11-ARIMA**

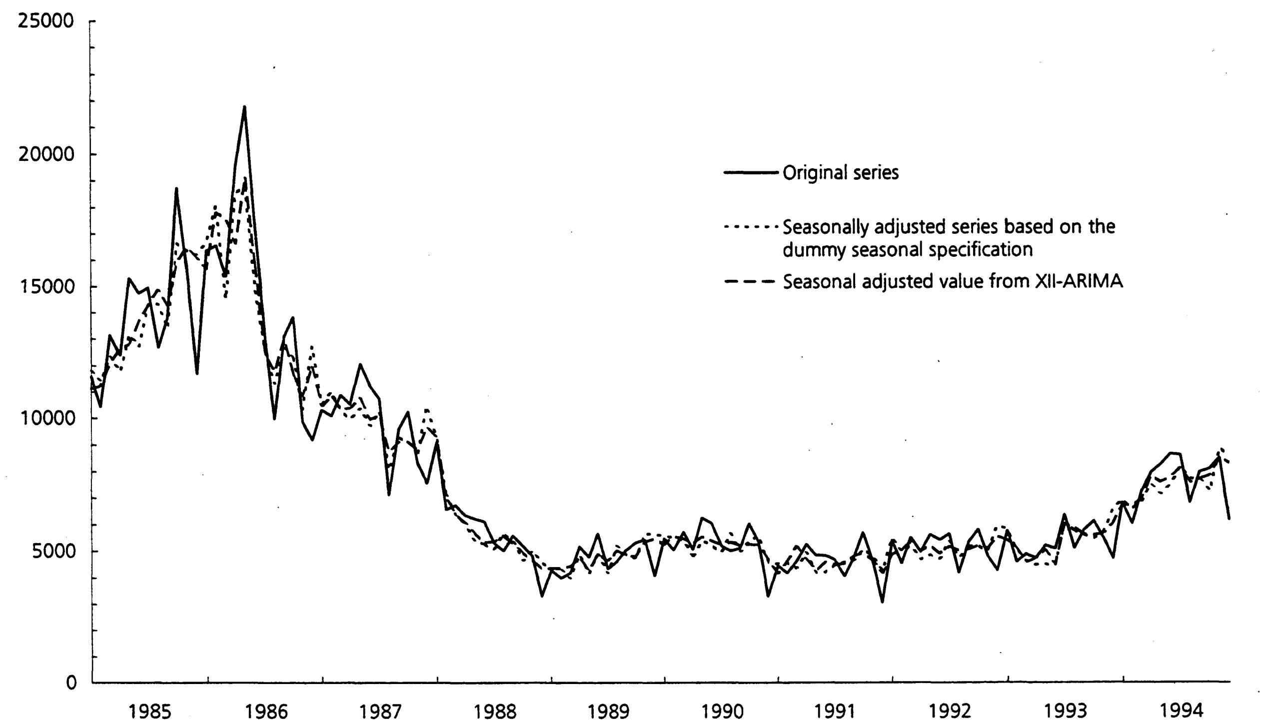
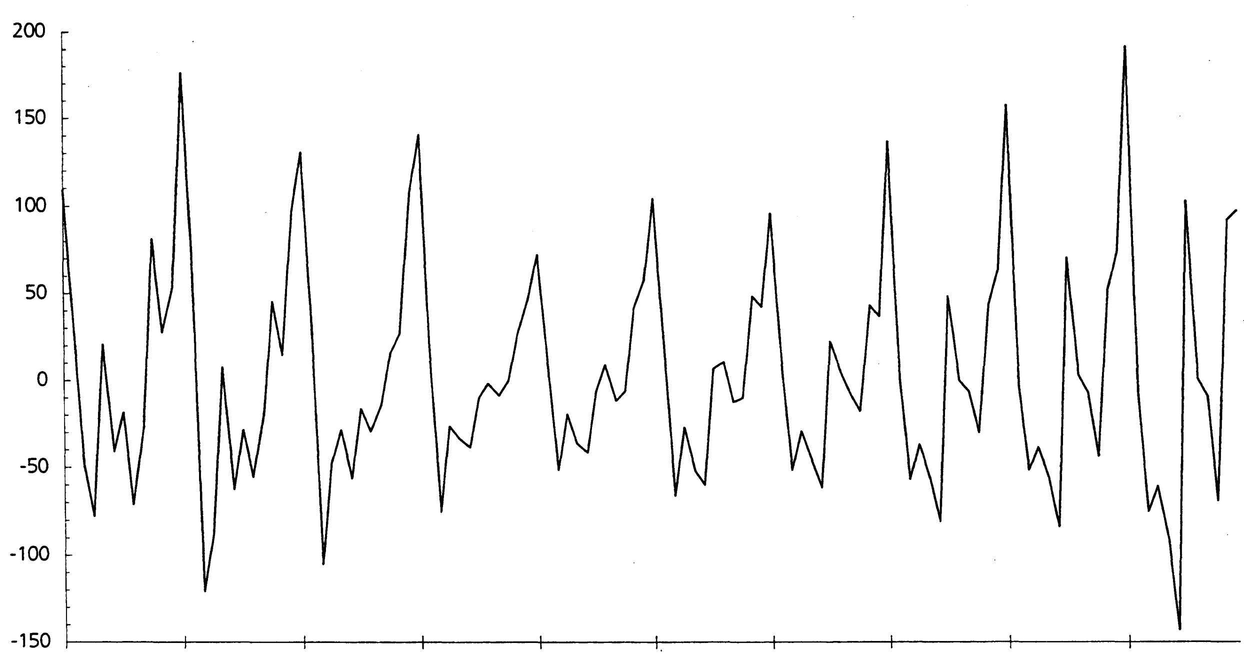



**Figure 2. The difference in the seasonally adjusted value based on the harmonic and dummy seasonal model specification**



# 7. Seasonal adjustment by structural time series models vs X11-ARIMA

It is not clear how one should compare seasonal adjustment by structural time series models and more mechanical seasonal adjustment made by X11-ARIMA. The main reason for this is that seasonal adjustment in X11-ARIMA to a large extent can be viewed as non-parametric.

However using the forecast option for extending the time serie implicitly means that an ARIMA model is fitted to the time series. A natural question to put forward is in which situations X11-ARIMA can be expected to estimate the seasonal component «optimally» assuming some reasonable definition of the seasonal component. This topic is analyzed by Wallis (1983) and Burridge and Wallis (1984). However, in order to achieve some theoretical results they had to assume the absence of data irregularities caused by calender effects and outliers which are often important properties of real time series. In this paper we will apply some practical (but formal) criteria for comparing the seasonal adjustment by the two approaches. These criteria are the same that is used by den Butter et al. (1985) and Butter and Mourik (1990) (cf also Butter and Fase (1991)) and have some interest from a practical point of view. The general idea is to measure how different methods score on different indicators for desireable properties (as viewed by practioners) of the seasonal adjusted serie and the seasonal component. However, it should be mentioned that in the above references a large amount of different time series are investigated and the interest is attached to systematic differences in performance of different methods for seasonal adjustment.

The results from measuring the scores of the different methods are given in table 3. Below we define and interprete the different indicators.

### 7.1 The average absolute percentage change of the seasonally adjusted serie

If the irregular component is important the seasonally adjusted serie will tend to be volatile and hence hard to interpret. Thus all other things equal the seasonal adjustment method which supports the smoother seasonally adjusted serie is preferable. To measure smoothness we employ the average absolute percentage change. Let $Y_t$ be the observed time serie (not log-transformed) and let $SA_t$ be the seasonal adjusted serie obtained by some method. For the structural time series model we have the following relationship between the observed value and the seasonally adjusted value:

$$(6.1) \qquad SA_t = \frac{Y_t}{e^{\tilde{\gamma}_t}},$$

where $\tilde{\gamma}_t$ denotes the extracted seasonal component obtained from the previous estimated structural time series models. The average absolute percentage change (ABPC) of the seasonally adjusted serie is now defined as:

$$(6.2) \qquad \text{ABPC}: \frac{1}{T-1}\sum_{t=2}^{T} 100\left|\frac{SA_t - SA_{t-1}}{SA_{t-1}}\right|.$$

In equation (6.2) T denotes the entire sample, which is 264. Not surprisingly the lowest value is obtained for the seasonally adjusted serie from X11-ARIMA. The omission of calender effects and outliers in the estimated structural time series models may explain why X11-ARIMA performs somewhat better.

### 7.2 Orthogonality between the seasonal component and the seasonally adjusted serie

It seems reasonable to require that the seasonal component and the seasonally adjusted serie should be uncorrelated or near uncorrelated. For the structural time series model it is easy to see that this means that the smoothed components should share a property of the theoretical variables. Remember that our stochastic assumptions postulate that the different components are theoretically uncorrelated. Formally orthogonality is defined as the correlation coefficient between the seasonal component and the seasonally adjusted serie. Let the symbol $S_t$ denote the seasonal component at period t. For the structural time series models it coincides with the denominator of equation (6.1). Thus we measure orthogonality in the following way:

(6.3) Orthogonality: $$\frac{\sum_{t=1}^{T}(S_t-\bar{S})(SA_t-\overline{SA})}{\sqrt{\sum_{t=1}^{T}(S_t-\bar{S})^2\sum_{t=1}^{T}(SA_t-\overline{SA})^2}}.$$

From table 3 we see that the correlation coefficient is small in all three cases so one can not claim that lack of orthogonality is a severe problem for any of the models.

### 7.3 Idempotency

The meaning of this property is the following. Assume that we have obtained a seasonally adjusted time serie by one arbitrary method. Let us next treat this variable as the observed variable, i. e. we try to decompose this variable in a trend component, a seasonal component and a irregular component. Since the left hand side variable is already seasonally adjusted the seasonal component should not be present. Let $SSA_t$ denote the seasonal component in level when the seasonally adjusted variable is treated as if it was the orginal serie. Formally idempotency is now defined as:

(6.4) Idempotency: $$\frac{1}{T}\sum_{t=1}^{T}100\left|\frac{SSA_t-1}{SA_t}\right|.$$

From table 3 it is evident that the sum in (6.4) is very close to zero. Thus one can argue that one can not use the idempotency criterion to distinguish between the seasonal adjustment methods.

### 7.4 Residual autocorrelation in the irregular component

As opposed to the standardized innovations which is uncorrelated by construction, there will be some autocorrelation in the smoothed estimate of the irregular component. However one do not want this to be too severe. To measure the residual autocorrelation in the irregular component obtained by different method we employ a Box-Ljung type statistic which utilizes the autocorrelations of the first twelve lags. Formally, we measure residual autocorrelation by:

(6.5) Residual autocorrelation: $T(T+2)\sum_{k=1}^{12}\frac{r_k^2}{T-k}$.

Table 3 reveals that the irregular component from X11-ARIMA contains less residual autocorrelation than the irregular component from the two structural time series models, but again this may be caused by calender effects and outliers.

### 7.5 Stability of seasonally adjusted value

The last property that will be commented on is stability. Again we need to define some new symbols. Let ly denote the latest year that is used for seasonal adjustment and let $S^{ly}$ denote the seasonal component obtained by seasonal adjustment over the indicated observation period. Let furthermore $SLT_{ly}^{ly-1}$ be defined as the mean percentage change in the seasonal components for year ly-1 by extending the adjustment period from year ly-1 to ly, i.e:

(6.6) $$SLT_{ly}^{ly-1} = \frac{1}{12}\sum_{t=T_{ly-2}}^{T_{ly-1}} 100\left|\frac{S_t^{ly} - S_t^{ly-1}}{Y_t}\right|.$$

The criterion for stability is now defined as:

Stability (latest year):

(6.7) $$\frac{1}{4}(SLT_{ly}^{ly-1} + SLT_{ly-1}^{ly-2} + SLT_{ly-2}^{ly-3} + SLT_{ly-3}^{ly-4}).$$

Table 3 shows that the two structural time series models are superior regarding stability. The stability measure is four times as high for the seasonal factors from X11-ARIMA compared to the seasonal factors calculated using the structural time series models.

**Table 3. Comparison of seasonal adjustment by structural time series models and X11-ARIMA**

| Properties | Structural time series models | | X11-ARIMA |
|---|---|---|---|
| | Dummy seasonal specificaton | Harmonic seasonal specification | |
| ABPC | 9.0674 | 8.8062 | 6.4751 |
| Orthogonality | 0.0056 | 0.0128 | 0.0105 |
| Idempotency | $2.097 \cdot 10^{-7}$ | $3.752 \cdot 10^{-7}$ | $4.382 \cdot 10^{-7}$ |
| Residual autocorrelation | 56.6 | 57.3 | 25.3 |
| Stability | 0.4751 | 0.6647 | 2.0045 |

# 8. Conclusions

In this paper we have shown how structural time series models can be used for decomposition of monthly time series in a trend, seasonal and irregular component. The term structural means that for each component of the time serie an explicit parametric statement of the underlying process is made. All the assumptions taken together ensures identification of the different components. In the most general case the trend component follows a random walk with a drift, which also is allowed to follow a random walk. Two different non-stationary stochastic processes were employed, which coincided in the degenerated deterministic seasonal case. In our empirical illustration the models were applied on the log of the number of registered new passenger cars. A model simplification in which both the drift of the trend component and the seasonality were fixed seemed reasonable, even if no formal argument was given. An important property of the models mentioned above are that they can be used for seasonal adjustment. After having estimated the system parameters, all components can be estimated using the entire sample information. Thus a trend, seasonal and irregular component are extracted, which means that the observed time serie can be corrected for seasonality. The time serie was also seasonally adjusted using X11-ARIMA and was mainly done in accordance with the way Statistics Norway officially adjusts the serie. Some practical, but formal criteria, were used in order to compare the output from the seasonal adjustment based on the two structural time series models and X11-ARIMA. The main result from this comparison was that the structural time series models and X11-ARIMA had their strengths in different areas. A feature which made the comparison of seasonal adjustment by structural time series models and X11-ARIMA difficult was the omission of calender effects and outlier adjustment of the former. This was not done deliberately, but was a result of some technical shortcomings. An alternative route could have been to run X11-ARIMA with no allowance of these effects, but then we would have removed from the official seasonal adjustment of this time series.

# References


Box, G. E. P. and G. C. Tiao (1975): Intervention analysis with applications to economic and environmental problems, *Journal of the American Statistical Association*, **70**, 70-79.

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

Burridge, P. and K. F. Wallis (1984): Unobserved-components models for seasonal adjustment filters, *Journal of Business and Economic Statistics,* **2**, *350-9.*

Dagum, E. B. (1988): The X11 ARIMA/88 seasonal adjustment method - foundations and user's manual, Time Series Research and Analysis Division, Statistics Canada, Ottawa.

Dagum, E. B., Quennville B. and B. Sudrathar (1992): Trading-day Variations Multiple Regression Models with Random Parameters, *International Statistical Review*, **60**, 1, 57-73.

den Butter, F. A. G., Coenen R. L. and F. J. J. S. van de Gevel (1985): The Use of Arima Models in Seasonal Adjustment - A Comparative Study of Census X11, X-11 ARIMA and Burman's Signal Extraction Method, *Empirical Economics,* **10**, 209-30.

den Butter, F. A. G. and M. M. G. Fase (1991): *Seasonal adjustment as a Practical Problem.* North Holland.

den Butter, F. A. G. and T. J. Mourik (1990): Seasonal Adjustment Using Structural Time Series Models: An Application with the Census X-11 Method, *Journal of Business & Economic Statistics*, **8**, 4, 385-94.

Engle, R.F. (1978): "Estimating Structural Models of Seasonality", in A. Zellner (ed.), *Applied Time Series Analysis of Economic Data*. U.S. Department of Commerce and Bureau of the Census.

Franzini, L. and A. C. Harvey (1983): Testing for deterministic trend and seasonal components in time series models, *Biometrika*, 70, 3, 673-82.

Harvey, A. C. (1989): *Forecasting, structural time series models and the Kalman filter.* Cambridge University Press.

Harvey, A. C. and P. H. J. Todd (1983): Forecasting Economic Time Series With Structural and Box-Jenkins Models: A Case Study, *Journal of Business and Economic Statistics,* **1**, 4, 299-315.

Koopman, S. J., Harvey, A. C., Doornik and N. Shephard (1995): *STAMP 5.0. Structural Time Series Analyser, Modeller and Predictor.* Chapman & Hall.

Lothian, J. and Morry, M. (1978): A set of Quality Control Statistics for the X11-ARIMA Seasonal Adjustment Program, Research Paper, Seasonal Adjustment and Time Series Staff, Statistics Canada.

Maravall, A. (1988): "The Use of ARIMA Models in Unobserved Components Estimation" in Barnett, W., Berndt E. and H. White: *Dynamic Econometric Modelling*, Cambridge: Cambridge University Press.

Maravall, A. and V. Gómez (1994): Program SEATS. Signal Extraction in ARIMA Time Series. Instruction for the User. European University Institute Working Paper ECO No 94/28, Florence.

Pfeffermann, D. (1993): A general method for estimating the variances of X11-seasonally adjusted estimators, *Journal of Time Series Analysis*, 15, 85-116.

Pfeffermann, D., Morry M. and P. Wong (1995): Estimation of the variances in X11 ARIMA seasonally adjusted estimators for a multiplicative decomposition and heteroscedastic variances, *International Journal of Forecasting*, **11**, 271-283.

Pham, D. Q. (1995): The use of intervention analysis and X11-ARIMA in the decomposition of the automobile sales in Norway. Unpublished paper, Statistics Norway.

Scott, S. (1992): An extended review of the X11ARIMA seasonal adjustment package, *International Journal of Forecasting*, **8**, 627-633.

Wallis, K. F. (1983):" Models for X-11 and X-11-Forecast procedures for preliminary and revised seasonal adjustments" in A. Zellner (ed.), *Applied Time Series Analysis of Economic Data*. U.S. Department of Commerce and Bureau of the Census.

# Appendix A. Properties of the basic structural model with the dummy seasonal specification

Let B be the backward operator such that for an arbitrary variable $x_t$ we have

(A.1) $B^i x_t = x_{t-i}$ .

Using this lag operator notation (1.2)-(1.4) can be written as:

(A.2) $(1-B)\mu_t = B\beta_t + u_t$,

(A.3) $(1-B)\beta_t = v_t$ and

(A.4) $S(B) = w_t$.

The expression S(B) in equation (A.4) denotes the summation filter and is, for monthly time series, defined by

(A.5) $S(B) = \sum_{i=0}^{11} B^i$.

If we furthermore let the difference operator at distance i, $\Delta_i$, be defined by:

(A.6) $\Delta_i = 1 - B^i$,

we obtain after having multiplied by $\Delta_1$ on both sides of (A.2):

(A.7) $\Delta_1^2 \mu_t = Bv_t + \Delta_1 u_t$.

If we now apply the combined filter $\Delta_1^2 S(B)$ on both sides of (1.1) we obtain:

(A.8) $S(B)\Delta_1^2 y_t = S(B)Bv_t + S(B)\Delta_1 u_t + \Delta_1^2 w_t + S(B)\Delta_1^2 \varepsilon_t$.

Since we have that:

(A.9) $\Delta_{12} \equiv S(B)\Delta_1$,

equation (A.8) may also be written as:

(A.10) $\Delta_1 \Delta_{12} y_t = S(B)Bv_t + \Delta_{12} u_t + \Delta_1^2 w_t + \Delta_{12}\Delta_1 \varepsilon_t$.

Equations (A.8) and (A.10) are termed stationary form representations of the nonstationary structural model. Note that the highest power of B on the right side of (A.8) and (A.10) is 13. Equation (1.15) may also be written down in its socalled reduced/canonical form. In this form the lefthand side of (A.10) is a function of a single error term. If one follows this route of action it can be shown that the reduced form of (A.10) is an MA(13)-process in which the 14 parameters are functions of the four hyperparametersin the structural time series model. In principle the constrained MA-parameters can be found by equating the autocovariances of $\Delta_1\Delta_{12}y_t$ following from the structural time series model and the MA(13) process respectively. This implies solving a simultaneous non-linear equation system. Such a system will have more than one solution, but only one of them will be consistent with invertibility of the MA-process. With the seasonal process used in this structural time series model the autocovariances are given by:

(A11.1) $\text{var}(\Delta_1\Delta_{12}y_t) = 12\sigma_{vv} + 2\sigma_{uu} + 6\sigma^2_{ww} + 4\sigma^2_{\varepsilon\varepsilon},$

(A11.2) $\text{cov}(\Delta_1\Delta_{12}y_t, \Delta_1\Delta_{12}y_{t-1}) = 11\sigma_{vv} - 4\sigma^2_{ww} - 2\sigma_{\varepsilon\varepsilon},$

(A11.3) $\text{cov}(\Delta_1\Delta_{12}y_t, \Delta_1\Delta_{12}y_{t-2}) = 10\sigma^2_{vv} + \sigma^2_{ww},$

(A11.4) $\text{cov}(\Delta_1\Delta_{12}y_t, \Delta_1\Delta_{12}y_{t-i}) = (12-i)\sigma^2_{\varepsilon\varepsilon}; 3 \leq i \leq 10,$

(A11.5) $\text{cov}(\Delta_1\Delta_{12}y_t, \Delta_1\Delta_{12}y_{t-11}) = \sigma^2_{vv} + \sigma^2_{\varepsilon\varepsilon},$

(A11.6) $\text{cov}(\Delta_1\Delta_{12}y_t, \Delta_1\Delta_{12}y_{t-12}) = -\sigma^2_{uu} - 2\sigma^2_{\varepsilon\varepsilon},$

(A11.7) $\text{cov}(\Delta_1\Delta_{12}y_t, \Delta_1\Delta_{12}y_{t-13}) = \sigma^2_{\varepsilon\varepsilon},$ and

(A11.8) $\text{cov}(\Delta_1\Delta_{12}y_t, \Delta_1\Delta_{12}y_{t-i}) = 0; i \geq 14.$

# Appendix B. Properties of the basic structural model with the harmonic seasonal specification

Further insight about the seasonal model based on harmonic functions can be obtained by solving for $\gamma_{i,t}$ in (1.7a) and (1.7b) and inserting the results together with the expression for $\gamma_{6,t}$ into (1.6). This yields when B again is the backward operator:

$$\text{(B.1)}\quad \gamma_t = \sum_{j=1}^{5}\{\frac{[1-\cos(\lambda_j)B]\omega_{j,t} + [\sin(\lambda_j)B]\omega^*_{j,t}}{[1-2\cos(\lambda_j)B + B^2]}\} + \frac{\omega_{6,t}}{1-B}.$$

The earlier introduced summation operator can be written as the product of the full complement of trigonometric operators.

$$\text{(B.2)}\quad S(B) = \prod_{j=1}^{6}\gamma_j(B), \text{where}$$

$$\text{(B.3a)}\quad \gamma_j(B) = 1-(2\cos(\lambda_j))B + B^2, j = 1,\ldots,5 \text{ and}$$

$$\text{(B.3b)}\quad \gamma_6(B) = 1 + B.$$

Let furthermore $S_j(B)$ be defined by:

$$\text{(B.4)}\quad S_j(B) = \frac{S(B)}{\gamma_j(B)}, j = 1,\ldots,6.$$

It can be shown that:

$$\text{(B.5)}\quad \gamma_t = \sum_{j=1}^{5} S_j(B)\{\frac{[1-\cos(\lambda_j)B]\omega_{j,t} + [\sin(\lambda_j)B]\omega^*_{j,t}}{S(B)}\} + \frac{S_6(B)\omega_{6,t}}{S(B)}.$$

Multiplying with S(B) on both sides of (B.5) yields:

$$\text{(B.6)}\quad \gamma_t = \sum_{j=1}^{5} S_j(B)\{[1-\cos(\lambda_j)B]\omega_{j,t} + [\sin(\lambda_j)B\omega^*_{j,t}\} + S_6(B).$$

From our stochastic assumptions it follows that only one hyperparameter is involved both for the dummy seasonal and the harmonic seasonal specification. For both seasonal models it is the case that the sum of the seasonal component taken over 12 consecutive months equals zero. However, whereas this sum follows a white noise distribution in the case with the dummy seasonal specification, it follows a smoother process in the case with the harmonic seasonality specification. Since the $S_j(B)$ terms can be written as:

(B7.1) $S_1(B) = [1 - B + B^2][1 + B^2][1 + B + B^2][1 + \sqrt{3}B + B^2][1 - B],$

(B7.2) $S_2(B) = [1 - \sqrt{3}B + B^2][1 + B^2][1 + B + B^2][1 + \sqrt{3}B + B^2][1 - B],$

(B7.3) $S_3(B) = [1 - \sqrt{3}B + B^2][1 - B + B^2][1 + B + B^2][1 + \sqrt{3}B + B^2][1 - B],$

(B7.4) $S_4(B) = [1 - \sqrt{3}B + B^2][1 - B + B^2][1 + B^2][1 + \sqrt{3}B + B^2][1 - B],$

(B7.5) $S_5(B) = [1 - \sqrt{3}B + B^2][1 - B + B^2][1 + B^2][1 + B + B^2][1 - B],$ and

(B7.6) $S_6(B) = [1 - \sqrt{3}B + B^2][1 - B + B^2][1 + B^2][1 + B + B^2][1 + \sqrt{3}B + B^2],$

it is evident that the maximum lag of the disturbance terms $\omega_{j,t}$ and $\omega^*_{k,t}$ is 10, thus $S(B)\gamma_t$ follows an MA(10) process in which the coefficients are constrained according to the above structural seasonal model. As with the dummy seasonal specification the contribution from the seasonal component in the stationary form is given by $(1-B)^2S(B)\gamma_t$. Multiplying by $(1-B)^2$ on both sides of (B.6) gives:

(B.8) $$(1 - B)^2 S(B)\gamma_t = \sum_{j=1}^{5}(1 - B)^2 S_j(B)\{[1 - \cos(\lambda_j)B]\omega_{j,t} + [\sin(\lambda_j)B]\omega^*_{j,t}\} + (1 - B)^2 S_6(B)\omega_{6,t}.$$

This means that the maximum of the disturbance terms on the right hand side of (B.8) is 12. Thus as for the model specification with dummy seasonals $(1-B)^2 S(B)y_t$ follows a constrained MA(13) process. Thus the autocorrelation function for $(1-B)^2S(B)y_t$ has a cutoff at lag 14 no matter whether one uses the dummy seasonal or the harmonic seasonal model or not. In principle one can solve for the theoretical autocorrelation function of $(1-B)^2S(B)y_t$, but because of the MA-structure of the seasonal component the algebraic calculations are much more tedious in the latter case and should be left to a programme such as Mathematica. It should also be emphasized that the number of hyperparameters are the same in both the structural time series models, i.e. 4 parameters.

# Appendix C. The number of first time registered new passenger cars in Norway 1973-1994

| | Month | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Year | Jan. | Feb. | Mar. | Apr. | May | June | July | Aug. | Sep. | Oct. | Nov. | Dec. |
| 1973 | 6219 | 6331 | 8656 | 8077 | 9289 | 10068 | 8583 | 8048 | 7186 | 8024 | 6276 | 3592 |
| 1974 | 5133 | 6644 | 8954 | 9330 | 9594 | 8616 | 8461 | 7258 | 8400 | 7737 | 6809 | 5275 |
| 1975 | 7168 | 7036 | 7579 | 9973 | 10034 | 10480 | 8641 | 7627 | 9985 | 11627 | 7969 | 6767 |
| 1976 | 8244 | 9059 | 11651 | 10872 | 12186 | 14183 | 12156 | 9747 | 11878 | 11357 | 10307 | 7989 |
| 1977 | 10739 | 10120 | 15050 | 13347 | 15174 | 16463 | 12170 | 10269 | 12358 | 11569 | 10763 | 10707 |
| 1978 | 8484 | 6918 | 7396 | 8225 | 7624 | 7101 | 5556 | 5515 | 6174 | 6666 | 6144 | 4270 |
| 1979 | 5965 | 6069 | 7818 | 7220 | 9687 | 9659 | 7165 | 6843 | 7640 | 9609 | 7689 | 5387 |
| 1980 | 7757 | 7106 | 8139 | 7627 | 9385 | 9066 | 8740 | 7297 | 8171 | 9934 | 7375 | 6213 |
| 1981 | 8139 | 8602 | 9234 | 8802 | 10251 | 9946 | 10509 | 7068 | 8518 | 9774 | 8906 | 6995 |
| 1982 | 10007 | 8836 | 10534 | 9542 | 9593 | 11919 | 10986 | 8062 | 11478 | 11220 | 9732 | 6164 |
| 1983 | 10027 | 9276 | 11156 | 9226 | 9992 | 10546 | 8731 | 8354 | 9980 | 10644 | 8730 | 5650 |
| 1984 | 8817 | 8152 | 9358 | 8319 | 10779 | 10097 | 9198 | 8261 | 8977 | 11570 | 9669 | 7191 |
| 1985 | 11563 | 10426 | 13130 | 12379 | 15304 | 14737 | 14938 | 12675 | 13794 | 18723 | 15658 | 11683 |
| 1986 | 16339 | 16548 | 15378 | 19572 | 21793 | 17241 | 13079 | 9982 | 13103 | 13816 | 9876 | 9185 |
| 1987 | 10320 | 10088 | 10892 | 10523 | 12068 | 11251 | 10747 | 7119 | 9594 | 10250 | 8319 | 7559 |
| 1988 | 9096 | 6572 | 6727 | 6352 | 6227 | 6110 | 5315 | 5014 | 5589 | 5194 | 4820 | 3292 |
| 1989 | 4288 | 3990 | 4180 | 5182 | 4774 | 5671 | 4343 | 4614 | 5037 | 5306 | 5428 | 4075 |
| 1990 | 5481 | 5054 | 5728 | 5071 | 6251 | 6048 | 5191 | 5025 | 5117 | 6047 | 5367 | 3289 |
| 1991 | 4448 | 4179 | 4607 | 5267 | 4858 | 4848 | 4684 | 4059 | 4778 | 5700 | 4636 | 3044 |
| 1992 | 5415 | 4553 | 5512 | 4967 | 5629 | 5432 | 5663 | 4194 | 5361 | 5821 | 4811 | 4275 |
| 1993 | 5769 | 4592 | 4889 | 4747 | 5228 | 5106 | 6377 | 5120 | 5797 | 6148 | 5488 | 4725 |
| 1994 | 6833 | 6044 | 7234 | 7966 | 8281 | 8679 | 8641 | 6823 | 7991 | 8096 | 8544 | 6197 |

# Previously issued on the subject

**Notater**

87/46 Lars Loe, Hilde Olsen og Tore Schweder: Kalmanfilteret som teknikk for økonomisk analyse.

85/19 Morten Jensen, Vidar Knudsen, Hilde Olsen og Tore Scweder: Sesongjustering ved X11-metoden. Rapport fra en analysegruppe.

# The most recent publications in the series Reports

94/30 Elin Berg: Estimering av investeringsrelasjoner med installasjonskostnader. 1994-86s. 95 kr ISBN 82-537-4078-6

94/31 Torbjørn Hægeland: En indikator for effekter av næringspolitiske tiltak i en økonomi karakterisert ved monopolitisk konkurranse. 1994-86s. 95 kr ISBN 82-537-4089-1

95/1 Trygve Kalve og Åne Osmunddalen: Kombinert bruk av sosialhjelp og trygdeytelser. 1995-45s. 80 kr ISBN 82-537-41057

95/2 Bjørn E. Naug: En økonometrisk modell for norsk eksport av industrielle råvarer. 1995-32s. 80 kr ISBN 82-537-4106-5

95/3 Morten Kjelsrud og Jan Erik Sivertsen: Flyktninger og arbeidsmarkedet 2. kvartal 1993. 1995-28s. 80 kr ISBN 82-537-4107-3

95/4 Inger Texmon: Ut av redet En demografisk analyse av flytting fra foreldrehjemmet. 1995-87s. 95 kr ISBN 82-537-4108-1

95/5 Resultatkontroll jordbruk 1995 Gjennomføring av tiltak mot forurensninger. 1995-90s. 95 kr ISBN 82-537-4129-4

95/6 Hilde H. Holte: Langtidsarbeidsløses levekår 1991. 1995-77s. 95 kr ISBN 82-537-4132-4

95/7 Geir Frengen, Frank Foyn, og Richard Ragnarsøn: Innovasjon i norsk industri og oljeutvinning i 1992. 1995-93s. 95 kr ISBN 82-537-4135-9

95/8 Annegrete Bruvoll og Gina Spurkland: Avfall i Noreg fram til 2010. 1995-33s. 80 kr ISBN 82-537-4136-7

95/9 Taran Fæhn, Leo Andreas Grünfeld, Erling Holmøy, Torbjørn Hægeland og Birger Strøm: Sammensetningen av den effektive støtten til norske næringer i 1989 og 1991. 1995-106s. 110 kr ISBN 82-537-4137-5

95/10 Ole Tom Djupskås og Runa Nesbakken: Energibruk i husholdningene 1993 Data fra forbruksundersøkelsen. 1995-46s. 80 kr ISBN 82-537-4138-3

95/11 Liv Grøtvedt og Liv Belsby: Barns helse Helseundersøkelsene. 1995-53s. 95 kr ISBN 82-537-4140-5

95/12 Kristin Rypdal: Anthropogenic Emissions of $SO_2$, $NO_x$, NMVOC and $NH_3$ in Norway. 1995-56s. 95 kr ISBN 82-537-4145-6

95/13 Odd Frank Vaage: Feriereiser 1993/94. 1995-48s. 80 kr ISBN 82-537-4149-9

95/14 Bodil Merethe Larsen og Runa Nesbakken: Norske $CO_2$-utslipp 1987-1993 En studie av $CO_2$-avgiftens effekt. 1995-40s. 80 kr ISBN 82-537-4158-8

95/15 Odd Frank Vaage: Kultur- og mediebruk 1994. 1995-68s. 95 kr ISBN 82-537-4159-6

95/16 Toril Austbø: Kommunale avløp Økonomi. 1995-39s. 80 kr ISBN 82-537-4162-6

95/17 Jan-Erik Lystad: Camping i Norge 1984-1994. 1995-80s. 95 kr ISBN 82-537-4170-7

95/18 Torstein Bye, Tor Arnt Johnsen og Mona Irene Hansen: Tilbud og etterspørsel av elektrisk kraft til 2020 Nasjonale og regionale framskrivinger. 1995-37s. 80 kr ISBN 82-537-4171-5

95/19 Marie W. Arneberg, Hanne A. Gravningsmyhr, Kirsten Hansen, Nina Langbraaten, Bård Lian og Thor Olav Thoresen: LOTTE - en mikrosimuleringsmodell for beregning av skatter og trygder. 1995-66s. 95 kr ISBN 82-537-4173-1

95/21 Nico Keilman and Helge Brunborg: Household Projections for Norway, 1990-2020 Part I: Macrosimulations. 1995-82s. 95 kr ISBN 82-537-4178-2

95/23 Hilde Rudlang: Bruk av edb i skolen 1995. 1995-77s. 95 kr ISBN 82-537-4181-2

95/24 Leif Brubakk, Morten Aaserud, Wilma Pellekaan and Fritz von Ostvoorn: SEEM - An Energy Demand Model for Western Europe. 1995-66s. 95 kr ISBN 82-537-4185-5

95/25 Hilde Lurås: Framskriving av miljøindikatorer. 1995-30s. 80 kr ISBN 82-537-4186-3

95/27 Knut H. Alfsen, Bodil M. Larsen og Haakon Vennemo: Bærekraftig økonomi? Noen alternative modellscenarier for Norge mot år 2030. 1995-62s. 95 kr ISBN 82-537-4190-1

C

*Returadresse:*
Statistisk sentralbyrå
Postboks 8131 Dep.
N-0033 Oslo

Publikasjonen kan bestilles fra:

Statistisk sentralbyrå
Salg-og abonnementservice
Postboks 8131 Dep.
N-0033 Oslo

Telefon: 22 00 44 80
Telefaks: 22 86 49 76

*eller:*
Akademika – avdeling for
offentlige publikasjoner
Møllergt. 17
Postboks 8134 Dep.
N-0033 Oslo

Telefon: 22 11 67 70
Telefaks: 22 42 05 51

ISBN 82-537-4200-2
ISSN 0806-2056

Pris kr 80,00

**Statistisk sentralbyrå**
**Statistics Norway**