ECONSTOR
Make Your Publications Visible.

A Service of
ZBW
Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Koopman, Siem Jan; Bos, Charles S.

Working Paper

Time Series Models with a Common Stochastic Variance for Analysing Economic Time Series

Tinbergen Institute Discussion Paper, No. 02-113/4

Provided in Cooperation with:
Tinbergen Institute, Amsterdam and Rotterdam

Suggested Citation: Koopman, Siem Jan; Bos, Charles S. (2002) : Time Series Models with a Common Stochastic Variance for Analysing Economic Time Series, Tinbergen Institute Discussion Paper, No. 02-113/4, Tinbergen Institute, Amsterdam and Rotterdam

This Version is available at:
https://hdl.handle.net/10419/86022

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

WWW.ECONSTOR.EU

Mitglied der Leibniz-Gemeinschaft

TI 2002-113/4
Tinbergen Institute Discussion Paper

Time Series Models with a Common Stochastic Variance for Analysing Economic Time Series

Siem Jan Koopman
Charles S. Bos

Faculty of Economics and Business Administration, Vrije Universiteit Amsterdam, and Tinbergen Institute

Tinbergen Institute
The Tinbergen Institute is the institute for economic research of the Erasmus Universiteit Rotterdam, Universiteit van Amsterdam and Vrije Universiteit Amsterdam.

Tinbergen Institute Amsterdam
Keizersgracht 482
1017 EG Amsterdam
The Netherlands
Tel.: +31.(0)20.5513500
Fax: +31.(0)20.5513555

Tinbergen Institute Rotterdam
Burg. Oudlaan 50
3062 PA Rotterdam
The Netherlands
Tel.: +31.(0)10.4088900
Fax: +31.(0)10.4089031

Most TI discussion papers can be downloaded at
http://www.tinbergen.nl

Time series models with a common stochastic variance for analysing economic time series

Siem Jan Koopman[*] and *Charles S. Bos* [†]
Department of Econometrics, Free University Amsterdam
and Tinbergen Institute Amsterdam

October 25, 2002

Abstract

The linear Gaussian state space model for which the common variance is treated as a stochastic time-varying variable is considered for the modelling of economic time series. The focus of this paper is on the simultaneous estimation of parameters related to the stochastic processes of the mean part and the variance part of the model. The estimation method is based on maximum likelihood and it requires the subsequent uses of the Kalman filter to treat the mean part and sampling techniques to treat the variance part. This approach leads to the evaluation of the exact likelihood function of the model subject to simulation error. The standard asymptotic properties of maximum likelihood estimators apply as a result. A Monte Carlo study is carried out to investigate the small-sample properties of the estimation procedure. We present two illustrations which are concerned with the modelling and forecasting of two U.S. macroeconomic time series: inflation and industrial production.

Keywords: Autoregressive integrated moving average; Importance sampling; Industrial production; Inflation; Kalman filter; Monte Carlo simulation; Simulation smoothing; State space; Stochastic volatility; Unobserved components time series.

JEL classification: C15, C32, C51, E23, E31

[*]Department of Econometrics, De Boelelaan 1105, NL-1081 HV Amsterdam. Email: `s.j.koopman@feweb.vu.nl`.
[†]visiting Nuffield College, Oxford OX1 1NF, UK. Email: `charles.bos@nuffield.ox.ac.uk`.

1 Introduction

The linear Gaussian state space model has become one of the standard modelling frameworks for the empirical analysis of economic time series. It is also used for forecasting and signal extraction in other fields such as engineering (from where it originates), statistics and empirical finance. The role of the state space framework for modelling macroeconomic time series and analysing business cycles is discussed in the textbooks of Harvey (1989), Hamilton (1994) and Kim and Nelson (1999). Other discussions of state space approaches to time series analysis can be found in the books of Brockwell and Davis (1987) and Shumway and Stoffer (2000). A recent account of the statistical analysis of the linear Gaussian state space model together with several non-Gaussian and nonlinear extensions is given by Durbin and Koopman (2001).

We propose a moderate generalisation of the standard linear Gaussian model: the common variance of the state space model is allowed to be stochastic and time-varying. Let us consider the stochastic local level model which is the simplest example of a state space model and is given by

$$y_t = \alpha_t + \varepsilon_t, \qquad \alpha_{t+1} = \alpha_t + \xi_t, \qquad t = 1, \ldots, n, \tag{1}$$

with observation y_t and unobservable level α_t which is modelled as a so-called random walk process. The disturbances are normally distributed with

$$\varepsilon_t \sim \mathcal{N}(0, \sigma^2), \qquad \xi_t \sim \mathcal{N}(0, \sigma^2 q^2), \qquad t = 1, \ldots, n,$$

where variance $\sigma^2 > 0$ and signal-to-noise ratio $q \geq 0$ are fixed and unknown. Further they are serially and mutually uncorrelated for $t = 1, \ldots, n$. For simplicity we assume that the random walk process is initialised by a fixed and known value for α_1. The forecast function of the local level model is the exponentially weighted moving average recursion where the discount coefficient only depends on the signal-to-noise ratio q. The common variance is represented by σ^2. The novelty is introduced by relaxing the assumption of σ^2 being fixed and instead allowing it to be stochastic and time-varying. Specifically, we replace σ^2 by σ_t^2 for which a time series process will be formulated. For example, an autoregressive model for $\log \sigma_t^2$ can be taken with the specification

$$\sigma_t^2 = \exp(h_t), \qquad h_{t+1} = (1 - \phi)d + \phi h_t + \sigma_\eta \eta_t, \qquad t = 1, \ldots, n, \tag{2}$$

with constant $d = \mathrm{E}(h_t)$, autoregressive parameter $|\phi| < 1$ and standard deviation $\sigma_\eta > 0$. The disturbances $\eta_t \sim \mathcal{N}(0, 1)$ are assumed to be serially uncorrelated and mutually uncorrelated with the local level model disturbances ε_t and ξ_t at all time points. The initial value of the autoregressive process is distributed as $h_1 \sim \mathcal{N}\{d, \sigma_\eta^2/(1 - \phi^2)\}$. A special case of the local level model (1) with the common stochastic variance specification (2) is obtained by imposing the restriction $q = 0$ (so $\alpha_{t+1} = \alpha_t = \alpha$) which reduces (1) and (2) to the basic stochastic volatility (SV) model which is considered by, for example, Taylor (1994), Harvey, Ruiz, and Shephard (1994), Danielson (1994),

Jacquier, Polson, and Rossi (1994), Shephard and Pitt (1997) and Sandmann and Koopman (1998). In this paper, however, we show that the stochastic level α_t of the model with $q > 0$ can be treated simultaneously with the common stochastic variance σ_t^2 in a statistical analysis based on maximum likelihood estimation.

Related models for this class of linear dynamic models with a common stochastic variance have been proposed in the literature. An example is the contribution of Shephard (1994) that discusses simultaneous inference of parameters related to stochastic mean and variance equations. The so-called local scale model of Shephard does not consider the common variance as stochastic but it treats the measurement variance as a stochastic variable. Further, the stochastic specification of the variance is based on the gamma-beta transition model and the estimation method is different from the maximum likelihood method presented in this paper. Other related contributions are presented by Nabeya and Tanaka (1988) who consider a linear regression model with the constant replaced by a random walk process and with the variances replaced by a deterministically time-varying common variance for both the regression disturbance and the random walk innovation, by Engle and Smith (1999) who present a model where the Stochastic Permanent Break (STOP-BREAK) process mingles transitory shocks and permanent shifts randomly in a local level model with a single GARCH error process, and by Bos, Mahieu, and van Dijk (2000) who carry out a Bayesian analysis of a local level model with different stochastic and deterministic specifications of variance processes for exchange rates.

Our proposed generalisation of the linear Gaussian state space model is partly motivated by the fact that economic time series such as inflation, interest rates, production indices and other monetary and financial series can be subject to time-varying heteroskedasticity that is potentially difficult to model explicitly. A standard solution for dealing with heteroskedasticity in econometrics is to specify a deterministic function of an exogenous variable, for example,

$$\sigma_t^2 = \sigma^2 \exp(x_t), \qquad t = 1, \ldots, n,$$

where x_t is an exogenous variable. More generally, a linear equation of explanatory variables can be used to obtain

$$\sigma_t^2 = \exp(\gamma_0 + \gamma_1 x_{1,t} + \ldots + \gamma_k x_{k,t}), \qquad t = 1, \ldots, n,$$

where $\gamma_0, \ldots, \gamma_k$ are fixed and unknown coefficients and $x_{1,t}, \ldots, x_{k,t}$ are exogenous explanatory variables. The coefficients can be estimated by numerically maximising the loglikelihood function of the underlying model. In practice it is hard to select appropriate exogenous variables that can give an adequate description of the heteroskedasticity. Furthermore this solution is clearly not practical for forecasting when no future values of explanatory variables are available.

The inclusion of a common stochastic variance can be regarded as a limited extension of the local level model since the stochastic model for the common variance applies to all variances

associated with the disturbances of the local level model. However, the proposed generalisation may be interesting for the following reasons. Firstly, many models used in practice including the autoregressive integrated moving average model and the dynamic regression model have state space representations with a single disturbance variable. In such cases the supposed restriction does not exist. This also applies to the local level model (1) with a single disturbance (that is, $\xi_t = q\varepsilon_t$ for $t = 1, \ldots, n$) as proposed by Ord, Koehler, and Snyder (1997). Secondly, when models with multiple disturbances are considered such as the local level model (1), it may not be straightforward to attribute the heteroskedasticity to a particular disturbance in practice. In such cases it is reasonable to let the common variance of the model to be stochastic and time-varying.

The state space model with a common stochastic variance poses an estimation problem since the loglikelihood function is not tractable by linear methods such as the Kalman filter due to the nonlinearities caused by the stochastic variance equation. Similar considerations apply to stochastic volatility models. Advanced but practical simulation methods have been developed to compute the loglikelihood function for such models; see, for example, the references to the stochastic volatility literature given earlier in this section. However, in the case of stochastic volatility models, estimation is only considered for parameters related to the stochastic variance and the deterministic mean equations. In this paper we consider the simultaneous estimation of parameters related to both the stochastic mean and common stochastic variance equations by the method of maximum likelihood. The key to the development in this paper is the notion that we can isolate the common variance in the statistical treatment of the mean equation. Subsequently, the common variance can then be treated as a stochastic variable and modelled separately as a result. In the case of the local level model with a common stochastic variance, the new results imply that we can consider the case of $q > 0$, rather than $q = 0$, and we are able to estimate q simultaneously with parameters ϕ and σ_ξ associated with the model of σ_t.

The extension of this model lead to a nonlinear "weighted" Kalman filter in which the weights associated with the observations y_t are determined by the model for the common stochastic variance. This becomes clear when we write the local level model (1) with a common stochastic variance as

$$y_t = \sigma_t(\alpha_t^* + \varepsilon_t^*), \qquad \alpha_{t+1}^* = \alpha_t^* + \xi_t^*,$$

where $\alpha_t^* = \alpha_t/\sigma_t$, $\varepsilon_t^* = \varepsilon_t/\sigma_t$ and $\xi_t^* = \xi_t/\sigma_t$ such that $\varepsilon_t^* \sim \mathcal{N}(0,1)$ and $\xi_t^* \sim \mathcal{N}(0, q^2)$. Thus the local level model for the weighted observations y_t/σ_t has the constant q as the signal-to-noise ratio and therefore the same forecast function as for model (1) with constant variances. However, the loglikelihood function for observation $y_1, \ldots, y_n$ generated by model (1) with common stochastic variance (2) depends on σ_t for $t = 1, \ldots, n$. The estimation of q is therefore subject to the model specification for σ_t. Further this leads to observation weights of forecast functions that change stochastically over time as a function of σ_t. The same conclusion applies to observation weights

of estimated components such as for the level component α_t

The remaining paper is organised as follows. The next section introduces the general model and discusses some examples of models of interest in economics, finance and other fields of empirical research. The estimation methodology is discussed in section 3. A Monte Carlo study is carried out to investigate the small sample properties of the estimation method and the encouraging results of this study are reported in section 4. In an empirical illustration in section 5 we show that a joint stochastic model for the mean and the variance equations may provide a solid basis for analysing monthly time series of U.S. inflation and of U.S. industrial production. The final section discusses the contribution of this paper and it discusses further extensions which we regard as future research.

2 Time series model with a common stochastic variance

2.1 General specification

The linear Gaussian state space model can be formulated with the inclusion of a common fixed variance σ^2, that is

$$\begin{aligned} y_t &= c_t + Z_t\alpha_t + G_t\varepsilon_t, \qquad \varepsilon_t \sim \mathcal{N}(0, \sigma^2 I_{r_\varepsilon}), \qquad \sigma^2 > 0, \\ \alpha_{t+1} &= \qquad T_t\alpha_t + H_t\varepsilon_t, \qquad\qquad\qquad\qquad\qquad\quad t = 1, \ldots, n, \end{aligned} \tag{3}$$

where y_t is a $p \times 1$ vector of observations, c_t is a $p \times 1$ vector of fixed effects, α_t is a $m_\alpha \times 1$ vector of unobserved states and ε_t is a $r_\varepsilon \times 1$ vector of disturbances. The system matrices Z_t, T_t, G_t and H_t are assumed to be fixed for all time points $t = 1, \ldots, n$. Unknown elements of the system matrices will be treated as parameters to be estimated by the method of maximum likelihood. The fixed effects vector c_t is given by $c_t = c + X_{c,t}\delta_c$ where c is a $p \times 1$ vector of constants while the regression coefficients associated with the $p \times k_c$ matrix $X_{c,t}$ of explanatory variables are collected in the $k_c \times 1$ vector δ_c. Initially we assume that $\alpha_1 \sim \mathcal{N}(a_1, \sigma^2 P_1)$ where a_1 and P_1 are known; later we will consider cases where elements of α_1 are generated by a diffuse density. Similar state space formulations are discussed in De Jong (1989).

The first equation is the observation equation in which the $p \times m_\alpha$ matrix Z_t selects (or weights) the appropriate elements of the state vector α_t relevant for elements in y_t. Similarly, the $p \times r_\varepsilon$ matrix G_t selects the appropriate elements of the disturbance vector ε_t. The second equation is the state equation with the $m_\alpha \times m_\alpha$ transition matrix T_t and the $m_\alpha \times r_\varepsilon$ disturbance selection matrix H_t. For many practical time series models the system matrices are time-invariant.

The time series state space model (3) can be extended by replacing σ^2 with the time-varying stochastic variable σ_t^2 so that

$$\varepsilon_t \sim \mathcal{N}(0, \sigma_t^2 I_r), \qquad \sigma_t^2 = \exp(h_t), \qquad t = 1, \ldots, n, \tag{4}$$

where h_t is a scalar and treated as the common log-variance of the model. The time series process of the stochastic variable h_t is formulated by the state space representation

$$\begin{aligned} h_t &= d_t + A_t\beta_t + C_t\eta_t, \qquad \eta_t \sim \mathcal{N}(0, I_{r_\eta}), \\ \beta_{t+1} &= \qquad B_t\beta_t + D_t\eta_t, \qquad\qquad\qquad\qquad t = 1, \ldots, n, \end{aligned} \tag{5}$$

where scalar d_t is a fixed effect, β_t is an $m_\beta \times 1$ vector of unobserved states for the log-variance equation and η_t is the $r_\eta \times 1$ vector of disturbances. The log-variance system vectors A_t and C_t and system matrices B_t and D_t are assumed to be fixed for $t = 1, \ldots, n$. In section 3 we will show how unknown elements in these system matrices can be estimated by the method of maximum likelihood. The fixed effect d_t can be modelled in the same way as c_t but with specification $d_t = d + X_{d,t}\delta_d$ where d is a constant, $X_{d,t}$ is a $1 \times k_d$ vector of explanatory variables and δ_d is the $k_d \times 1$ regression coefficient vector. The initial log-variance state vector is given by $\beta_1 \sim \mathcal{N}(b_1, Q_1)$ where initially b_1 and Q_1 are assumed known.

2.2 Gaussian linear model with fixed time-varying variances

Let us first consider the time series model (3) where the common variance is deterministic, that is $\sigma_t^2 = \exp(d + X_{d,t}\delta_d)$ for $t = 1, \ldots, n$. A special case of the general model is the regression model that is considered by Nabeya and Tanaka (1988) and is given by

$$\begin{aligned} y_t &= X_{c,t}\delta_c + \sigma_t(\mu_t + \varepsilon_t), \qquad \varepsilon_t \sim \mathcal{N}(0, \sigma_\varepsilon^2), \\ \mu_{t+1} &= \mu_t + \xi_t, \qquad\qquad\qquad\qquad \xi_t \sim \mathcal{N}(0, \sigma_\xi^2), \end{aligned} \tag{6}$$

with stochastically time-varying constant μ_t and deterministically time-varying common standard deviation σ_t for $t = 1, \ldots, n$. They used this model for developing asymptotic tests for parameter constancy over time. This special case of the linear Gaussian state space model can be analysed using the Kalman filter and the associated methods of maximum likelihood estimation, diagnostic checking, signal extraction and forecasting; see, for example, Harvey (1989), Shumway and Stoffer (2000) and Durbin and Koopman (2001, Part I). These treatments also allow for cases where other elements of the system matrices vary over time deterministically.

Model (6) for which the common variance σ^2 is allowed to vary over time is considered in this paper and maximum likelihood estimation of its parameters is treated generally in section 3.

2.3 Stochastic volatility model

Another special case of the general model (3), (4) and (5) is obtained by taking $\alpha_t = 0$ such that the mean specification is fixed and that only the variance specification of the model is stochastic. The restriction leads to the specification of a stochastic volatility (SV) model of the form

$$y_t = c_t + \sigma_\varepsilon \varepsilon_t, \qquad \varepsilon_t \sim \mathcal{N}(0, \sigma_t^2), \qquad \sigma_t^2 = \exp(h_t), \qquad t = 1, \ldots, n, \tag{7}$$

where σ_ε replaces G_t in equation (3). The log-volatility h_t can be modelled as the stationary autoregressive process (2) that we have as a special case of (5) with $d_t = (1-\phi)d$, $A_t = 1$, $B_t = \phi$, $C_t = 0$, $D_t = \sigma_\eta > 0$, $b_1 = d$, $Q_1 = \sigma_\eta^2/(1-\phi^2)$ and with $|\phi| < 1$. Various contributions have appeared in the literature for estimating SV models by maximum likelihood using importance sampling techniques; see Danielson (1994), Shephard and Pitt (1997), Durbin and Koopman (1997), Sandmann and Koopman (1998) and Durham and Gallant (2002). Other methods of inference have also been considered in the context of SV models including Bayesian methods, *e.g.* Kim, Shephard, and Chib (1998), and efficient method of moments procedures, *e.g.* Gallant, Hsieh, and Tauchen (1997).

2.4 ARMA models with stochastic variances

It is well-known that the autoregressive moving average (ARMA) model can be represented in state space. The general model allows the ARMA(p^*, q^*) model to have a stochastic variance specification, that is

$$y_t = \varphi_1 y_{t-1} + \ldots + \varphi_{p^*} y_{t-p^*} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \ldots + \theta_{q^*} \varepsilon_{t-q^*}, \qquad \varepsilon_t \sim \mathcal{N}(0, \sigma_t^2), \tag{8}$$

for $t = 1, \ldots, n$ and with orders $p^* \geq 0$ and $q^* \geq 0$. The log-variance $h_t = \log \sigma_t^2$ can be modelled by an autoregressive process such as in (2) or by any other time series process that can be represented in the general state space form (5).

2.5 Unobserved components models with common stochastic variances

The local level model (1) is a special case of the unobserved components time series model that consists of the level α_t and the irregular ε_t as the two unobservables. More elaborate specifications within this class of time series models are discussed in Harvey (1989). The so-called basic structural time series model consists of trend, seasonal and irregular components. This model with a common stochastic variance is given by

$$\begin{aligned} y_t &= \mu_t + \gamma_t + q_1 \varepsilon_{1t}, \\ \mu_{t+1} &= \mu_t + \beta_t + q_2 \varepsilon_{2t}, \\ \beta_{t+1} &= \beta_t + q_3 \varepsilon_{3t}, \\ S(L)\gamma_{t+1} &= q_4 \varepsilon_{4t}, \end{aligned} \tag{9}$$

for $t = 1, \ldots, n$ and where $S(L)$ is the seasonal sum operator $1 + L + \ldots + L^{s-1}$ for seasonal length s and lag operator L such that $Ly_t = y_{t-1}$. Variable ε_{it} is the i-th element of the disturbance vector ε_t which is distributed as in (4) and with the common log-variance given by (5), for $i = 1, \ldots, 4$. The unknown fixed coefficients $q_1, \ldots, q_4$ are unknown and can be estimated by maximum

likelihood. The variances are uniquely identified in this model provided that the common log-variance h_t follows a stationary process with zero mean and some constant variance σ_h^2. In this case we note that the variance of the ith disturbance is $q_i^2\sigma_t^2$ with unconditional expectation

$$q_i^2 \exp\{2\mathrm{E}(h_t) + \mathrm{var}(h_t)\} = q_i^2 \exp \sigma_h^2,$$

for $i = 1, \ldots, 4$. When h_t is modelled as a nonstationary process, the restriction of the zero mean for this process is replaced by the restriction that $h_0 = 0$ is fixed.

The state space representation of the mean equation of the model (9) for quarterly data is based on the 5×1 state vector

$$\alpha_t = \begin{pmatrix} \mu_t & \beta_t & \gamma_t & \gamma_{2t} & \gamma_{3t} \end{pmatrix}',$$

where γ_{2t} and γ_{3t} are auxiliary variables required for the formulation of the seasonal component in state space. Other specifications for the components, in particular the seasonal component, can be considered and they are discussed in detail by Harvey (1989). We finally note that the initial state vector is diffuse and requires specific modifications for the analysis which are discussed in section 3.5.

2.6 Single source of error models

Ord, Koehler, and Snyder (1997) consider unobserved components models such as the ones discussed in the previous section but with $\varepsilon_{1t} = \varepsilon_{2t} = \varepsilon_{3t} = \varepsilon_{4t}$; they refer to this class of models as single source of error models. Nonlinear extensions of this class of models are also considered. They argue that the single source of error models cover a larger extent of the parameter space in the stationary representation of the model and that it therefore will produce more accurate model-based forecasts. In a general treatment of time series models with correlated disturbances, Harvey and Koopman (2000) argue that single source of error models (or models with perfectly correlated disturbances) are not appropriate for signal extraction. It is noted that the common stochastic variance discussed in this paper apply naturally to the single source of error models.

A specific example of a nonlinear single source of error model is discussed by Engle and Smith (1999) and is known as the stochastic permanent break (STOPBREAK) model. In its simplest form the model is given by

$$y_t = \mu_t + q_t\varepsilon_t, \qquad \mu_{t+1} = \mu_t + \varepsilon_t, \qquad t = 1, \ldots, n, \tag{10}$$

where q_t is a function of past realisations of ε_t. The STOPBREAK model can be casted into the general specification given earlier in section 2.1 when q_t is deterministic and not a function of lagged ε_t's and when σ^2 is replaced by the common stochastic variance σ_t^2 as modelled by (4) and (5).

3 Estimation by maximum likelihood

In this section we develop a method for computing the Monte Carlo estimate of the loglikelihood function for state space models with common stochastic variance. In the standard situation, with deterministic variances, the density $p(y|\psi)$ of the data can be computed through a prediction-error decomposition (see section 3.1). The loglikelihood is then the sum of the individual contributions of the prediction errors v_t. With stochastic variance however, the Kalman equations cannot be used for extracting the v_t's as the variances are not available in the filtering equations. Only after conditioning on the common stochastic variance can the loglikelihood be computed from the output of the Kalman filter. Importance sampling is used to adjust for the bias introduced by conditioning, as explained in subsequent sections.

3.1 Kalman filter

Consider model (3), (4) and (5) and define vector ψ as the collection of unknown elements of the system matrices for the mean and variance equations. Given a realised sequence for $\sigma^* = (\sigma_1, \ldots, \sigma_n)'$, the joint density of the observations $y = (y_1', \ldots y_n')'$ can be obtained via the prediction error decomposition. In particular, the conditional logdensity of y is given by

$$\log p(y|\sigma^*; \psi) = -\frac{nN}{2} \log 2\pi - \frac{1}{2} \sum_{t=1}^{n} \left(\log |\sigma_t^2 F_t| + \sigma_t^{-2} v_t' F_t^{-1} v_t \right), \tag{11}$$

where v_t is the vector of one-step ahead prediction errors and $\sigma_t^2 F_t$ is its variance matrix; see Schweppe (1965) and Harvey (1989). Both v_t and F_t are produced by the Kalman filter as given by

$$\begin{aligned} v_t &= y_t - c_t - Z_t a_t, & F_t &= Z_t P_t Z_t' + G_t G_t', \\ & & K_t &= T_t P_t Z_t' F_t^{-1}, \\ a_{t+1} &= T_t a_t + K_t v_t, & P_{t+1} &= T_t P_t (T_t - K_t Z_t)' + H_t H_t', \end{aligned} \tag{12}$$

for $t = 1, \ldots, n$ and where a_1 and P_1 are known values and represent the unconditional mean and variance of the initial state vector, respectively. The vector a_t is the linear estimator or predictor of the state α_t conditional on $\{y_1, \ldots, y_{t-1}\}$ with variance matrix P_t, for $t = 1, \ldots, n$. The intermediate matrix K_t is known as the Kalman gain. Proofs of the Kalman filter are given by Anderson and Moore (1979) and Durbin and Koopman (2001, §4.2.1). We note that the Kalman filter equations as given by (12) do not depend on σ^*. This implies that v_t and F_t do not change when σ^* vary for $t = 1, \ldots, n$.

The innovations v_t can be regarded as linear transformations of y for a given ψ; see Harvey (1989). The consequence of the well-known Gaussian properties is that the innovation vectors are

also normally distributed and given by

$$v_t \sim \mathcal{N}(0, \sigma_t^2 F_t), \qquad t = 1, \ldots, n. \tag{13}$$

It is further noticed that the Jacobian of the transformation from y to $v = (v_1', \ldots v_n')'$ is unity. Therefore the density of y is the same as the density of v, that is

$$p(y|\sigma^*; \psi) = p(v|\sigma^*; \psi),$$

where we can take v as being generated by the Gaussian model (13) for given σ^*.

The key to the simultaneous estimation of parameters related to the stochastic mean and variance equations of model (3)–(5) is the result that the statistical properties of the innovations can be summarised by (13). This shows that σ_t^2 is the common denominator of the variance matrix of v_t. When σ_t^2 is the constant σ^2 in model (3), it is a well-known fact that σ^2 can be *concentrated out* of the loglikelihood function and that equation (13) holds for $\sigma_t^2 = \sigma^2$; see Harvey (1989). The maximum likelihood estimate of σ^2 is then simply given by

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{t=1}^{n} v_t' F_t^{-1} v_t.$$

3.2 Importance sampling

When σ_t follows a stochastic process as given by (4) and (5) we can not express the Gaussian density $p(y; \psi)$ in analytical terms. This also applies to density $p(v; \psi)$ associated with model (13). The difficulty is well recognised in the econometric and statistical literature on stochastic volatility models; see, for example, the overview articles by Ghysels, Harvey, and Renault (1996) and Shephard (1996). We note that model (13) with σ_t modelled as (2) can be regarded as the SV model for the innovations v_t. A standard tool in evaluating non-tractable densities is importance sampling and is applied to SV models by Danielson (1994), Shephard and Pitt (1997), Sandmann and Koopman (1998) and Durham and Gallant (2002). The same approach will be taken in this paper. The technique involves approximating the solution of the density via averages of simulations from an approximating model. Importance sampling was first used in econometrics by Kloek and Van Dijk (1978) in their work on computing posterior densities. Geweke (1989) proposes to use the Lindeberg-Lévy central limit theory in order to assess the accuracy of the importance sampler. Koopman and Shephard (2002) discuss practical procedures based on extreme value theory for empirically testing the asymptotic normality of the importance sampling estimator.

Given the innovations and their scaled variance matrices (both can be computed without observing σ^*), we express the density of the observations as

$$\begin{aligned} p(y; \psi) &= \int p(y|\sigma^*; \psi) p(\sigma^*; \psi) d\sigma^* \\ &= \int p(v|\sigma^*; \psi) p(\sigma^*; \psi) d\sigma^*, \end{aligned} \tag{14}$$

where $p(v|\sigma^*;\psi)$ is the density of model (13) as given by (11) for a realised value of σ^*. Given a realised value of y, the evaluation of the resulting likelihood function (14) via importance sampling is based on an importance sampling density $g(\sigma^*|v;\psi)$ from which it is relatively easy to simulate from. For considerations of feasibility, the importance density $g(\cdot)$ is chosen to be close to the density $p(\cdot)$. The observation density associated with the true model can then be represented by

$$p(y;\psi) = \int \frac{p(v|\sigma^*;\psi)p(\sigma^*;\psi)}{g(\sigma^*|v;\psi)} g(\sigma^*|v;\psi)d\sigma^*, \tag{15}$$

which we approximate via importance sampling. This requires generating Monte Carlo simulations from the importance density $g(\sigma^*|v;\psi)$ to obtain the estimator

$$\widehat{p(y;\psi)} = \frac{1}{M}\sum_{i=1}^{M} \frac{p(v|\sigma^{*(i)};\psi)p(\sigma^{*(i)};\psi)}{g(\sigma^{*(i)}|v;\psi)}, \tag{16}$$

where $\sigma^{*(i)}$ is a realisation from the importance density $g(\sigma^*|v;\psi)$. The Monte Carlo approximation (16) of the true density (15) will be based on a Gaussian importance density $g(\sigma^*|y;\psi)$ in this paper. We note that the estimator (16) is only subject to simulation error.

The importance density $g(\sigma^*|y;\psi) = g(\sigma^*|v;\psi)$ for model (13) can be based on the approximating linear Gaussian state space model

$$v_t = h_t + u_t, \qquad u_t \sim \mathcal{NID}(r_t, s_t), \qquad t = 1,\ldots,n, \tag{17}$$

where v_t is obtained from the time series $y_1,\ldots,y_n$ by the Kalman filter as described in §3.1 and where h_t is given by the model specification for $h_t = \log\sigma_t^2$ as in (5). The time-varying mean r_t and variance s_t are identified via a recursive procedure for which the details are given by Shephard and Pitt (1997) and Durbin and Koopman (2001). Generating simulations from the importance density $g(\sigma^*|v;\psi)$ are in effect based on the approximating model (17) and is referred to as simulation smoothing. Various simulation smoothing methods exist to compute conditional draws from the linear Gaussian state space model. A recent account of simulation smoothing together with the development of a new, simple and efficient device for this is given by Durbin and Koopman (2002).

In practice, focus is on the logdensity and in the case of a Gaussian importance density it can be shown that

$$\log\widehat{p(y;\psi)} = \log g(v;\psi) + \log\left[\frac{1}{M}\sum_{i=1}^{M}\frac{p(v|\sigma^{*(i)};\psi)}{g(v|\sigma^{*(i)};\psi)}\right] + O(M^{-3/2}), \tag{18}$$

where $\log g(v;\psi)$ is the logdensity of the approximating model that can be evaluated by the Kalman filter; the details of this result are given by Durbin and Koopman (1997).

3.3 Estimation procedure

It has been shown that we can evaluate the loglikelihood function for a linear Gaussian model with a common stochastic time-varying variance using importance sampling methods. The full

procedure is relatively easy to implement and can be summarised by the following five steps:

1. Apply the Kalman filter to model (3) without any consideration of $\sigma_1^2, \ldots, \sigma_n^2$; store v_t and F_t for $t = 1, \ldots, n$.

2. Consider model (13) and obtain an approximating model; compute the loglikelihood function of the approximating model using the Kalman filter and generate M samples from the resulting Gaussian importance density $g(\sigma^*|y;\psi)$ using the device of Durbin and Koopman (2002).

3. Compute the Monte Carlo estimate of the loglikelihood function given by

$$\log \widehat{\ell(y;\psi)} = \log g(v;\psi) + \log \left[\frac{1}{M} \sum_{i=1}^{M} \frac{p(v|\sigma^{*(i)};\psi)}{g(v|\sigma^{*(i)};\psi)} \right].$$

4. Maximise this loglikelihood with respect to ψ; the evaluations of the Monte Carlo estimator $\log \widehat{\ell(y;\psi)}$ for different values of ψ is based on the same set of random numbers for the sampling of σ^* from the Gaussian density $g(\sigma^*|y;\psi)$ to ensure a smooth loglikelihood function in ψ.

This procedure is numerically stable and is implemented for the object-oriented matrix programming language Ox of Doornik (1999) using the state space functions in the Ox library SsfPack as documented by Koopman, Shephard, and Doornik (1999). The programs written for this paper can be obtained from `www.ssfpack.com`.

3.4 Signal extraction and forecasting

The state vector α_t and disturbance vector ε_t of model (3) with known σ_t^2 can be estimated by standard state space methods; see Durbin and Koopman (2001). For example, state smoothing recursions exist for the linear Gaussian state space model that compute $\hat{\alpha}_t = \mathrm{E}(\alpha_t|y)$ and $V_t = \mathrm{var}(\alpha_t|y)$ for $t = n, \ldots, 1$. In the case of a stochastic time-varying process for σ_t we need to take account of the variation of σ_t because a realisation of α_t depends partially on σ_t. In a similar way as for the construction of the likelihood function, we can develop an importance estimator for α_t which is given by

$$\hat{\alpha}_t = \bar{\omega}^{-1} \sum_{i=1}^{M} \omega_i \hat{\alpha}_t^{(i)}, \qquad \omega_i = \frac{p(v|\sigma^{*(i)};\psi)}{g(v|\sigma^{*(i)};\psi)}, \qquad \bar{\omega} = \sum_{i=1}^{M} \omega_i, \tag{19}$$

where $\hat{\alpha}_t^{(i)}$ is the smoothed estimate $\mathrm{E}(\alpha_t|y)$ for α_t in model (3) with σ replaced by time-varying standard deviations in the $n \times 1$ vector $\sigma^* = \sigma^{*(i)}$. The smoothed state variance is given by

$$\mathrm{var}(\alpha_t|y) = \bar{\omega}^{-1} \sum_{i=1}^{M} \omega_i V_t^{(i)},$$

where $V_t^{(i)}$ is the smoothed state variance V_t for model (3) with σ replaced by time-varying standard deviations in $\sigma^* = \sigma^{*(i)}$.

In a similar fashion we can compute the filtered estimate of the state vector $\mathrm{E}(\alpha_t|y_1, \ldots, y_t)$. However, the importance sampling weights ω_i are ratios of densities of v which is a linear transformation of y. This estimate can therefore not be properly defined as a filtered estimate. Other simulation techniques such as Monte Carlo filtering (or particle filtering) that is reviewed in Doucet, deFreitas, and Gordon (2000) should be employed for filtering. These techniques shall not be considered in this paper. Finally, forecasts can be computed by importance estimation techniques since forecasts are conditional on y.

Algorithms for computing observation weights for the construction of filtered and smoothed estimates of the state vector are reviewed in Harvey and Koopman (2000) for the linear Gaussian unobserved components model. These weights can also be computed for models with a common stochastic variance and they can be constructed in a similar way as in (19). In particular, the observation weights of the estimated state vector $\hat{\alpha}_t$ are given by

$$w_{t,j} = \bar{\omega}^{-1} \sum_{i=1}^{M} \omega_i w_{t,j}^{(i)}, \qquad j = 1, \ldots, n,$$

where $w_{t,j}^{(i)}$ is the jth observation weight of the smoothed state $\hat{\alpha}_t$ for the model (3) with σ replaced by time-varying standard deviations in $\sigma^* = \sigma^{*(i)}$.

3.5 Diffuse initialisation

Time series models with nonstationary dynamic specifications require diffuse initial conditions of the state vector when such models are casted in state space form. The consequences of diffuse initial conditions for Kalman filtering and smoothing have been discussed elsewhere; see, for example, the treatment given by Durbin and Koopman (2001, Chapter 5) and the references therein. In the case of simulation smoothing, diffuse initialisations can be accounted for in a straightforward fashion for which the details are given by Durbin and Koopman (2002).

4 Small-sample properties: A Monte Carlo study

Since the unknown parameters are estimated by exact maximum likelihood, subject to simulation error, the asymptotic properties of the estimators apply as usual. Given the numerical involvement of the estimation procedure, it is interesting to investigate the finite sample properties of the estimators. We therefore carry out a Monte Carlo study.

4.1 Design of the Monte Carlo study

We consider the following three different model specifications for the mean equation: the stationary autoregressive moving average (ARMA) model (8), the local level model (1) and the unobserved components model (9). Further, the common stochastic variance (CSV) is based on the log-variance $h_t = \log \sigma_t$ which is either fixed at zero, $h_t = 0$, or is modelled by the autoregressive model (2). The different data generation processes (DGPs) and their parameter values are presented in Table 1.

Table 1: Data generation processes of the Monte Carlo study

	No CSV	CSV
ARMA	y_t as in (8) with $p^* = q^* = 1$, $\varphi_1 = 0.8,\ \ \theta_1 = -0.6$, $h_t = 0$.	y_t as in (8) with $p^* = q^* = 1$, $\varphi_1 = 0.8,\ \ \theta_1 = -0.6$, h_t as in (2) with $d = 0,\ \ \phi = 0.9,\ \ \sigma_\eta = 0.2$.
LL	y_t as in (1) with $q = 0.5$, $h_t = 0$.	y_t as in (1) with $q = 0.5$, h_t as in (2) with $d = 0,\ \ \phi = 0.9,\ \ \sigma_\eta = 0.2$.
BSM	y_t as in (9) with $s = 4$, $q_3 = 0$, $\beta_1 = 0$, $q_1 = 1,\ \ q_2 = 0.5,\ \ q_4 = 0.2$, $h_t = 0$.	y_t as in (9) with $s = 4$, $q_3 = 0$, $\beta_1 = 0$, $q_1 = 1,\ \ q_2 = 0.5,\ \ q_4 = 0.2$, h_t as in (2) with $d = 0,\ \ \phi = 0.9,\ \ \sigma_\eta = 0.2$.

For each DGP, $M = 500$ series are generated with three different number of observations: $n = 100, n = 500$ and $n = 1000$. Each generated series is estimated twice using two different model specifications. Both models adopt the "true" mean equation of the DGP while the common log-variance is either fixed at zero (see column No CSV) or stochastic and modelled as an autoregressive process (see column CSV). In the remaining part of this section we discuss the simulation results that are based on this simulation design and that are presented in the Tables 2–8.

4.2 ARMA simulation results

The simulation results for the ARMA model are presented in Table 2 which consists of two panels: the first panel considers the DGP without a common stochastic variance (No CSV) and the second panel considers the DGP with CSV. Each panel has four vertical blocks of which the last three are associated with the three different sample sizes (100, 500 and 1000). In each of these blocks a summary of the estimation results are presented for the $M = 500$ simulated series without CSV and with CSV. The sample mean of the estimated parameters for the M series is given together with its sample standard deviation. Further, averages of two diagnostic statistics are reported together with the percentages of M diagnostics that have not passed the corresponding critical value. The first diagnostic test statistic is the normality test of Doornik and Hansen (1994),

which is an adapted version of the test for normality of Bowman and Shenton (1975), and is χ^2 distributed with two degrees of freedom. The second diagnostic is the heteroskedasticity test $H(h) = \sum_{t=T-h+1}^{T} v_t^2 / \sum_{t=d+1}^{d+1+h} v_t^2$ statistic and is $F_{h,h}$ distributed with $h \approx n/3$. Finally the average value of the M maximised log-likelihood values is reported for both estimated models in each block.

Table 2: Small sample results for the ARMA model

		Generating no SV											
		$n = 100$				$n = 500$				$n = 1000$			
Estimating		No SV		SV		No SV		SV		No SV		SV	
	p	$\overline{p}$	s.d.	$\overline{p}$	s.d.	$\overline{p}$	s.d.	$\overline{p}$	s.d.	$\overline{p}$	s.d.	$\overline{p}$	s.d.
φ	0.8	0.72	0.24	0.70	0.27	0.79	0.08	0.78	0.08	0.79	0.05	0.79	0.05
ϑ	-0.6	-0.53	0.28	-0.51	0.30	-0.59	0.10	-0.59	0.11	-0.60	0.07	-0.59	0.07
σ_{ARMA}	1	0.98	0.07	0.96	0.08	1.00	0.03	0.99	0.04	1.00	0.02	1.00	0.02
σ_{SV}				0.16	0.23			0.11	0.13			0.09	0.11
ϕ_{SV}				0.36	0.33			0.37	0.31			0.33	0.27
		Stat	f_{Rej}	Stat	f_{Rej}	Stat	f_{Rej}	Stat	f_{Rej}	Stat	f_{Rej}	Stat	f_{Rej}
Normality-DH		2.09	0.07	2.03	0.07	2.00	0.05	1.90	0.05	2.04	0.06	1.88	0.05
H($n/3$)		1.10	0.35	1.10	0.34	1.02	0.14	1.02	0.13	1.01	0.04	1.01	0.06
LL		-139.70		-139.06		-707.35		-708.98		-1416.60		-1419.20	
		Generating SV											
		$n = 100$				$n = 500$				$n = 1000$			
Estimating		No SV		SV		No SV		SV		No SV		SV	
	p	$\overline{p}$	s.d.	$\overline{p}$	s.d.	$\overline{p}$	s.d.	$\overline{p}$	s.d.	$\overline{p}$	s.d.	$\overline{p}$	s.d.
φ	0.8	0.70	0.30	0.69	0.29	0.78	0.09	0.78	0.09	0.79	0.05	0.79	0.05
ϑ	-0.6	-0.50	0.33	-0.49	0.33	-0.59	0.12	-0.58	0.11	-0.59	0.07	-0.59	0.07
σ_{ARMA}	1	1.03	0.12	0.97	0.12	1.05	0.06	1.00	0.06	1.05	0.04	1.00	0.04
σ_η	0.2			0.28	0.25			0.22	0.10			0.21	0.08
ϕ	0.9			0.64	0.34			0.84	0.16			0.87	0.08
		Stat	f_{Rej}	Stat	f_{Rej}	Stat	f_{Rej}	Stat	f_{Rej}	Stat	f_{Rej}	Stat	f_{Rej}
Normality-DH		4.04	0.23	2.49	0.09	10.90	0.56	2.28	0.05	20.89	0.84	2.85	0.10
H($n/3$)		1.23	0.41	1.07	0.33	1.04	0.27	1.01	0.02	1.02	0.20	1.00	0.01
LL		-144.37		-142.84		-732.86		-727.81		-1469.69		-1453.81	

Reported are the average value of the estimated parameters ($\overline{p}$) together with the standard deviations (s.d.) of the estimates. The average value of the Normality (DH) and heteroskedasticity (H) test statistics are reported, together with the fraction of rejections (f_{Rej}), using a significance level of 5%. Furthermore, the average loglikelihood value (LL) is given.

The true DGP model that forms the basis of the results in Table 2 is the ARMA model (8) with $p^* = q^* = 1$ and with coefficients $\varphi = 0.8, \vartheta = -0.6$ and $\sigma_t^2 = 1$ for $t = 1, \ldots, n$ for the case without CSV. Estimating the coefficients using a realised time series from the true model may be difficult due to the existence of strong (negative) correlation between estimates of φ and ϑ caused by possible root cancellations. For the DGP with CSV we consider model (2) for σ_t^2 with coefficients $d = 0$, $\phi = 0.9$ and $\sigma_\eta = 0.2$.

For the smaller sample sizes, estimates of φ and ϑ have relatively large standard errors but they become smaller for larger samples. The estimated coefficients for the mean equation of the model are not biased irrespective of the specification of the common variance part of the model. The estimated coefficients for the variance equation of the model tend to be insignificant when

the DGP is without CSV.

When the DGP has a fixed variance (No CSV), the normality test is rejected in about 5% of the cases, so the size of the test is correct. When the DGP includes the CSV specification but it is not estimated, normality is rejected in 23% of the cases for $n = 100$ and in 84% of the cases for $n = 1000$. The heteroskedasticity test appears to be a less reliable test statistic for detecting of model misspecification due to the omission of a common stochastic variance. When CSV is both generated and estimated, the size of the normality test is also not correct for all sample sizes; the rejection rate of the normality test can be 10%.

The estimation of state space models with a common stochastic variance is more demanding compared to the estimation of standard state space models. However, the extra computational burden due to the use of simulation methods is not excessive. As an overall conclusion we emphasize that estimation results can be obtained quickly and the precision is good.

4.3 Other simulation results

A Monte Carlo study has also been carried out for the class of unobserved components models with and without a common stochastic variance equation. Tables 7 and 8 (included at the end of this paper) report the results for DGPs based on model (9). Table 7 concerns the local level model (1) with $\sigma^2 = 1$ and $q = 0.5$ or, equivalently, the basic structural model (9) with $q_1 = 1$, $q_2 = 0.5$ and $\beta_t = \gamma_t = 0$ for $t = 1, \ldots, n$. The DGP with CSV has the same specification for the common stochastic variance σ_t^2 as for the ARMA model of the previous section. Table 8 concerns the local level model with a stochastic seasonal component, that is model (9) with $q_3 = 0$ and $\beta_1 = 0$ such that $\beta_t = 0$ for $t = 1, \ldots, n$. The parameter values for this DGP are selected as $q_1 = 1$, $q_2 = 0.5$ and $q_4 = 0.2$. The same specification for the CSV is taken as for the other models.

The Monte Carlo results for the two models produce similar results as for the ARMA model reported in the previous section. It is surprising to see that when the DGP is taken with no CSV (that is, $h_t = 0$) and estimation is based on the model with CSV, the estimated coefficients of the mean equation have no bias while the estimated coefficients of the variance equation appear not to be significant. The reverse case (CSV is part of DGP but it is not estimated) does also produce unbiased estimates for the mean equation but in this case the normality test statistics are unsatisfactory and on these grounds most of the estimated models would have correctly been rejected when sample sizes are sufficiently large. The heteroskedasticity tests are also unreliable in this context.

5 Analysing macro-economic time series: two illustrations

5.1 U.S. monthly inflation rate

We consider the first difference of the logarithm of the monthly U.S. CPI index[1] between 1957:1 and 2001:9 and is taken as the indicator of U.S. inflation. The data is presented in figure 1. In terms of an unobserved components model, the time series plot of inflation suggests the inclusion of level and seasonal components. Therefore we base our analysis on model (9) with $s = 12$ and $\beta_t = 0$ for $t = 1, \ldots, n$ and $n = 537$. The common variance specification for σ_t^2 is based on the log-variance specification (2) with $d = 0$. The parameter vector ψ for this model is given by

$$\psi = (q_1, q_2, q_4, \phi, \sigma_\eta)',$$

where the coefficients q_1, q_2 and q_4 refer to equation (9) and ϕ and σ_η refer to (2). The mean equation requires a diffuse initial condition for the state vector.

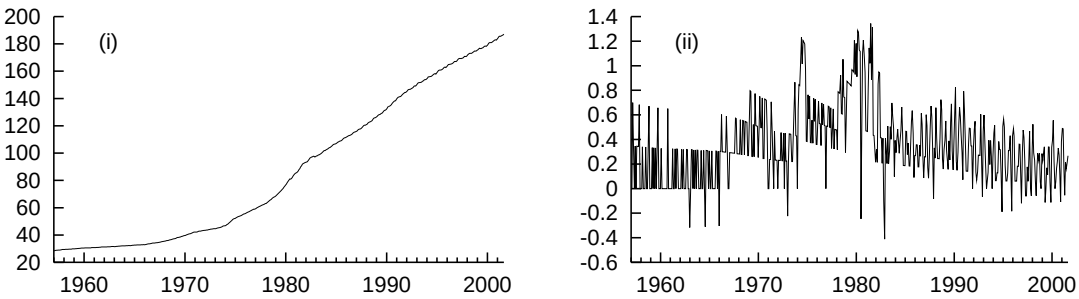

Figure 1: (i) The U.S. monthly consumer price index (CPI) and (ii) the difference of log CPI (inflation)

The estimation results reported in table 3 have been obtained by numerically maximising the loglikelihood function which is computed either by the Kalman filter only (for the fixed variance case) or by the Kalman filter and importance sampling methods described in section 3 (for the general model with a common stochastic variance). The numerical optimisation is carried out by the BFGS method as implemented in the the Ox functions library of Doornik (1999). In all cases convergence was obtained rapidly and required 10 to 50 loglikelihood evaluations depending on the specification of the model and the starting value of the parameter vector.

The parameter estimates of three model specifications are reported in table 3 together with 95% (asymmetric) confidence intervals of the estimates and some diagnostic test statistics. The final rows report the normality test of Doornik and Hansen (1994), the Box-Ljung Q test of Ljung and Box (1978) for residual correlation (with 21 degrees of freedom for the LL-Seas model, 14 degrees

[1] Source: Bureau of Labor Statistics, series `CUUR0000SA0L1E`, with the city average CPI index, all items less food and energy, not seasonally adjusted.

Table 3: Estimation results of models for the U.S. monthly inflation rate

Parameter	LL-Seas		LL-Seas-CSV		LL-Seas-CSV-δ	
q_1 (s.d. irr)	0.1579	[0.144, 0.173]	0.2493	[0.197, 0.316]	0.2892	[0.235, 0.356]
q_2 (s.d. level)	0.0474	[0.035, 0.063]	0.0554	[0.038, 0.080]	0.0538	[0.037, 0.078]
q_4 (s.d. seas)	0.0249	[0.018, 0.035]	0.0703	[0.052, 0.096]	0.0853	[0.064, 0.115]
φ (CSV)			0.9935	[0.961, 0.999]	0.9994	[0.988, 1.000]
σ_η (CSV)			0.2222	[0.159, 0.311]	0.0809	[0.040, 0.164]
$\delta_{\mathrm{L,\ 1974:2}}$					0.5605	[0.372, 0.749]
$\delta_{\mathrm{L,\ 1974:11}}$					−0.4744	[−0.667, −0.282]
$\delta_{\mathrm{I,\ 1980:7}}$					−1.2057	[−1.627, −0.785]
$\delta_{\mathrm{L,\ 1981:9}}$					−0.5260	[−0.794, −0.258]
$\delta_{\mathrm{L,\ 1982:8}}$					−0.3683	[−0.643, −0.094]
LL	85.52		146.33		179.18	
Test statistic						
Normality-DH	102.085	(0.000)	0.354	(0.838)	0.930	(0.628)
Q	54.474	(0.000)	26.204	(0.125)	29.553	(0.009)
ARCH	4.904	(0.270)	7.097	(0.229)	8.121	(0.215)
H($n/3$)	0.586	(1.000)	0.693	(0.992)	0.719	(0.985)
Time	0.97		1:03.72		3:53.88	

Parameter estimates with 95% confidence intervals of three models for U.S. inflation 1957-2001. The three models consist of a combination of level, seasonal plus irregular components (LL-Seas), common stochastic variance components (LL-Seas-CSV) and interventions, (LL-Seas-CSV-δ), respectively. The loglikelihood (LL) value and various diagnostic test statistics (with p-values) are also reported. The last row reports the estimation time.

of freedom for the LL-Seas-CSV-δ model), the ARCH-test of Engle (1982) for correlation in the squared residuals and the heteroskedasticity test statistics $H(n/3)$ together with the corresponding p-values between square brackets. More details of the test statistics have been given in section 4.

In the second column, labelled as LL-Seas, the estimation results for the local level model plus a stochastic seasonal component and with a fixed common variance are presented. The test statistics based on the standardised residuals reveal that the model does not capture all dynamic properties of the time series accurately. The residuals have relatively large values in the year 1974 and in the earlier years of the 1980s than in the other years. We therefore consider the inclusion of a common stochastic variance such that the overall variance of the model can be different for different time periods in the sample.

The third column with label LL-Seas-CSV in table 3 reports the estimated parameters for the model with a common stochastic variance. This specification produces satisfactory results and, based on the diagnostic test statistics, we conclude that a more appropriate model for the U.S. inflation is found compared to model LL-Seas. Some concern is raised by the relative large value for the Box-Ljung test but we emphasise that we only consider a simple parsimonious model for inflation. In fact, it is surprising that the inclusion of the common stochastic variance can improve the fit dramatically for such a simple model. Similar gains may be obtained when richer dynamic mean specifications for the U.S. inflation series are considered.

The autoregressive parameter ϕ of the common stochastic variance is estimated at a value of 0.9935 which is close to the unit root. This type of near-random walk behaviour does not cause a problem for the estimation method employed and may be caused by neglecting outliers and shifts in the inflation series. To investigate the robustness of the new model and to ensure that our results do not rely on a few a-typical observations we have included some dummy variables that take account of an outlier in July 1980 and some possible level shifts in the inflation series. The estimation results can be found in the column labelled as LL-Seas-CSV-δ and we observe that a further increase of the loglikelihood value is obtained. The estimated value of ϕ is even closer to unity. Further, even though all breaks are significant, the diagnostics do not indicate that the model with dummy variables for outliers and breaks is superior compared to the model without the dummy variables. We do not continue the analysis by considering other possible specifications for the common variance such as a random walk model or a stochastic spline model.

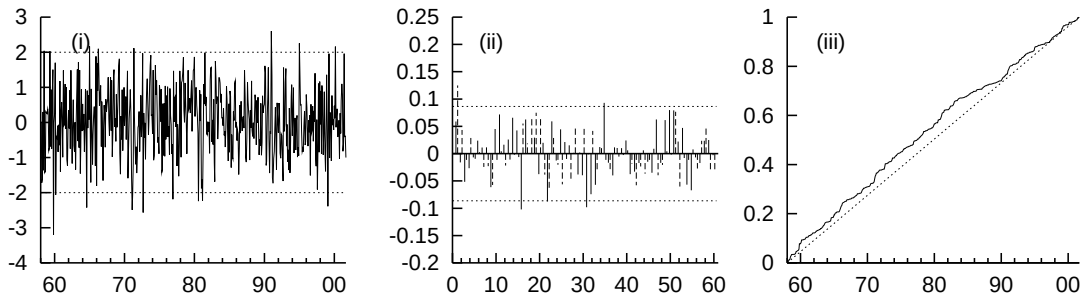

Figure 2: Prediction residuals for the estimated LL-Seas-CSV-δ model: (i) standardised residuals; (ii) autocorrelogram of residuals (continuous lines) and squared residuals; (iii) scaled cumulative sum of squared residuals.

Figure 2 presents graphs of the standardised prediction residuals together with the autocorrelation functions of the residuals and their squares and the cumulative sum (CUSUM) of the squared residuals. These graphical diagnostic tests are based on one-step ahead prediction residuals and they raise no concern with respect to the appropriateness of the LL-Seas-CSV-δ model for the inflation series. From these results, we may conclude that the common stochastic variance is an important characteristic of U.S. inflation rates. However, a more elaborate model with a richer dynamic specification for the mean equation may well give a better fit and may provide a more economic meaningful interpretation.

Figure 3 displays the estimated time-varying unobserved components trend μ_t and seasonal γ_t together with the estimated time-varying standard deviation of the observation equation $q_1\sigma_t = q_1 \exp(h_t/2)$ for the final model reported in table 3. The estimation of the components are discussed in section 3.4. In the first graph of figure 3 the jumps in the estimated level component are associated with the level shifts in the model LL-Seas-CSV-δ. We note that when the common

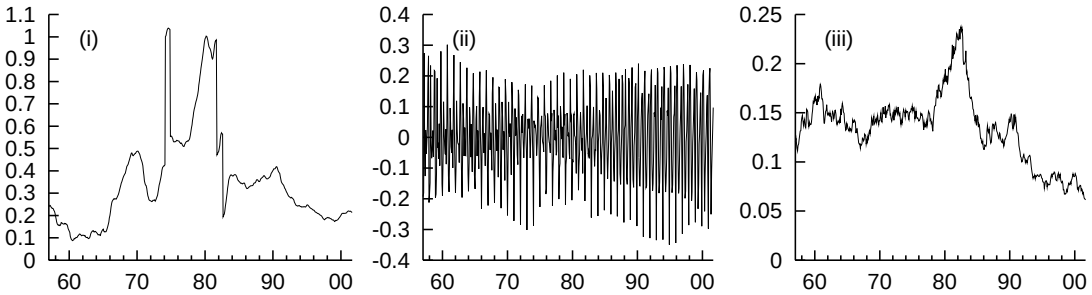

Figure 3: Estimation results for the LL-Seas-CSV-δ: (i) estimated level (including level breaks), (ii) estimated seasonal and (iii) estimated standard deviation of the irregular component (scaled by the estimated time-varying common stochastic standard deviation)

stochastic variance component is included, the model is flexible and can account for those shifts and the breaks in the model are not absolutely necessary. The seasonality in the second graph is smooth but clearly time-varying. The final graph of figure 3 presents the estimated series $q_1\hat{\sigma}_t$ which corresponds to the estimate of q_1 in the LL-Seas model (that is approximately 0.16). The common stochastic variance component displays a higher level in the early 1980s corresponding to a higher uncertainty in the early period od the 1980s which is sometimes referred to as the "second oil crisis". Further it shows a gradual reduction in the level in the early 1990s corresponding to the less volatile period associated with the Volcker-Greenspan regime at the U.S. Federal Reserve. We therefore conclude that some interesting features of the U.S. inflation series are effectively and realistically represented by the common stochastic variance in this basic model.

Figure 4 plots the observation weights for the trend plus seasonal signal $\mu_t + \gamma_t$ which are computed as discussed in section 3.4. The weights are presented for the estimated models LL-Seas (upper panels) and LL-Seas-CSV-δ (bottom panels). The panels on the left hand side are associated with time point 1981:1 and we observe that the distribution of the weights have heavier tails for the LL-Seas model with a constant variance compared to the model with a common stochastic variance. The latter model allows for the increased variability at the beginning of the 1980s and therefore the observations surrounding time point 1981:1 are more relevant than observations in more tranquile periods and are given more weight. Also the weighting pattern is asymmetric whereas for the model with a constant variance the weighting pattern is symmetric. The weights associated with time point 1995:8 are the same as for time point 1981:1 in the upper panel (apart from end-point effects) because the model has constant variances. In the bottom panel it is shown that the less weight is given to observations in the 1980s and the beginning of the 1990s whereas relative more weight is given to observations after 1995:8. This reflects that observations from 1993 onwards can be relied upon with more confidence.

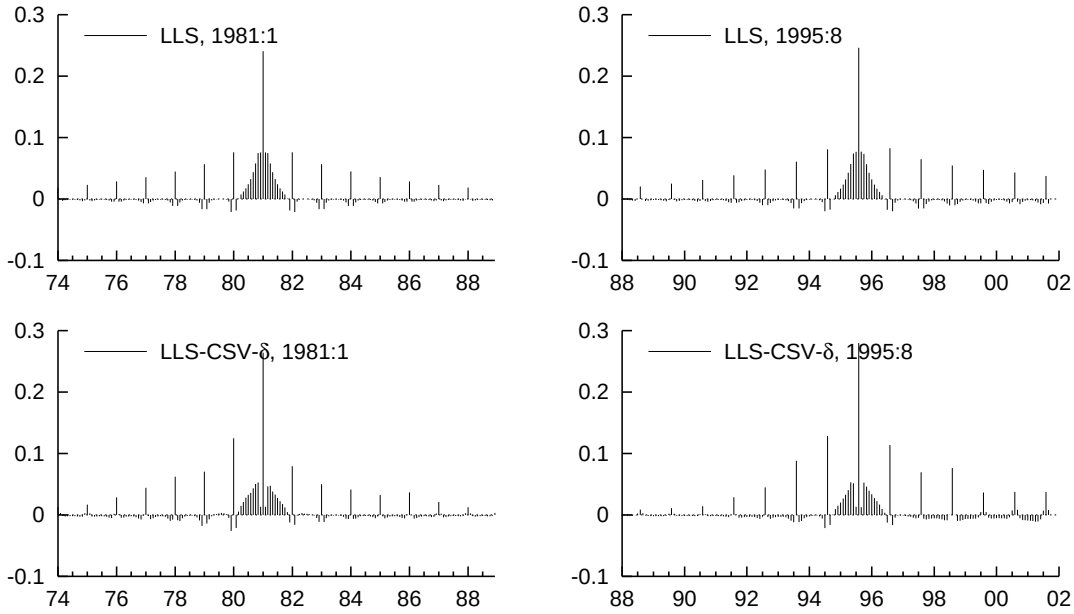

Figure 4: Observation weights for the signal of the LL-Seas and LL-Seas-CSV-δ models at periods 1981:1 and 1995:8

Table 4: Forecasting results for the inflation data

LL-Seas	1	3	6	12
MFE	−0.004	−0.007	−0.005	−0.011
MAPE	0.123	0.131	0.120	0.104
RMSE	0.154	0.163	0.149	0.133
LL-Seas-CSV	1	3	6	12
MFE	−0.006	−0.007	−0.007	−0.012
MAPE	0.108	0.110	0.102	0.099
RMSE	0.138	0.140	0.132	0.127
LL-Seas-CSV-δ	1	3	6	12
MFE	−0.004	−0.008	−0.006	−0.011
MAPE	0.109	0.112	0.105	0.101
RMSE	0.139	0.142	0.134	0.130

The table reports the mean forecast error (MFE), the mean absolute prediction error (MAPE) and the root mean squared forecast error (RMSE), at horizons of 1, 3, 6 and 12 months ahead. Forecasting starts at 1985:1.

To investigate whether the common stochastic variance component improves the forecasting performance, we have carried out the following forecasting exercise. Starting with the sample 1957:1-1985:1, the model is estimated and h-step forecasts are computed for $h = 1, 3, 6, 12$; for the details, see section 3. The four forecast errors $e_{m,h} = y_{m+h} - \hat{y}_{m+h|m}$ are computed where m refers to the final observation in the estimation sample and $\hat{y}_{t|m}$ is the forecast of y_t given the observations of the estimation sample for $t > m$. The estimation and forecasting steps are repeated for samples that are increased by one observation at a time until sample 1957:1-2000:9 is reached. Finally, some statistics are computed for the collection of forecast errors.

Table 4 reports the mean forecast error (MFE), the mean absolute prediction error (MAPE) and the root mean squared error (RMSE) which are defined by

$$\text{MFE} = \sum e_{t+h}/k, \qquad \text{MAPE} = \sum |e_{t+h}|/k, \qquad \text{RMSE} = \sqrt{\sum e_{t+h}^2/k}.$$

where k is the number of forecasts errors for horizons $h = 1, 3, 6, 12$ months. We compare results for the three inflation models as considered in table 3. The models with the common stochastic variance outperform the forecast statistics of the model with constant variance for all forecasting horizons. This is mainly due to the gradual change of the variance after 1985 which seems a salient feature of US inflation data. By allowing for these changes in the common variance of the model, more accurate forecasts can be produced. The last panel of table 4 gives the results for the LL-Seas-CSV model with the pre-1985 level breaks. These statistics are similar to the ones based on same the model without the breaks. This is due to the fact that no breaks occur in the forecasting periods.

We further checked the coverage probabilities that were implied by the forecasts for the series and the corresponding forecast variances. Even though the unconditional coverage probability and independence tests of Christoffersen (1998) have not very high power given the limited number of time periods for which we could make our forecasts, a clear indication was given that models incorporating the common stochastic variance component fitted the data significantly better according to both tests. Finally we note that for all models the MAPE and RMSE statistics are lower at horizons 6 and 12 than at horizons 1 and 3 although this is less true for the two models with a common stochastic variance. It appears that the seasonal effect is strong and therefore good forecasts can be produced at yearly horizons in particular.

5.2 US industrial production

Another macroeconomic time series of interest is U.S. industrial production[2]. The monthly time series from 1960 to 2001 is presented in levels, in logarithms and as percentages of month-to-month

[2]Industrial Production – Market Group – Total Index – Not Seasonally Adjusted. Source: Economagic, series b500001_ipnsa, Federal Reserve, Board of Governors. Available is 1919:1-2001:12 but we use only 1960:1-2001:12.

growth rates in figure 5. The time series plot of growth reveals that the variability of the series is lower after the early 1980s.

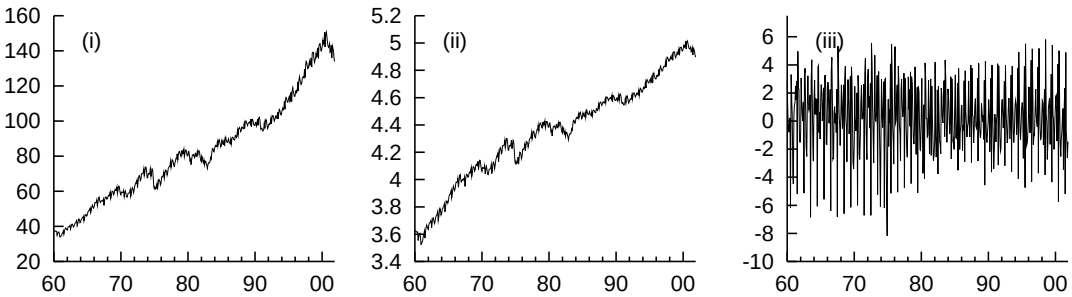

Figure 5: The U.S. monthly industrial production: (i) in levels, (ii) in logs and (iii) in percentage growth

Table 5: Estimation results for U.S. industrial growth $\Delta \log(\text{IP})$

Parameter	Seas		AR-Seas		AR-Seas-CSV	
q_1 (s.d. irr)	0.9499	[0.885, 1.019]	0.7248	[0.648, 0.811]	0.7463	[0.594, 0.938]
q_4 (s.d. seas)	0.0311	[0.024, 0.040]	0.0365	[0.029, 0.046]	0.0344	[0.025, 0.046]
q_5 (s.d. ar1)			0.5434	[0.426, 0.693]	0.4980	[0.352, 0.704]
φ (ar1)			0.8756	[0.704, 0.935]	0.9471	[0.884, 0.970]
ϕ (CSV)					0.9432	[0.810, 1.077]
σ_η (CSV)					0.2072	[0.051, 0.848]
LL	-753.89		-715.56		-699.42	
Test statistics						
Normality-DH	32.260	(0.000)	17.397	(0.000)	5.559	(0.062)
Box-Ljung Q	82.733	(0.000)	24.271	(0.186)	29.261	(0.032)
ARCH	35.894	(0.105)	17.540	(0.149)	1.829	(0.405)
H(n/3)	0.477	(1.000)	0.480	(1.000)	0.785	(0.939)
Time	0.58		1.48		1:19.34	

Parameter estimates with 95% confidence intervals of three models for U.S. monthly industrial growth 1960-2001. The three models consist of a combination of seasonal plus irregular components (Seas), autoregressive (AR-Seas) component and common stochastic variance (AR-Seas-CSV) component respectively. The loglikelihood (LL) value and various diagnostic test statistics (with p-values) are also reported. The last row reports the estimation time.

The graph of production growth shows further that seasonal variation is the most prominent feature. We therefore start to estimate an unobserved components time series model with seasonal and irregular components, that is model (9) with $\mu_t = \beta_t = 0$ (and $q_2 = q_3 = 0$). The estimation results of this simple model are reported in the column of table 5 indicated by Seas. The large Box-Ljung Q portmanteau test[3] for serial correlation indicates that the model is not appropriate for the growth series. A closer investigation shows that the serial correlation coefficient of the prediction residuals at the first lag is relatively large. Therefore an autoregressive component is

[3]The Q-statistic has 21 degrees of freedom for the smallest model, 17 degrees of freedom for the largest.

included in the model that is given by

$$y_t = \rho_t + \gamma_t + \varepsilon_{1t}, \qquad t = 1, \ldots, n,$$

where γ_t and ε_{1t} are specified in (9) and ρ_t is the autoregressive component given by

$$\rho_{t+1} = \varphi\rho_t + q_5\varepsilon_{5t}, \qquad t = 1, \ldots, n,$$

with $\rho_1 \sim \mathcal{N}\{0, q_5^2/(1-\varphi^2)\}$. The estimation results of this model for the production growth series is reported in the column of table 5 indicated by AR-Seas. The increase of the loglikelihood value and the decrease of the Box-Ljung Q test-statistic for serial correlation are substantial compared to the results of the seasonal plus irregular model. The estimate of the autoregressive φ is close to 0.9 and provides the expected persistency in U.S. growth whereas the time-variation of the seasonal component remains unchanged. The normality test statistic however raises some concern about the specification of the current model for U.S. growth in production. Some graphical diagnostics are presented in figure 6 and they are reasonable although the prediction residuals are subject to some unusual large values during the oil crises in the middle of the 1970s and the early 1980s. They appear even more clearly in the final plot of figure 6, that is the cumulative sum of squared prediction residuals.

It is observed earlier that the overall variation of growth is locally varying over time and therefore we introduce the common stochastic variance. The estimation of this model does not give numerical problems and the loglikelihood value increases by a value of almost 16 at the expense of two additional unknown parameters and more computing time. The loglikelihood ratio clearly indicates that this more elaborate model is preferable. The reported estimated parameter values show that U.S. growth in the last, say, forty years is very persistent (the estimated autoregressive coefficient for ρ_t is close to 0.95) and the persistency of the uncertainty in growth, as indicated by the autoregressive coefficient ϕ of the CSV, is also substantial given the estimated value 0.94. It is surprising to observe that the estimated confidence intervals for the seasonal component and the irregular component remain close to their estimated values for the model without CSV. This may indicate that the time-varying heteroskedasticity is most applicable to the autoregressive component.

It is shown that some improvements in the modelling of macroeconomic time series can be obtained by allowing the common variance to change stochastically over time. This is confirmed by the graphical diagnostic plots for the prediction residuals of the final estimated model presented in figure 7. The residuals are well-behaved and by considering a richer dynamic specification for the mean equation we may well reduce the remaining serial correlation in the residuals further.

Finally, the plots of the estimated components of interest in figure 8 are informative and they display the main characteristics of the dynamics in U.S. growth. Further analyses may consider an

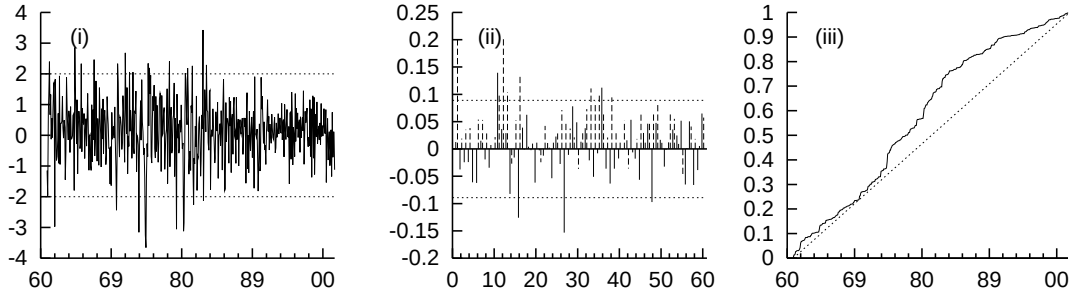

Figure 6: Prediction residuals of the estimated model "AR-Seas" with AR(1), seasonal and irregular components: (i) plot of standardised residuals; (ii) autocorrelogram of residuals (continuous lines) and squared residuals; (iii) scaled cumulative sum of squared residuals.

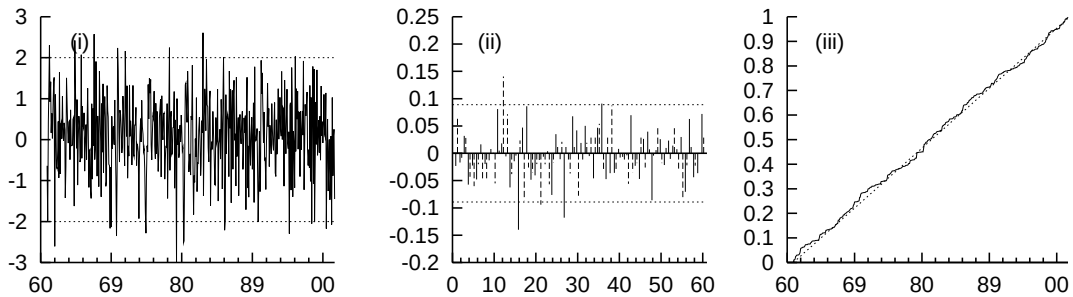

Figure 7: Residuals of the model with AR(1), seasonal, irregular and SV components: (i) plot of standardised residuals; (ii) autocorrelogram of residuals (continuous lines) and squared residuals; (iii) scaled cumulative sum of squared residuals.

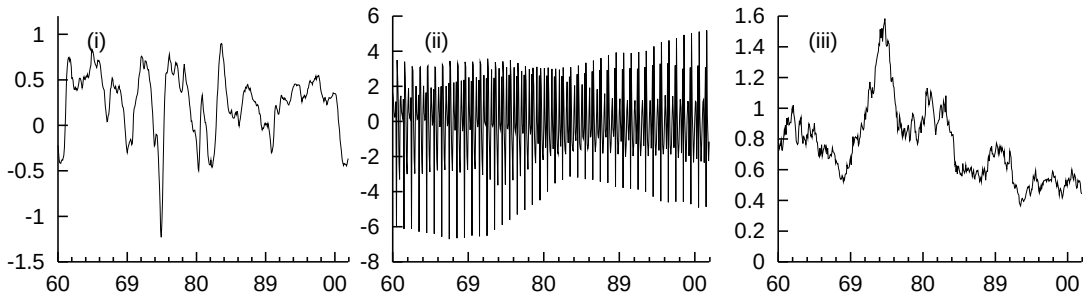

Figure 8: Estimation results for the production growth model: (i) estimated autoregression, (ii) estimated seasonal and (iii) estimated standard deviation of the irregular component (scaled by the estimated time-varying common stochastic standard deviation)

autoregressive component ρ_t that is based on a second order autoregressive process with complex roots or on a cyclical component based on time-varying trigonometric terms. The introduction of (lagged) explanatory variables may lead to an empirical model that is of more interest from an economic point of view.

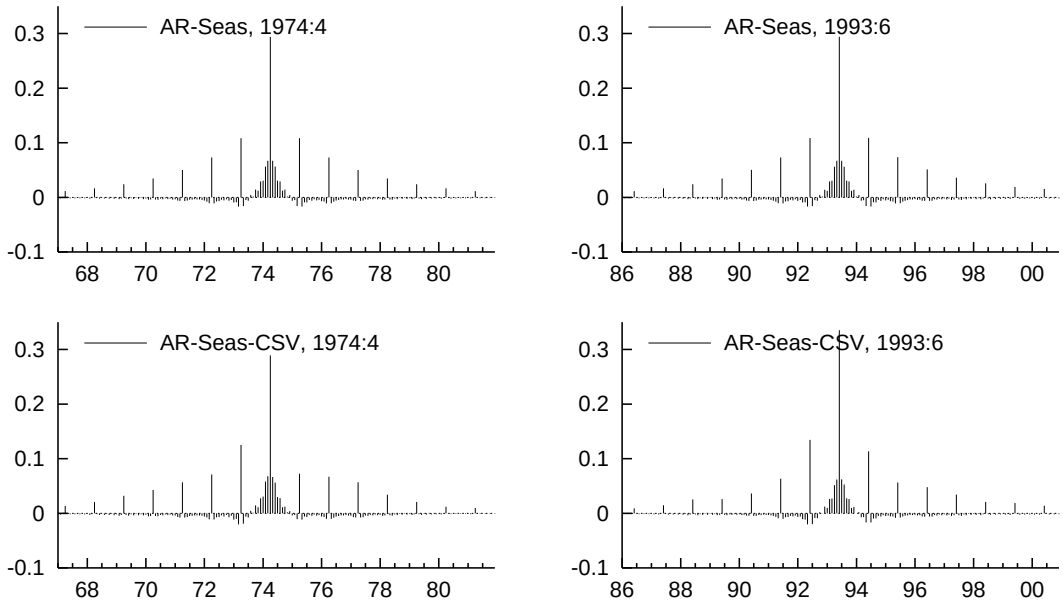

Figure 9: Weights for the AR-Seas and AR-Seas-CSV models, at the high-variability period 1974:4 and the low uncertainty period 1993:6

The observation weights of the estimated signal $\gamma_t + \rho_t$ are presented for the time points 1974:4 and 1993:6 in figure 9. The top panels display the weights for the AR-Seas model with a fixed variance and are therefore equivalent. The bottom panels present the weight patterns implied by the model with a common stochastic variance and they are different from each other but also from the associating weights of the top panels. The main difference is that the patterns of the bottom panels are not symmetric and therefore different weights are assigned to observations with an equal distance to the time index of the signal. This is caused by the (common) heteroskedasticity modelled as a stochastic function of time.

The same forecasting exercise as for US inflation is carried out for US industrial growth where forecasts are made based on the data until 1980:1 upto 2000:12. For each subsample the model with a fixed variance and with a common stochastic variance is estimated and forecasts are computed for horizons $h = 1, 3, 6, 12$. The results are reported in table 6. The mean of the forecasts errors reduces considerably for the model with a common stochastic variance leading to less bias in the forecasts. Note that the bias is still higher than in the case of US inflation. It is surprising that the variation of the forecast errors is equivalent for both models. This is probably due to the fact that

Table 6: Forecasting results for US industrial growth

AR-Seas	1	3	6	12
MFE	0.103	0.136	0.191	0.161
MAPE	0.697	0.707	0.696	0.677
RMSE	0.921	0.925	0.886	0.859
AR-Seas-CSV	1	3	6	12
MFE	0.071	0.089	0.136	0.117
MAPE	0.704	0.705	0.693	0.688
RMSE	0.930	0.928	0.893	0.881

The table reports the mean forecast error (MFE), the mean absolute prediction error (MAPE) and the root mean squared forecast error (RMSE), at horizons of 1, 3, 6 and 12 months ahead. Forecasting starts at 1980:1.

the main feature of US growth is described only by the seasonal component. The autoregressive component and the heteroskedastic nature of the time series seem less important.

6 Conclusions

In this paper, we have considered the possibility of simultaneously modelling both the mean equation and the common stochastic variance equation within a general class of linear time series models. The state space framework is flexible and useful in, for example, dissecting time series into unobserved components so that level, slope and seasonal effects in a time series can be analysed separately. With the inclusion of a stochastic volatility component into the model, further flexibility is attained that can be useful for the modelling of macroeconomic time series.

After an introduction of the general model together with a discussion of some special cases in section 2, a short theoretical exposition is given in section 3. A Monte Carlo experiment is carried out to show the small sample properties of the estimates produced by the importance sampling estimation method. These simulation based maximum likelihood estimates are obtained subject to simulation error which can be made arbitrarily small by enlarging the number of simulations.

The methodology is illustrated by modelling a monthly time series of U.S. inflation between 1957–2001 (537 observations). It is found that a simple local level model with a seasonal component and a common stochastic variance is appropriate and that it can incorporate some important dynamic features of inflation that is difficult to model since the series has been subject to some shocks in the last forty years. The autoregressive common log-variance has been able to capture

(i) the volatile period of the 1970s and early 1980s associated with oil crises and (ii) the less volatile period after the end of the 1980s associated with the Volcker-Greenspan regime at the U.S. Federal Reserve. The role that the common stochastic variance play in the analysis is illustrated by presenting the observations weights associated with the estimated signal of the time series and by presenting forecasting results. We show that US inflation forecasts that are based on a simple descriptive model can be improved by incorporating a common stochastic variance component.

A similar analysis is carried out for the monthly percentage growth of US industrial production in the years 1960 to 2001. The in-sample fit of the model with a common stochastic variance component improves significantly although the forecast results are less affected by the extension. This may be due to the highly seasonal nature of the time series.

The framework of a common stochastic variance can be exploited into other directions. In this paper we have focused on stochastic heteroskedasticity as a function of time although it can also be related to other characteristics of the observation y_t. For example, the actual value of the observation or the value of another variable associated with y_t may be a source of heteroskedasticity. Let us assume that heteroskedasticity is a stochastic function of the variable x_t associated with y_t. The treatment of a common stochastic variance as a function of x_t is similar to the method described in section 3. The Kalman filter is used to obtain innovations based on the state space model (3) conditional on the observations and on a given value of σ^*. It is noticed that the innovations are independently distributed. An intermediate step is to re-order the innovations v_t and their scaled variances F_t subject to the magnitude of, say, x_t. Denote these re-ordered innovations by v_j^o and their variances by F_j^o. It is still valid to treat the re-ordered innovations as being generated by the model

$$v_j^o \sim \mathcal{N}(0, \sigma_j^{o2}), \qquad j = 1, \ldots, n.$$

We then can use model (5) for the re-ordered quantities (so using index j instead of time index t) and obtain a Monte Carlo estimate of the loglikelihood function since the equality $p(y|\psi) = p(v^o|\psi)$ where v^o is the stack of the re-ordered innovations still applies. Further research is required to investigate the effectiveness of this development but this short exposition reveals the flexibility of the proposed methodology for modelling stochastic heteroskedasticity.

References

Anderson, B. D. O. and J. B. Moore (1979). *Optimal Filtering*. Englewood Cliffs: Prentice-Hall.

Bos, C. S., R. J. Mahieu, and H. K. van Dijk (2000). Daily exchange rate behaviour and hedging of currency risk. *J. Applied Econometrics 15*(6), 671–696.

Bowman, K. O. and L. R. Shenton (1975). Omnibus test contours for departures from normality based on $\sqrt{b_1}$ and b_2. *Biometrika 62*, 243–250.

Brockwell, P. J. and R. A. Davis (1987). *Time Series: Theory and Methods*. New York: Springer-Verlag.

Christoffersen, P. F. (1998). Evaluating interval forecasts. *International Economic Review 39*(4), 841–862.

Danielson, J. (1994). Stochastic volatility in asset prices: Estimation with simulated maximum likelihood. *Journal of Econometrics 61*, 375–400.

De Jong, P. (1989, December). Smoothing and interpolation with the state-space model. *J. American Statistical Association 84*(408), 1082–1088.

Doornik, J. A. (1999). *Object-Oriented Matrix Programming using Ox* (3rd ed.). London: Timberlake Consultants Ltd. See `http://www.nuff.ox.ac.uk/Users/Doornik`.

Doornik, J. A. and H. Hansen (1994). An omnibus test for univariate and multivariate normality. Technical report, Nuffield College, Oxford, UK.

Doucet, A., J. F. G. deFreitas, and N. J. Gordon (Eds.) (2000). *Sequential Monte Carlo methods in practice*. New York: Springer-Verlag.

Durbin, J. and S. J. Koopman (1997). Monte Carlo maximum likelihood estimation of non-Gaussian state space model. *Biometrika 84*, 669–84.

Durbin, J. and S. J. Koopman (2001). *Time Series Analysis by State Space Methods*. Oxford: Oxford University Press.

Durbin, J. and S. J. Koopman (2002). A simple and efficient simulation smoother for state space time series analysis. *Biometrika 89*, 603–616.

Durham, G. B. and A. Gallant (2002). Numerical techniques for maximum likelihood estimation of continuous-time diffusion. *J. Business and Economic Statist. 20*(3), 297–316.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflations. *Econometrica 50*, 987–1008.

Engle, R. F. and A. D. Smith (1999). Stochastic permanent breaks. *Review of Economics and Statistics 81*(4), 553–574.

Gallant, A. R., D. Hsieh, and G. Tauchen (1997). Estimation of stochastic volatility models with diagnostics. *Journal of Econometrics 81*(1), 159–192.

Geweke, J. (1989). Bayesian inference in econometric models using Monte Carlo integration. *Econometrica 57*(6), 1317–1339.

Ghysels, E., A. C. Harvey, and E. Renault (1996). Stochastic volatility. In G. Maddala and C. Rao (Eds.), *Handbook of Statistics*, Volume 14, Statistical Methods in Finance, pp. 119–191. North-Holland, Amsterdam.

Hamilton, J. (1994). *Time Series Analysis*. Princeton: Princeton University Press.

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

Harvey, A. C. and S. J. Koopman (2000). Signal extraction and the formulation of unobserved components models. *Econometrics Journal 3*, 84–107.

Harvey, A. C., E. Ruiz, and N. Shephard (1994). Multivariate stochastic variance models. *Rev. Economic Studies 61*, 247–264.

Jacquier, E., N. G. Polson, and P. E. Rossi (1994). Bayesian analysis of stochastic volatility models. *J. Business and Economic Statist. 12*, 371–417.

Kim, C. J. and C. R. Nelson (1999). *State Space Models with Regime Switching*. Cambridge, Massachusetts: MIT Press.

Kim, S., N. Shephard, and S. Chib (1998). Stochastic volatility: Likelihood inference and comparison with ARCH models. *Rev. Economic Studies 64*, 361–393.

Kloek, T. and H. K. Van Dijk (1978). Bayesian estimates of equation system parameters: An application of integration by Monte Carlo. *Econometrica 46*, 1–20.

Koopman, S. J. and N. Shephard (2002). Testing the assumptions behind the use of importance sampling. Discussion paper, Nuffield College, Oxford.

Koopman, S. J., N. Shephard, and J. A. Doornik (1999). Statistical algorithms for models in state space using SsfPack 2.2. *Econometrics Journal 2*, 107–160.

Ljung, G. M. and G. E. P. Box (1978). On a measure of lack of fit in time series models. *Biometrika 66*, 67–72.

Nabeya, S. and K. Tanaka (1988). Asymptotic theory of a test for the constancy of regression coefficients against the random walk alternative. *Annals of Statistics 16*, 218–235.

Ord, J. K., A. B. Koehler, and R. D. Snyder (1997). Estimation and prediction for a class of dynamic nonlinear statistical models. *J. American Statistical Association 92*, 1621–1629.

Sandmann, G. and S. J. Koopman (1998). Estimation of stochastic volatility models via Monte Carlo maximum likelihood. *Journal of Econometrics 87*, 271–301.

Schweppe, F. (1965). Evaluation of likelihood functions for Gaussian signals. *IEEE Transactions on Information Theory 11*, 61–70.

Shephard, N. (1994). Partial non-Gaussian state space. *Biometrika 81*(1), 115–131.

Shephard, N. (1996). Statistical aspects of ARCH and stochastic volatility. In D. Cox, D. Hinkley, and O. Barndorff-Nielsen (Eds.), *Time Series Models in Econometrics, Finance and Other Fields*, Number 65 in Monographs on Statistics and Applied Probability, pp. 1–67. Chapman and Hall, London.

Shephard, N. and M. K. Pitt (1997). Likelihood analysis of non-Gaussian measurement time series. *Biometrika 84*, 653–667.

Shumway, R. H. and D. S. Stoffer (2000). *Time Series Analysis and its Applications*. Springer Texts in Statistics. New York: Springer.

Taylor, S. J. (1994). Modeling stochastic volatility: A review and comparative study. *Mathematical Finance 4*(2), 183–204.

Table 7: Small sample results for the Local Level model

		Generating no SV											
		$n = 100$				$n = 500$				$n = 1000$			
Estimating		No SV		SV		No SV		SV		No SV		SV	
	p	$\overline{p}$	s.d.	$\overline{p}$	s.d.	$\overline{p}$	s.d.	$\overline{p}$	s.d.	$\overline{p}$	s.d.	$\overline{p}$	s.d.
q_1	1.0	1.00	0.10	0.98	0.11	1.00	0.04	0.99	0.04	1.00	0.03	0.99	0.03
q_2	0.5	0.48	0.11	0.48	0.11	0.50	0.05	0.49	0.05	0.50	0.04	0.50	0.03
σ_η				0.15	0.23			0.11	0.13			0.08	0.11
ϕ				0.38	0.31			0.34	0.30			0.35	0.28
		Stat	f_{Rej}	Stat	f_{Rej}	Stat	f_{Rej}	Stat	f_{Rej}	Stat	f_{Rej}	Stat	f_{Rej}
Normality-DH		1.96	0.04	1.31	0.02	1.99	0.05	1.29	0.01	2.01	0.05	1.52	0.03
H($n/3$)		1.12	0.33	1.05	0.29	1.01	0.14	1.02	0.12	1.00	0.04	1.00	0.02
LL		-164.12		-164.20		-830.95		-831.98		-1663.95		-1664.71	
		Generating SV											
		$n = 100$				$n = 500$				$n = 1000$			
Estimating		No SV		SV		No SV		SV		No SV		SV	
	p	$\overline{p}$	s.d.	$\overline{p}$	s.d.	$\overline{p}$	s.d.	$\overline{p}$	s.d.	$\overline{p}$	s.d.	$\overline{p}$	s.d.
q_1	1.0	1.05	0.15	0.98	0.14	1.05	0.07	1.00	0.06	1.05	0.05	1.00	0.05
q_2	0.5	0.51	0.13	0.49	0.13	0.52	0.06	0.50	0.06	0.52	0.04	0.50	0.04
σ_η	0.2			0.23	0.22			0.22	0.10			0.20	0.07
ϕ	0.9			0.68	0.33			0.85	0.13			0.88	0.08
		Stat	f_{Rej}	Stat	f_{Rej}	Stat	f_{Rej}	Stat	f_{Rej}	Stat	f_{Rej}	Stat	f_{Rej}
Normality-DH		3.86	0.19	1.49	0.02	9.42	0.51	2.42	0.06	17.92	0.80	3.75	0.17
H($n/3$)		1.20	0.37	1.02	0.27	1.08	0.28	1.00	0.03	1.03	0.20	1.00	0.00
LL		-168.64		-165.79		-856.56		-850.34		-1713.49		-1699.89	

See table 2 for a description of the entries in the table.

Table 8: Small sample results for the Local Level plus Seasonal model

		Generating no SV											
		$n = 100$				$n = 500$				$n = 1000$			
Estimating		No SV		SV		No SV		SV		No SV		SV	
	p	$\overline{p}$	s.d.	$\overline{p}$	s.d.	$\overline{p}$	s.d.	$\overline{p}$	s.d.	$\overline{p}$	s.d.	$\overline{p}$	s.d.
q_1	1.0	0.99	0.15	0.98	0.18	0.99	0.06	0.99	0.06	1.00	0.04	0.99	0.04
q_2	0.5	0.48	0.12	0.48	0.13	0.50	0.05	0.49	0.05	0.50	0.03	0.50	0.03
q_4	0.2	0.20	0.05	0.19	0.06	0.20	0.02	0.20	0.02	0.20	0.01	0.20	0.01
σ_η				0.18	0.24			0.11	0.13			0.09	0.11
ϕ				0.39	0.34			0.33	0.31			0.29	0.28
		Stat	f_{Rej}	Stat	f_{Rej}	Stat	f_{Rej}	Stat	f_{Rej}	Stat	f_{Rej}	Stat	f_{Rej}
Normality-DH		2.04	0.05	1.24	0.01	1.91	0.05	1.35	0.02	1.89	0.05	1.46	0.03
H($n/3$)		1.06	0.14	1.02	0.11	1.02	0.01	0.99	0.00	1.01	0.00	1.00	0.00
LL		-182.91		-182.73		-939.40		-938.98		-1884.43		-1882.91	

		Generating SV											
		$n = 100$				$n = 500$				$n = 1000$			
Estimating		No SV		SV		No SV		SV		No SV		SV	
	p	$\overline{p}$	s.d.	$\overline{p}$	s.d.	$\overline{p}$	s.d.	$\overline{p}$	s.d.	$\overline{p}$	s.d.	$\overline{p}$	s.d.
q_1	1.0	1.03	0.21	1.12	0.52	1.04	0.08	1.01	0.08	1.05	0.06	1.00	0.06
q_2	0.5	0.51	0.13	0.56	0.31	0.53	0.06	0.51	0.06	0.53	0.04	0.51	0.04
q_4	0.2	0.20	0.06	0.22	0.12	0.21	0.02	0.20	0.03	0.21	0.02	0.20	0.02
σ_η	0.2			0.24	0.23			0.19	0.09			0.17	0.06
ϕ	0.9			0.70	0.32			0.87	0.13			0.90	0.07
		Stat	f_{Rej}	Stat	f_{Rej}	Stat	f_{Rej}	Stat	f_{Rej}	Stat	f_{Rej}	Stat	f_{Rej}
Normality-DH		3.18	0.15	1.33	0.02	7.87	0.39	1.65	0.02	13.05	0.63	2.38	0.06
H($n/3$)		1.21	0.28	0.96	0.07	1.05	0.10	0.98	0.00	1.03	0.03	0.99	0.00
LL		-186.52		-184.97		-962.74		-959.00		-1937.07		-1922.74	

See table 2 for a description of the entries in the table.