Quaderni di Statistica
Vol. 1, 1999

A state space framework for forecasting non-stationary economic time series

Giuseppe Storti
Dipartimento di Scienze Economiche, Università degli Studi di Salerno
E-mail: storti@unisa.it

Summary: Aim of this paper is to present a state space approach to forecasting complex non-stationary economic time series with conditional heteroskedasticity and a conditional mean decomposed into local polynomial trend, stationary autoregressive, seasonal and irregular noise components. The order of the model is chosen to minimize the value of the Akaike's Information Criterion while the unknown parameters are estimated maximizing a Gaussian log-likelihood function in the classical prediction error decomposition form. The proposed model is applied to study a series of U.S. monthly inflation data from 1971.01 to 1999.10.

Keywords: State Space Models, Kalman Filter, Conditional Heteroskedasticity, Seasonal Adjustment, Inflation

1. Introduction

The mean of a wide class of non-stationary economic time series $\{ y_t,\ t=1,...,T\}$ can be typically decomposed into a trend c_t, a stationary component w_t, a seasonal component s_t, a trading day effect d_t and an error or noise component e_t

$$y_t = c_t + s_t + w_t + d_t + e_t \tag{1}$$

with e_t being a zero mean, serially uncorrelated white noise series. Equation (1) has a general interpretation as a signal plus noise decomposition

$$y_t = SE_t + e_t$$

where the signal SE_t is given by the sum of a cyclical component CT_t and a seasonal-periodic component ST_t, including the trading day and seasonal effects,

$$SE_t = CT_t + ST_t$$

The cyclical component CT_t, in turn, is given by the sum of a long term structural component $STR_t = c_t$ and a short-medium term component $AR_t = w_t$

$$CT_t = STR_t + AR_t$$

A further generalization of this framework is to allow the noise component e_t to have a time varying conditional variance or *volatility*,

$$(e_t | \mathbf{e}^{t-1}) \sim (0, h_t^2) \qquad \textit{with } \mathbf{e}^{t-1} = \{ e_1, \ldots, e_{t-1} \}.$$

This work proposes an integrated approach to forecasting non-stationary economic time series admitting a decomposition of type (1) in which the error term is Conditionally Heteroskedastic (CH).

The structure of the paper is as follows. Section 2 illustrates a class of state space models for structural time series of the type described in (1). In section 3 a state space model for the estimation of the conditional variance in CH time series is presented. Section 4 unifies these two approaches in order to propose a new model which allows for simultaneous modelling of the conditional mean components and of the time varying conditional variance. The proposed approach is fully developed in a state space setting. The structural components of the series and the time varying conditional variance are estimated by Kalman filtering while the unknown hyperparameters in the model are estimated maximizing a Gaussian log-likelihood function via the EM algorithm. Throughout the paper, the results of some applications to series of monthly and annual inflation data are shown. Some concluding remarks are given in section 5.

2. A state space model for forecasting and seasonal adjustment

2.1. Modelling issues

Let $\{ y_t,\ t=1,...,T\}$ be an univariate time series admitting a structural decomposition of the type described in eq. (1). A general state space representation for y_t can be obtained through the model

$$y_t = \boldsymbol{C}_t \boldsymbol{x}_t + e_t \qquad (2a)$$

$$\boldsymbol{x}_t = \boldsymbol{A}\boldsymbol{x}_{t-1} + \boldsymbol{D}\boldsymbol{q}_t \qquad (2b)$$

where $\boldsymbol{C}_t$, $\boldsymbol{A}$ and $\boldsymbol{D}$ are $(1\times k)$, $(k\times k)$ and $(k\times m)$ matrices, respectively, $\boldsymbol{x}_t$ is a k-dimensional state vector and e_t (1×1) and $\boldsymbol{q}_t$ $(m\times 1)$ are Gaussian white noise series of mutually independently distributed errors with known variances $Var(e_t) = \sigma_e^2$ and $Var(\boldsymbol{q}_t) = \boldsymbol{Q}$. Also, the error series e_t and $\boldsymbol{q}_t$ are assumed to be uncorrelated with the initial state $\boldsymbol{x}_0$, $\boldsymbol{E}(\boldsymbol{q}_t \boldsymbol{x}_0^{'}) = \boldsymbol{0}$ and $\boldsymbol{E}(\boldsymbol{x}_0 e_t) = \boldsymbol{0}$, $\forall t$, with $\boldsymbol{x}_0 \sim N(\boldsymbol{m}_0, \boldsymbol{P}_0)$. The transition matrix $\boldsymbol{A}$ and the observation matrix $\boldsymbol{C}_t$ have a block structure. The blocks in $\boldsymbol{C}_t$ pick up the relevant components of the state vector while each block in $\boldsymbol{A}$ determines the dynamic behaviour of the corresponding component in (1).

Kitagawa and Gersch (1984), henceforth KG, propose the following model including local polynomial trend, stationary AR and seasonal components, trading day effect and observation error

$$y_t = \begin{bmatrix} C_1 & C_2 & C_3 & C_{4,t} \end{bmatrix} x_t + e_t \tag{3a}$$

$$x_t = \begin{bmatrix} A_1 & 0 & 0 & 0 \\ 0 & A_2 & 0 & 0 \\ 0 & 0 & A_3 & 0 \\ 0 & 0 & 0 & A_4 \end{bmatrix} x_{t-1} + \begin{bmatrix} D_1 \\ D_2 \\ D_3 \\ D_4 \end{bmatrix} q_t \tag{3b}$$

Model (3) admits an orthogonal decomposition into four different component models $\{ C_j, A_j, D_j \}$, with j=1,...,4, each corresponding to one of the structural components in (1). The trend model assumes that c_t satisfies the stochastic difference equation

$$\nabla^q c_t = q_{1,t} \tag{4}$$

The updating equation for the autoregressive component is

$$w_t = \sum_{i=1}^{p} \varphi_i w_{t-i} + q_{2,t} \tag{5}$$

while the seasonal model is assumed to satisfy the condition

$$s_t = -\sum_{i=1}^{S-1} s_{t-i} + q_{3,t} \tag{6}$$

obtained as a stochastic generalization of the usual constraint $\sum_{i=0}^{S-1} s_{t-i} = 0$, with S being the seasonal period. Similarly, the model for the trading day effect is derived from the constraint $\sum_{i=1}^{7} \gamma_{i,t} = 0$ and defined as

$$d_t = \sum_{i=1}^{6} \gamma_{i,t} d_{i,t} \qquad (7)$$

where $\gamma_{i,t}$ is the trading day effect factor and $d_{i,t}$ is equal to the number of i-th days of the week minus the number of 7-th days of the week in the tth period. The transition equation for the factors $\gamma_{i,t}$ directly follows from the nonperturbed difference equation constraint $\gamma_{i,t} = \gamma_{i,t-1}$, $i=1,...,6$.

The Kalman filter algorithm (Kalman, 1960) can be used to obtain a Minimum Mean Square Error (MMSE) estimate of the expected value of the state *predictive* density $p(\mathbf{x}_{t+k} \mid \mathbf{y}^t)$ and of the *smoothing* density $p(\mathbf{x}_t \mid \mathbf{y}^T)$. The estimation of the *smoothing* density is essential for seasonal adjustment and, in general, in any signal estimation problem.

2.2. *An application to monthly inflation data*

As an illustration, I show the results of an application of the KG modelling procedure to the analysis of the U.S. monthly inflation rate from January 1971 to October 1999, measured as the first difference of the log-transformed U.S. Consumer Price Index (base 1982-1984=100) for all urban consumers (CPI-U). This index differs from the CPI for Urban Wage Earners and Clerical Workers (CPI-W) by including the buying patterns of all urban households regardless of the consumer units' occupational status. The data have been plotted in Fig. 1.

For the identification and the estimation of the model, I use only observations up to December 1998 while the last ten observations are left for out of sample forecast evaluation. The model considered includes a trend, a seasonal (S=12) and an autoregressive component with no trading day effect. Following Kitagawa and Gersch (1984), the orders of the polynomial trend and of the autoregressive component have been chosen to minimize the value of the Akaike Information Criterion (AIC). On the basis of the results of the exploratory analyses performed, including preliminary ARIMA modelling of the series, the search has been restricted within the ranges $q=\{1,2\}$, for the trend, and $0 \leq p \leq 7$ for the autoregressive component.

Figure 1. First difference of the log-transformed U.S. CPI-U from 1971.01 to 1999.10 (Source: U.S. Bureau of Labour Statistics).

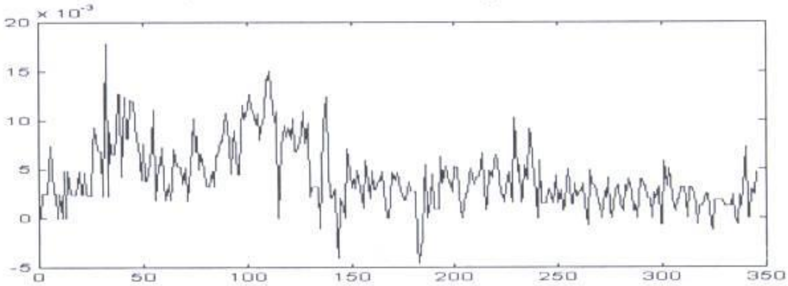
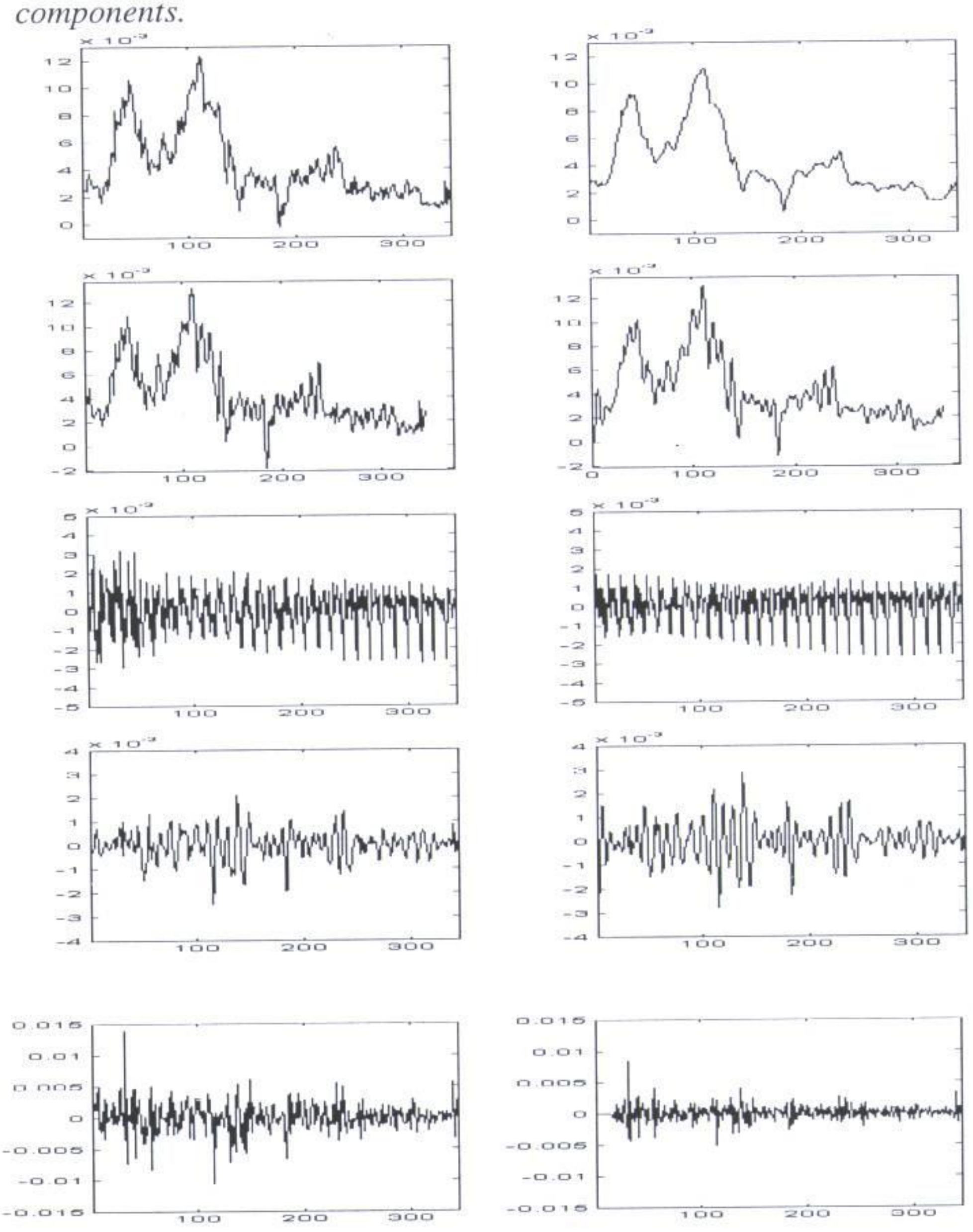


Table 1. AICs of Trend plus Seasonal plus AR models

AR order	*Trend order q=1*	*Trend order q=2*
0	-9.1013	-8.8551
1	-9.1027	-8.8577
2	-9.1057	-8.8667
3	**-9.1224**	-8.8764
4	-9.1154	-8.8662
5	-9.0990	-8.8521
6	-9.0924	-8.8373
7	-9.0829	-8.8193

The AIC (Table 1) indicates a model with a polynomial trend of order $q=1$ and an AR(3) component. The Maximum Likelihood (ML) estimates of the model parameters and the corresponding asymptotic standard errors have been reported in Table 2. For the numerical maximization of the Gaussian likelihood function, I have used a version of the EM algorithm tailored for state-space models by Wu *et al.* (1996). The initial value of the state has been assumed fixed, with $\boldsymbol{P_0} = \boldsymbol{0}$, and its components have been estimated by the EM algorithm as extra parameters in the model. Figure 2 shows the *predicted* and *smoothed* model components.

Figure 2. First difference of the log-transformed U.S. CPI-U. From left to right and from top to bottom: predicted and smoothed cycle-trend, cycle-trend + AR, seasonal, autoregressive and irregular components.



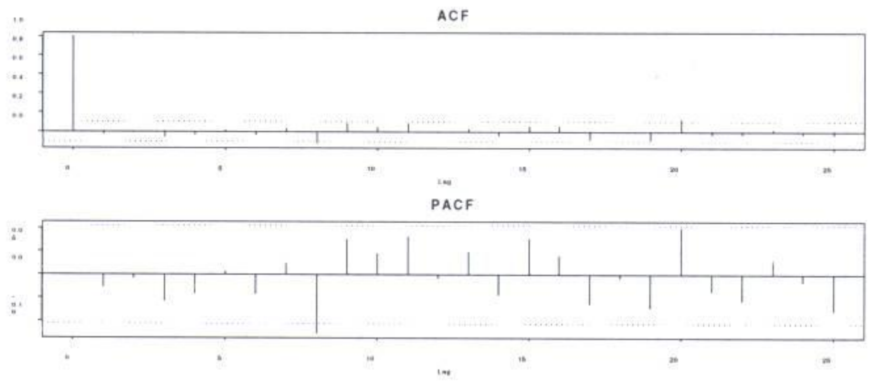
The former have been estimated using only information up to time (*t-1*) with a single Kalman filter run while the latter have been estimated by *Fixed Interval Smoothing (FIS)* using the whole set of information up to time (*T*). The predicted components are of interest for forecasting while the smoothed estimates should be considered in a time series decomposition problem (e.g. seasonal adjustment).

Table 2. Parameter estimates and asymptotic standard errors (in parentheses) for the model with trend of order 1 and AR(3) component

ϕ_1	ϕ_2	ϕ_3	var($q_{1,t}$)
0.7783	0.1537	-0.5577	4.2548×10^{-7}
(0.0500)	(0.0483)	(0.0394)	(1.2918×10^{-7})
var($q_{2,t}$)	var($q_{3,t}$)	σ_e^2	
3.1909×10^{-7}	3.9197×10^{-8}	2.9003×10^{-6}	
(1.0713×10^{-7})	(2.1579×10^{-8})	(3.1464×10^{-7})	

The two high peaks in the trend component are representative of the effects of the 1973 and 1979 oil shocks on the level of the inflation rate while the dip at observation 183 can be related to the 1986 oil price decrease.

Figure 3. Noise global and partial autocorrelation functions for the model with trend of order 1, seasonal and AR(3) component.



As shown in Fig. 3, the global and partial autocorrelation functions of the predicted noise component are not significantly different from 0 at any lag.

3. A state space approach to the estimation of conditional variance

3.1. The class of CPV models

This section introduces a class of state space models for the estimation of the conditional variance in CH time series. Let u_t be an univariate series of prediction errors such that $(u_t|\boldsymbol{u}^{t-1}) \sim N(0,h_t^2)$ and $Cov(u_t, u_{t-d}) = 0$, $\forall d \neq 0$. Also assume that u_t has finite moments up to the fourth order. A Changing Parameters Volatility (CPV) model (Storti, 1999) of order (r,s), with r and s integers, is defined as

$$u_t = \sum_{i=1}^{r} a_{i,t} u_{t-i} + \sum_{j=1}^{s} b_{j,t} h_{t-j} + e_t = \boldsymbol{C}_t^* \boldsymbol{x}_t^* + e_t \tag{8a}$$

$$\boldsymbol{x}_t^* = \boldsymbol{A}^* \boldsymbol{x}_{t-1}^* + \boldsymbol{q}_t^* \tag{8b}$$

where e_t is a Gaussian white noise observation error $e_t \sim N(0,\sigma_e^2)$, $\boldsymbol{q}_t^*$ $(n \times 1)$, with $n=r+s$, is a Gaussian serially uncorrelated system error, $\boldsymbol{q}_t^* \sim N(\boldsymbol{0},\boldsymbol{Q}^*)$, and $\boldsymbol{x}_t^*$ $(n \times 1)$ is an n-dimensional state vector with state variables given by the stochastically varying parameters $a_{i,t}$ $(i=1,\ldots,r)$ and $b_{j,t}$ $(j=1,\ldots,s)$. The observation matrix is

$$\boldsymbol{C}_t^* = [u_{t-1},\ldots,u_{t-r} \mid h_{t-1},\ldots,h_{t-s}]$$

while $\boldsymbol{A}^*$ is an $(n \times n)$ transition matrix of unknown coefficients. The specification of the model is completed by the usual assumptions

$$E(q_t^* e_z) = 0 \ , \ \forall\{t,z\}$$

$$x_0^* \sim N(m_0^*, P_0^*) \ with \ E[q_t^*(x_0^*)'] = 0 \ \ and \ E(x_0^* e_t) = 0 \ , \ \forall t$$

Under the above assumptions the model is *conditionally Gaussian* and the Kalman filter can be used to obtain a MMSE estimate of the state vector. The conditional variance is recursively estimated as

$$h_t^2 = C_t^* P_{t|t-1}^* (C_t^*)' + \sigma_e^2 \tag{9}$$

where $P_{t|t-1}^* = Var(x_t^* \mid u^{t-1})$. The conditional variance equation (9) can be also written as

$$\begin{aligned} h_t^2 = \sigma_e^2 &+ \sum_{i=1}^{r} p_{i,i(t)} u_{t-i}^2 + \sum_{j=r+1}^{r+s} p_{j,j(t)} h_{t-(j-r)}^2 + \sum_{i=1}^{r}\sum_{j=1}^{r} p_{i,j(t)} u_{t-i} u_{t-j} \\ &+ \sum_{i=r+1}^{s}\sum_{j=r+1}^{s} p_{i,j(t)} h_{t-(i-r)} h_{t-(j-r)} + \sum_{i=1}^{r}\sum_{j=r+1}^{s} p_{i,j(t)} u_{t-i} h_{t-(j-r)} \end{aligned} \tag{10}$$

with $p_{i,j(t)}$ being the element of place (i,j) in $P_{t|t-1}^*$. Compared to conventional approaches, the CPV model has two main advantages. First, it allows for time varying parameters in the conditional variance specification (10). Second, interaction terms between past innovations and volatilities are easily included in the model. The choice $A^* = 0$ yields a more parsimonious random coefficient version of the CPV model that we will call the *constrained CPV* model or, abbreviated, CPV-C. In a CPV-C model the conditional variance parameters are constant but the interaction terms are still present. It can be shown (Storti, 1999) that, if $A^* = 0$ and the covariance matrix Q^* is diagonal, the resulting CPV-C model will have the same conditional variance as a Generalized Autoregressive Conditionally Heteroskedastic (GARCH) model (Bollerslev, 1986) of the same order. Similarly, for $s=0$ the model is a random coefficient autoregressive model of order r. A generalization of model (8) is given by the *regression CPV* model

$$y_t = \boldsymbol{M}_t \boldsymbol{\beta} + u_t = M_t \boldsymbol{\beta} + \boldsymbol{C}_t^* \boldsymbol{x}_t^* + e_t \qquad (11a)$$

$$\boldsymbol{x}_t^* = \boldsymbol{A}^* \boldsymbol{x}_{t-1}^* + \boldsymbol{q}_t^* \qquad (11b)$$

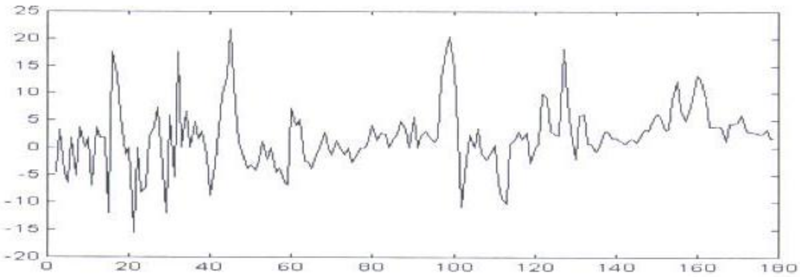
where y_t is an observed time series, $\boldsymbol{M}_t$ is a vector $(1 \times g)$ of endogenous or exogenous regressors and $\boldsymbol{\beta}$ a vector $(g \times 1)$ of unknown parameters. Again, as in section 2, the model parameters $\boldsymbol{A}^*$, $\boldsymbol{Q}^*$, σ_e^2 and $\boldsymbol{\beta}$ can be estimated maximizing a Gaussian log-likelihood function expressed in the classical prediction error decomposition form. Finally, it is worth noting that, when forecasting from a CPV model, the conditional variance h_t^2 affects the estimate of the conditional mean $E(u_t \mid \boldsymbol{u}^{t-1})$ in two different ways. First, h_t^2 enters the state updating equation, second, the estimated $E(\boldsymbol{x}_t^* \mid \boldsymbol{u}^{t-1})$ does not necessarily have to be equal to zero.

3.2. An application to the estimation of the conditional variance of the U.S. annual inflation rate

In order to illustrate the actual modelling procedure, the CPV regression model is here applied to estimate the time varying conditional variance of the U.S. annual rate of inflation from 1820 to 1998 (Fig. 4) calculated as the annual percentage change in the consumer price index (base 1982-84=100). This series has been obtained as a combination of three indices. From 1820 through 1874, the annual cost-of-living index calculated by the Federal Reserve Bank is used. From 1875 until 1912, it uses a monthly Index of General Prices calculated by the the Federal Reserve Bank of New York which was weighted between wholesale commodity prices (20%), wage payments (35%), the cost of living (35%) and rents (10%). For more information on this index, the interested reader may refer to Snyder (1924). From 1913 on, the Bureau of Labour's Consumer Price Index is used (CPI-U).

As a first step, I perform an ARCH-LM test (Engle, 1982) on the residuals of a first order autoregression. The order of the autoregression has been chosen on the basis of the analysis of the global and partial autocorrelation functions of the series. The LM test for ARCH is highly significant at any reasonable level up to lag 8, suggesting the presence of autoregressive conditional heteroskedasticity in the data.

Figure 4. U.S. annual inflation rate from 1820 to 1998 (Source: Global Financial Data).



As in the previous section, the order of the CPV model to be fitted can be chosen to minimize the value of the AIC or some other criterion such as the Schwarz Criterion (SC). The latter penalizes overparameterised models in a more severe way than the AIC. Table 3 reproduces the values of the AIC and SC for different model specifications. The model structure considered includes an AR(1) specification for the conditional mean with CPV(r,s) residuals. The search has been restricted within the intervals $1 \leq r \leq 3$ and $0 \leq s \leq 1$.

Table 3. AIC and SC values for different CPV regression models

Model order (r,s)	**AIC**	**SC**
(1,0)	6.0336314	6.1240539
(2,0)	6.0483059	6.2291508
(3,0)	6.0822615	6.4077823
(1,1)	5.9351725	**6.1160742**
(2,1)	**5.9136737**	6.2391945

Also, several different CPV-C specifications have been considered and the corresponding AIC and SC values have been reported in Table 4. For the unconstrained CPV model, the AIC indicates a model of order (2,1) while a (1,1) model is selected on the basis of the SC.

Table 4. AIC and SC values for different CPV-C regression models

Model order (r,s)	AIC	SC
(1,0)	6.0300828	6.1024208
(2,0)	6.0518270	6.1603339
(3,0)	6.0239173	6.1866777
(1,1)	**5.9398935**	**6.0484004**
(2,1)	5.9479811	6.1107415

Furthermore, the SC value for the CPV(1,0) model is only slightly higher than that registered for the more complex (1,1) model. The CPV(2,1) model is likely to be an overparameterised model for this series as suggested by the high number of not statistically significant parameters in its specification.

Table 5. ML parameter estimates and asymptotic s.e. (in parentheses)

AR(1)-CPV(1,0)	σ_e^2	Q	A	ϕ_0	ϕ_1
	15.7690 (2.2466)	0.3727 (0.1619)	0.3705 (0.1849)	0.7439 (0.3788)	0.4851 (0.0939)
AR(1)-CPV(1,1)	σ_e^2	Q(1,1)	Q(2,2)	Q(1,2)	A(1,1)
	1.4823 (0.5192)	0.3948 (0.1852)	0.2859 (0.1185)	0.0927 (0.0976)	-0.1956 (0.2375)
	A(2,2)	A(1,2)	A(2,1)	ϕ_0	ϕ_1
	0.2875 (0.1264)	0.9963 (0.1857)	0.1851 (0.0874)	1.4775 (0.5109)	0.5210 (0.1266)
AR(1)-CPV-C(1,1)	σ_e^2	Q(1,1)	Q(2,2)	Q(1,2)	ϕ_0
	1.8024 (0.8649)	0.2540 (0.1003)	0.6878 (0.0923)	-0.0153 (0.0645)	1.1628 (0.4605)
	ϕ_1				
	0.5704 (0.0724)				

For the sake of brevity, the estimates relative to the CPV(2,1) model have not been reported here.

Differently, for the constrained model, both criteria designate a CPV-C model of order (1,1). Hence, I consider estimating two different unconstrained CPV specifications of orders (1,0) and (1,1), respectively, and a constrained CPV-C model of order (1,1) (Table 5). The results have been compared with those obtained using the conditionally Gaussian AR(1)-GARCH (1,1) model

$$y_t = \underset{(0.928)}{3.016} + \underset{(0.066)}{0.624}\, y_{t-1} + u_t$$

$$h_t^2 = \underset{(0.384)}{0.749} + \underset{(0.074)}{0.250}\, u_{t-1}^2 + \underset{(0.056)}{0.749}\, h_{t-1}^2$$

and the AR(1)-ARCH(1) model

$$y_t = \underset{(0.735)}{1.572} + \underset{(0.074)}{0.523}\, y_{t-1} + u_t$$

$$h_t^2 = \underset{(1.091)}{15.233} + \underset{(0.130)}{0.476}\, u_{t-1}^2$$

The performance of all the models considered in estimating the conditional variance of the annual U.S. inflation rate has been evaluated on the basis of four different loss functions, namely

$$MSE_v = T^{-1}\sum_{t=1}^{T}[u_t^2 - \hat{h}_t^2]^2 \qquad LSE_v = T^{-1}\sum_{t=1}^{T}[ln(u_t^2) - ln(\hat{h}_t^2)]^2$$

$$MAE_v = T^{-1}\sum_{t=1}^{T}|u_t^2 - \hat{h}_t^2| \qquad LAE_v = T^{-1}\sum_{t=1}^{T}|ln(u_t^2) - ln(\hat{h}_t^2)|.$$

The results have been reported in Table 6. The last column gives the value of the maximised log-likelihood and, in parentheses, the number of estimated parameters for each model. The CPV models considered always perform better than GARCH models in terms of MSE_v and

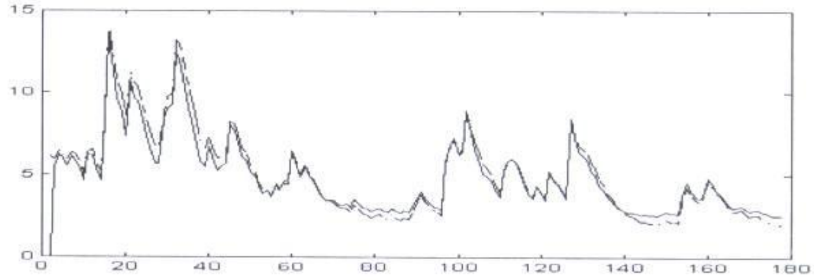
MAE_v even if the GARCH(1,1) model performs better than all the other models in terms of LAE_v.

Table 6. Values of four different loss functions

Model	MSE_v	MAE_v	LSE_v	LAE_v	*Lik.*
AR(1)-CPV(1,0)	3826.3	30.789	10.266	9.137	-522.96 (5)
AR(1)-CPV(1,1)	4057.1	31.023	8.249	2.403	-509.68 (10)
AR(1)-CPV-C(1,1)	3951.8	31.565	9.463	3.196	-514.10 (6)
AR(1)-ARCH(1)	4105.8	31.585	12.175	2.428	-527.36 (4)
AR(1)-GARCH(1,1)	4354.2	33.526	8.677	2.211	-516.37 (5)

Figure 5 shows the estimated conditional standard deviations obtained using a CPV-C (1,1) and a GARCH(1,1) model.

Fig. 5. Conditional s.d. from the CPV-C(1,1) (solid) and the GARCH(1,1) (dashed) models.



4. A structural state space model with conditional heteroskedasticity

4.1 The S-CPV model

Some economic time series exihibit features which cannot be separately captured by one of the two models we have considered in the two previous sections. In these cases, in order to allow for simultaneous modelling of the conditional mean components and the changing conditional variance, a feasible approach is to define a new model obtained nesting a CPV(r,s) structure into the KG model defined in section 2. The final model will be called a *Structural CPV* (S-CPV) model of order (q, p, S, 1, r, s) and defined as

$$y_t = [C_1 \quad C_2 \quad C_3 \quad C_{4,t} \quad C_t^*]x_t^\dagger + e_t \tag{12a}$$

$$x_t^\dagger = \begin{bmatrix} A_1 & 0 & 0 & 0 & 0 \\ 0 & A_2 & 0 & 0 & 0 \\ 0 & 0 & A_3 & 0 & 0 \\ 0 & 0 & 0 & A_4 & 0 \\ 0 & 0 & 0 & 0 & A^* \end{bmatrix} x_{t-1}^\dagger + \begin{bmatrix} D_1 \\ D_2 \\ D_3 \\ D_4 \\ D^* \end{bmatrix} q_t^\dagger \tag{12b}$$

where

$$x_t^\dagger = [\, x_t' \quad (x_t^*)' \,]' \qquad and \qquad q_t^\dagger = [\, q_t' \quad (q_t^*)' \,]',$$

x_t^*, q_t^*, C_t^* and A^* are defined as in section 3. The system error has multivariate Gaussian distribution $q_t^\dagger \sim N(0, Q^\dagger)$ with

$$Q^\dagger = \begin{bmatrix} Q & 0 \\ 0 & Q^* \end{bmatrix}$$

Example. A S-CPV model of order (2, 2, 4, 1, 1, 1) i.e. with a trend of order 2, an AR(2) component, a seasonal component of period $S=4$, a trading day effect and a CPV(1,1) error will have system matrices given by

$$A_1 = \begin{bmatrix} 2 & -1 \\ 1 & 0 \end{bmatrix} \qquad A_2 = \begin{bmatrix} \phi_1 & \phi_2 \\ 1 & 0 \end{bmatrix} \qquad A_3 = \begin{bmatrix} -1 & -1 & -1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

$$A_4 = I_6 \qquad\qquad A^* = \begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{bmatrix}$$

$$C_t^{\dagger} = [1 \quad 0 | \quad 1 \quad 0 | \quad 1 \quad 0 \quad 0 | \quad d_{1,t},\dots,d_{6,t} \,|\, \quad u_{t-1} \quad h_{t-1}\,],$$

where I_6 denotes an identity matrix of order 6; D_1 (2×5), D_2 (2×5) and D_3 (3×5) are matrices of zeros with the first, the second and the third element of the first row, respectively, equal to 1; D_4 (6×5) is a matrix of zeros and D^* (2×5) has all the elements equal to 0 except for the fourth element of the first row and the fifth element of the second row which are both equal to 1. The state vector and the system error are given by

$$x_t^{\dagger} = [\, c_t \quad c_{t-1} \,|\, \quad w_t \quad w_{t-1} \,|\, \quad s_t \quad s_{t-1} \quad s_{t-2} \,|\, \quad \gamma_{1,t},\dots,\gamma_{6,t} \,|\, \quad a_{1,t} \quad b_{1,t}\,]'$$

$$q_t^{\dagger} = [\, q_{1,t} \quad q_{2,t} \quad q_{3,t} \,|\, \quad q_{1,t}^{*} \quad q_{2,t}^{*}\,]'$$

The log-likelihood function of a S-CPV model can be written as

$$\ell(\,y; A^{\dagger}, \sigma_e^2, Q^{\dagger}\,) = -\frac{T}{2}\log(2\pi) - \frac{1}{2}\sum_{t=1}^{T}\log\left[(\,h_t^{\dagger}\,)^2\right] - \frac{1}{2}\sum_{t=1}^{T}\frac{(\,e_{t|t-1}^{\dagger}\,)^2}{(\,h_t^{\dagger}\,)^2} \quad (13)$$

where

$$e_{t|t-1}^{\dagger} = y_t - C_t^{\dagger} x_{t|t-1}^{\dagger}$$

and

$$(\,h_t^{\dagger}\,)^2 = C_t^{\dagger} P_{t|t-1}^{\dagger} \left(C_t^{\dagger}\right)' + \sigma_e^2$$

with $x^{\dagger}_{t|t-1} = E(x^{\dagger}_t \mid y^{t-1})$ and $P^{\dagger}_{t|t-1} = Var(x^{\dagger}_t \mid y^{t-1})$. The unknown parameters in $\{A^{\dagger}, \sigma^2_e, Q^{\dagger}\}$ can be estimated by maximizing the Gaussian log-likelihood (13) via the EM algorithm. Alternatively *scoring* or *quasi-Newton* methods could also be used (see Watson and Engle, 1983, for a discussion on the application of the method of scoring in the context of state space models).

4.2. The monthly inflation data revisited

Let us consider the predicted irregular component of the model estimated in section 2.2, $e_{t|t-1}$. Looking at the global and partial autocorrelations of $e^2_{t|t-1}$, it turns out that the 1st, 5th, 6th and 8th all exceed the confidence limit of two asymptotic standard errors. Also, the ARCH-LM test is significant at the 99% level up to lag 3. These results suggest the presence of a CH component which has not been modelled in the previous analysis. Hence, I include a CPV(*r,s*) component in the model estimated in section 2. Again, the order of the CH component is chosen to minimize the value of the AIC. The results of the identification procedure are compared with those obtained using the SC (Table 7). Both the criteria suggest a CPV component of order (1,0). Table 8 reproduces the ML estimates of the model parameters and their respective standard errors.

Table 7. AIC and SC values for Structural CPV models

CPV		
(r,s)	**AIC**	**SC**
(1,0)	**-9.1877**	**-8.9012**
(2,0)	-9.1054	-8.7501
(1,1)	-9.0964	-8.7411
CPV-C		
(r,s)	**AIC**	**SC**
(1,1)	-9.1315	-8.8564
(2,1)	-9.1090	-8.7995
(1,1)	-9.1017	-8.7922

The predictive performance of the model has been assessed calculating out of sample k-steps ahead forecasts (k=1,...,10) for the period from 1999.01 to 1999.10 (Fig. 6) conditioning on the information available at December 1998. The predicted values follow the data closely for lead times not greater than 4 months. For k>4 they tend to interpolate the observed data failing to capture the short term behaviour of the series. Also, 1-month ahead forecasts have been recursively calculated for the same period (Fig. 6). These shorter term predictions result very close to the observed data performing extremely well in capturing the turning points of the series. The results (Table 9) have been compared with those obtained using the GK approach considering four different indices and, namely, the Mean Square Error (MSE), the Mean Absolute Error (MAE), the Mean Absolute Percentage Error (MAPE) and the Theil inequality coefficient (THEIL).

Table 8. ML parameter estimates and asymptotic standard errors (in parentheses) for the S-CPV (1,3,12,0,1,0) model

ϕ_1	ϕ_2	ϕ_3
0.7328 (0.0465)	0.0585 (0.0761)	-0.4347 (0.0619)
$var(q_{1,t})$	$var(q_{2,t})$	$var(q_{3,t})$
4.6747*10-7 (1.3513*10-7)	6.0238*10-7 (1.6668*10-7)	4.9952*10-8 (2.3364*10-8)
σ_e^2	$var(q^*_{1,t})$	α
1.3387*10-6 (2.3573*10-7)	0.1585 (0.0561)	0.2045 (0.0968)

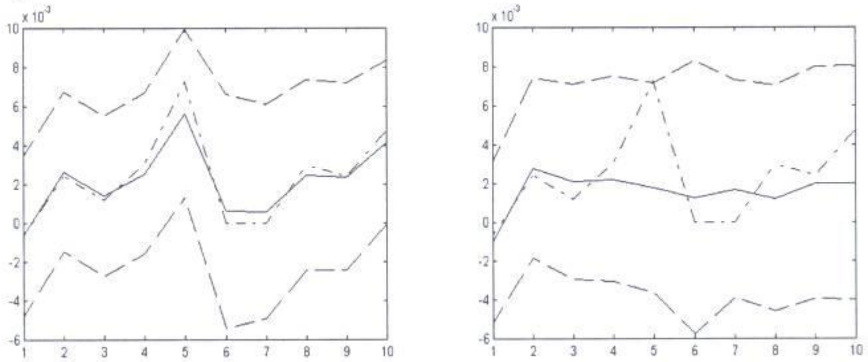
Table 9. Forecast error summary statistics for GK and S-CPV models

	Static(1-step ahead)		*Dynamic (k-steps ahead)*	
	S-CPV	*GK*	*S-CPV*	*GK*
RMSE	0.0006196	0.0012	0.0021	0.0021
MAE	0.0004406	0.0008621	0.0014	0.0015
MAPE	0.0143	0.0261	0.0503	0.0517
THEIL	0.1073	0.2172	0.4201	0.4319

The first two forecast error statistics depend on the scale of the dependent variable. The remaining two statistics are scale invariant. The Theil inequality coefficient always lies between zero and one, where zero indicates a perfect fit.

It can be easily observed how the S-CPV model always performs better than the GK model using only two more parameters than the latter. The gain obtained by the S-CPV model is particularly evident if we focus on 1 step ahead forecasts.

Figure 6. ∇ log(CPI-U) series. 1-step ahead (left) and k-steps ahead (right) predictions from the S-CPV(1,3,12,0,1,0) model for the period from 99.01 to 99.10; (—) predicted, (-.-) observed, (--) ±2 s.e. confidence bands.



5. Concluding remarks

The proposed modelling approach has been found useful for decomposing and forecasting conditionally heteroskedastic structural time series with trend, seasonal and autoregressive components. These features are usually observed in monthly and quarterly inflation data such as the U.S. monthly inflation series which has been analysed in this paper. The results obtained show that the S-CPV model not only gives a satisfactory forecasting performance but also produces a structural decomposition of the series which is consistent with the suggestions of economic theory.

Furthermore, the S-CPV model perfectly fits in a linear Gaussian state space context. The Kalman filter algorithm is used to compute, at each time step, the value of the log-likelihood function which is then maximised using the EM algorithm. This is, at the same time, computationally efficient and simple to apply, since it requires only the implementation of a Fixed Interval Smoother and some straightforward regression calculations.

Acknowledgements: This paper is supported by MURST 1998: "Modelli Statistici per l'Analisi delle Serie Temporali".

References

Bollerslev T. (1986) Generalized Autoregressive Conditional Heteroskedasticity, *Journal of Econometrics*, 31, 307-327.

Engle R. F. (1982) Autoregressive conditional heteroskedasticity with estimates of the variance of U.K. inflation, *Econometrica*, 50, 987-1008.

Kalman R. E. (1960) A new approach to linear filtering and prediction problems, *Trans. ASME, Journal of Basic Engineering*, D, 80, 35-45.

Kitagawa G., Gersch W. (1984) A smoothness priors-state space modelling of time series with trend and seasonality, *Journal of the American Statistical Association*, 79, 378-389.

Snyder C. (1924) A New Index of the General Price Level from 1875, *Journal of the American Statistical Association*, 19, 110-117.

Storti G. (1999) The CPV model: a state space generalization of GARCH processes, *Proceedings of the conference SCO 99, "Modelli Complessi e Metodi Computazionali Intensivi per la Stima e la Previsione"*, 248-253.

Watson M. W., Engle R. F. (1983) Alternative algorithms for the estimation of dynamic factor, MIMIC and varying coefficient regression models, *Journal of Econometrics*, 23, 385-400.
Wu L. S., Pai J. S., Hosking J. R. M. (1996) An algorithm for estimating parameters of state-space models, *Statistics & Probability Letters*, 28, 99-106.