Comput Stat (2009) 24:459–479
DOI 10.1007/s00180-008-0139-3

ORIGINAL PAPER

An omnibus noise filter

Claudio Morana

Received: 7 May 2007 / Accepted: 2 October 2008 / Published online: 28 October 2008

Abstract A new noise filtering approach, based on flexible least squares (FLS) estimation of an unobserved component local level model, is introduced. The proposed FLS filter has been found to perform well in Monte Carlo analysis, independently of the persistence properties of the data and the size of the signal to noise ratio, ouperforming in general even the Wiener Kolmogorov filter, which, theoretically, is a minimum mean square estimator. Moreover, a key advantage of the proposed filter, relatively to available competitors, is that any persistence property of the data can be handled, without any pretesting, being computationally fast and not demanding, and easy to be implemented as well.

Keywords Signal–noise decomposition · Long memory · Structural breaks · Flexible least squares · Exchange rates volatility

JEL Classification C32

This paper was partially written when the author was visiting the Department of Economics at Michigan State University. The author gratefully acknowledges funding from the Fulbright Commission and Michigan State University for hospitality. Many thanks to two anonymous referee and the associate editor for very constructive comments.

C. Morana (✉)
Facoltà di Economia, Dipartimento di Scienze Economiche e Metodi Quantitativi,
Università del Piemonte Orientale, Via Perrone 18,
28100 Novara, Italy
e-mail: claudio.morana@eco.unipmn.it

C. Morana
International Centre for Economic Research
(ICER, Torino), Torino, Italy

1 Introduction

Several recent contributions to the literature have been concerned with noise filtering for single or multivariate long memory processes (see Baillie 1996, for an introduction to long memory processes). Among the seminal contributions, Harvey (1998) has proposed the use of the Wiener–Kolmogorov (WK) filtering approach, generalized to the multivariate common long memory components case by Morana (2002). Beltratti and Morana (2006) have also proposed a semiparametric version of the WK filter, showing a similar performance to the parametric version of Harvey (1998) in Monte Carlo simulations, but with the advantage of not requiring the specification of the exact parametric structure of the process under investigation. Moreover, Arino and Marmol (2004) have generalized the Beveridge-Nelson (1981) permanent-transitory decomposition to the case of univariate nonstationary long memory processes, while in Morana (2004) and in Morana (2006) two approaches for the estimation of the common long memory components in fractionally cointegrated processes, based on the Kasa (1992) and the Gonzalo and Granger (1995) decompositions, respectively, have been proposed (see Engle and Granger 1987; Robinson and Yajima 2002; Marinucci and Robinson 2001, for details about the concept of fractional cointegration). Finally, in Morana (2007) a permanent-persistent–nonpersistent decomposition for univariate or multivariate long memory processes, based on Fourier transform filtering and principal components analysis, has been introduced. Relatively to previous contributions, the proposed approach has the advantage of being computationally fast and easily implemented also when the temporal or cross-sectional dimension is very large, requiring however the estimation of the fractional differencing parameter of the series investigated.

In this paper a new noise filtering method is proposed. The method is found to perform well also when the inverse signal to noise ratio is very large, independently of the persistence properties of the series to be filtered, i.e. short memory ($I(d)$, $d = 0$), stationary long memory ($I(d)$, $0 < d < 0.5$), nonstationary long memory ($I(d)$, $0.5 < d < 1$), stochastic nonstationary ($d = 1$), with or without deterministic nonstationarity, i.e. structural breaks. Relatively to the other available approaches, the proposed method does not require pretesting, i.e. unit root or structural breaks testing, or the estimation of the fractional differencing parameter, since filtering is based on flexible least squares estimation (FLS, Kalaba and Tesfatsion 1989) of an unobserved component local level model (Harvey 1989). In general the FLS filter is found to perform even better than the WK filter, which, theoretically, is a minimum mean square error estimator.

After this introduction, the paper is organized as follows. In sections two the econometric methodology is introduced, while is section three the Monte Carlo evidence is discussed; finally in section four an empirical application is provided, while in section five conclusions are drawn.

2 An omnibus noise filter

Consider the multivariate real valued stochastic process $\{\mathbf{y}_t\}_{t=0}^{T-1}$, of dimension k, integrated of order $I(d)$ $(0 \leq d \leq 1)$, subject or not to structural change of unknown form, observed with noise

$$\mathbf{y}_t = \mathbf{b}_t + \mathbf{a}_t + \mathbf{u}_t, \quad (1)$$
$$\Delta(d)\mathbf{a}_t = \mathbf{v}_t, \quad (2)$$
$$\mathbf{b}_t = \mathbf{f}(t), \quad (3)$$
$$\mathbf{u}_t, \mathbf{v}_t \sim i.i.d(\mathbf{0}, \Sigma_i) \;\; i = u, v, \quad (4)$$
$$E[\mathbf{u}_t \mathbf{v}_s'] = \mathbf{0}\; \forall t, s, \quad (5)$$

where $\mathbf{a}_t$ is the $k \times 1$ vector of unobserved persistent signal components, $\mathbf{b}_t$ is the $k \times 1$ vector of unobserved break processes with $\mathbf{f}(t)$ a $k \times 1$ vector of generic nonlinear deterministic functions of time, $\Delta(d)$ is a $k \times k$ diagonal matrix of fractional filters, with ith element $(1 - L)^{d_i}$, $\mathbf{u}_t$ a white noise component uncorrelated with the signal innovation component $\mathbf{v}_t$ at all leads and lags.

Noise filtering can be performed by means of FLS estimation of the model

$$\mathbf{y}_t = \mathbf{H}_t \beta_t + \mathbf{u}_t, \quad (6)$$
$$\beta_t = \mathbf{F}_t \beta_{t-1} + \mathbf{v}_t, \quad (7)$$
$$\mathbf{u}_t \simeq \mathbf{0}, \mathbf{v}_t \simeq \mathbf{0},$$

where $\mathbf{H}_t$ is a $k \times k$ diagonal matrix, $\mathbf{F}_t$ is a $k \times k$ diagonal transition matrix, set equal to the identity matrix in the current application ($\mathbf{F}_t = \mathbf{I}_k$), β_t is a $k \times 1$ vector of parameters.[1] The above specification can be regarded as approximately linear ($\mathbf{u}_t \simeq \mathbf{0}$), where the evolution of the unobserved signal component is assumed to take place only gradually ($\mathbf{v}_t \simeq \mathbf{0}$).

Alternatively, in a stochastic formulation, by setting $\mathbf{H}_t = \mathbf{F}_t = \mathbf{I}_k$, $\mathbf{u}_t \sim iidN(\mathbf{0}, \Sigma_u)$, $\mathbf{v}_t \sim iidN(\mathbf{0}, \Sigma_v)$, $E[\mathbf{u}_t \mathbf{v}_t'] = \mathbf{0}$, it can be interpreted in terms of a multivariate local level unobserved component model (Harvey 1989).

The FLS problem can then be stated as

$$\min_{\gamma} c_M^2(\beta) + \mu c_D^2(\beta), \quad (8)$$

where β is the vector referring to the entire story $\{\beta_t\}_{t=1}^T$, and the two components in the objective function can be thought of as the cost associated with measurement errors (measurement cost)

$$c_M^2(\beta) = \sum_{t=2}^{N} (\mathbf{y}_t - \mathbf{H}_t \beta_t)' \mathbf{M}_t (\mathbf{y}_t - \mathbf{H}_t \beta_t),$$

[1] As correctly observed by one of the referees, setting the $\mathbf{F}$ matrix equal to the identity matrix takes some "flexibility" away from the FLS approach. Yet, also in the light of the Monte Carlo results of Kladroba (2005), an identity transition matrix is a working solution to the problem of sensitivity to misspecification.

and state dynamics (dynamic cost)

$$c_D^2(\beta) = \sum_{t=2}^{N} (\beta_t - \beta_{t-1})' \mathbf{D}_t (\beta_t - \beta_{t-1}),$$

respectively, where μ is the known penalization term for parameters dynamics and $\mathbf{D}_t$ and $\mathbf{M}_t$ are symmetric and positive definite weighting matrices. The FLS solution to the above problem is the set of sequences $\{\beta_t\}_{t=1}^{T}$ minimizing both the measurement and dynamic costs, each conditional to a given choice for the penalization parameter μ.

A sequential updating procedure is in practice employed to solve the above problem, implemented by means of the following equations (Kalaba and Tesfatsion 1990b)

$$\beta_t = \mathbf{U}_{t-1} \mathbf{z}_t, \tag{9}$$

$$\mathbf{U}_t = \mathbf{H}_t' \mathbf{M}_t \mathbf{H}_t + \mathbf{Q}_{t-1}, \tag{10}$$

$$\mathbf{V}_t = \left[\mu \mathbf{F}_t' \mathbf{D}_t \mathbf{F}_t + \mathbf{U}_t\right]^{-1}, \tag{11}$$

$$\mathbf{z}_t = \mathbf{H}_t' \mathbf{M}_t \mathbf{y}_t + \mathbf{p}_{t-1}, \tag{12}$$

$$\mathbf{G}_t = \mu \mathbf{V}_t \mathbf{F}_t' \mathbf{D}_t, \tag{13}$$

$$\mathbf{Q}_t = \mu \mathbf{D}_t (\mathbf{I} - \mathbf{F}_t \mathbf{G}_t), \tag{14}$$

$$\mathbf{p}_t = \mathbf{G}_t' \mathbf{z}_t, \tag{15}$$

$$\mathbf{s}_t = \mathbf{V}_t \mathbf{z}_t, \tag{16}$$

initialized by setting $\mathbf{Q}_0 = \mathbf{0}$ and $\mathbf{p}_0 = \mathbf{0}$. After filtering, smoothing can be implemented through the formula

$$\beta_{t|T} = \mathbf{s}_t + \mathbf{G}_t \beta_{t+1|T},$$

with $\beta_{T|T} = \beta_T$.

A key issue for the implementation of the noise filtering approach is the selection of the "optimal" value for the penalization parameter μ. The proposed solution to this problem is to run the filter assuming several values for the penalization parameter $\mu = \{0.001, \ldots, 1000\}$, computing the signal-noise decomposition in each case, and evaluating the noise persistence properties by means of serial correlation tests. The optimal decomposition is obtained as the one characterized by the smallest penalization parameter which yields a nonserially correlated noise component, assessed by means of standard serial correlation tests, as for instance the Box–Pierce test or the Bartlett test. This selection procedure grants to the estimated signal process maximal variability, conditional to extracting all the relevant systematic dynamics from the series to be filtered and to following only a gradual dynamic path. Monte Carlo results provide full support to the proposed approach, independently of the persistence properties of the data and of the size of the inverse signal to noise ratio. In the rest of the section, the proposed filter and its relation to alternative methods proposed in the literature is discussed.

2.1 Flexible least squares and Kalman filtering

It should be noted that the Kalman Filter and the FLS filter are related, since the Kalman filter can be derived from the FLS filter in the case of linear and Gaussian dynamic and measurement equations, with mutually independent white noise shock terms, by setting the weight matrix $\mu\mathbf{D}_t$ equal to $\mathbf{\Sigma}_v^{-1}$ and the weight matrix $\mathbf{M}_t$ equal to $\mathbf{\Sigma}_u^{-1}$. Then the matrix $\mathbf{U}_t$ corresponds to $\mathbf{\Sigma}_{t,t}^{-1}$, i.e. the inverse of the error covariance matrix for the state $\beta_{t|t}$, and the vector $\mathbf{z}_t$ corresponds to the modified state estimate $\mathbf{\Sigma}_{t,t}^{-1}\beta_{t|t}$, yielding together the information filters version of the Kalman filter equations. The penalization parameter can then be interpreted as the inverse signal to noise ratio (see Kalaba and Tesfatsion 1990a for additional details).

Applying the Kalman filter to a linear state space form yields the minimum mean square error estimator of the state vector under Gaussianity. If Gaussianity does not hold the Kalman filter is a minimum mean square error linear estimator (see Harvey 1998 for additional details).

There are however other important differences between the two filtering approaches, which can lead to different results in estimation. In fact, the Kalman filter is a point estimation technique, yielding the most likely state sequence estimate for a correctly specified stochastic model, usually implemented by means of the prediction error decomposition of the Gaussian log-likelihood function. Differently, FLS is concerned with a multicriteria model specification and estimation problem, fully independent of probability assumptions, yielding the set of all state sequence estimates characterized by vector-minimal incompatibility between imperfectly specified theoretical relations and empirical observations. Hence, there is no guarantee, neither theoretical, nor from simulation, that the same state estimate would be provided by the two approaches. The simulations carried out in the paper actually show that in general this is not the case. Similarly, Kladroba (2005) shows by simulation how the relative performance of FLS relative to Kalman filtering crucially depends on the selection of the penalization parameter, and therefore on the selection of the inverse signal to noise ratio. As already pointed out, while in Kalman filtering the latter parameter is selected by means of a likelihood function based approach, in the proposed FLS implementation the penalization parameter is determined on the basis of a residuals whiteness criterion.

2.2 Wiener–Kolmogorov filtering

The WK filter of Harvey (1998) is based on the spectral generating function

$$g_j = \sigma_\eta^2[2(1-\cos\lambda_j)]^{-d} + \sigma_\varepsilon^2 \tag{17}$$

for the stationary case and

$$g_j = \sigma_\eta^2[2(1-\cos\lambda_j)]^{1-d} + 2(1-\cos\lambda_j)\sigma_\varepsilon^2 \tag{18}$$

for the nonstationary case, where d is the fractional differencing operator, $\sigma_\varepsilon^2/\sigma_\eta^2$ is the inverse signal to noise ratio, and $\lambda_j = 2\pi j/T$. Estimation can then be carried out by maximizing the frequency-domain (quasi) log-likelihood function

$$\log L = -\frac{1}{2}\sum_{j=1}^{T-1}\log g_j - \pi\sum_{j=1}^{T-1}\frac{I(\lambda_i)}{g_j}, \tag{19}$$

where $I\left(\lambda_j\right)$ is the sample periodogram.

Two sided time domain weights to filter the long memory signal from the observed process are then computed from the inverse Fourier transform of the transfer function $h\left(\lambda_j\right) = g(\lambda_i)/I(\lambda_i)$ or by the formula

$$w(L) = \frac{1}{1+(\sigma_\varepsilon^2/\sigma_\eta^2)|1-L|^{2d}}. \tag{20}$$

This yields the minimum mean square estimator of the signal component under Gaussianity and the minimum mean square linear estimator if Gaussianity does not hold (see Harvey 1998 for further details).

3 Monte Carlo results

In this section, FLS is compared with KF and WK by means of Monte Carlo simulations. In the exercise, the following data generating process has been assumed for the process y_t

$$\begin{aligned} y_t &= \beta_t + b_t + \varepsilon_t \\ (1-L)^d\beta_t &= v_t \\ & v_t \sim n.i.d.(0,1) \\ & \varepsilon_t \sim n.i.d.(0,\sigma^2) \\ b_t &= (b_{1,t},\ b_{2,t}) \\ b_{1,t} &= 0\ \ \forall t \\ b_{2,t} &= \begin{cases} 1 & 1 \le t \le T/4 \\ 3 & (T/4)+1 \le t \le T/2 \\ -2 & (T/2)+1 \le t \le (3T/4) \\ 1 & (3T/4)+1 \le t \le T \end{cases} \end{aligned}$$

for the long memory case and

$$\begin{aligned} y_t &= \beta_t + b_t + \varepsilon_t \\ \beta_t &= \pi\beta_{t-1} + v_t \\ & v_t \sim n.i.d.(0,1) \\ & \varepsilon_t \sim n.i.d.(0,\sigma^2) \end{aligned}$$

$$
\begin{aligned}
b_t &= (b_{1,t},\ b_{2,t}) \\
b_{1,t} &= 0\ \ \forall t \\
b_{2,t} &= \begin{cases} 1 & 1 \le t \le T/4 \\ 3 & (T/4)+1 \le t \le T/2 \\ -2 & (T/2)+1 \le t \le (3T/4) \\ 1 & (3T/4)+1 \le t \le T \end{cases}
\end{aligned}
$$

for the short memory case, with $\sigma^2 = (s/n)^{-1} = \{10, 7, 5, 3, 2, 1, 0.5, 0.25, 0.125\}$, $d = \{0.2, 0.4, 0.6, 0.8, 1\}$, $\pi = \{0.2, 0.4, 0.6, 0.8\}$, $t = 1, \ldots, 500$, and penalization parameter for FLS estimation $\mu = \left(0.001,\ \ldots,\ 1000\right)$. The number of replications has been set to 500.

The performance of the signal extraction technique is assessed with reference to the ability of recovering the unobserved signal components $\beta_t^* = \beta_t + b_t$, without distinguishing between the long memory and deterministic break components when $b_{2,t}$ is employed. The root mean square forecast error (RMSFE), the Theil inequality coefficient (IC), and the correlation coefficient are employed in the evaluation:

$$
\begin{aligned}
\text{RMSFE} &= \sqrt{\frac{1}{T}\sum_{t=1}^{T}\left(\beta_t^* - \hat{\beta}_t\right)^2}, \\
\text{IC} &= \frac{\text{RMSFE}}{\sqrt{\frac{1}{T}\sum_{t=1}^{T}\beta_t^{*2} + \sqrt{\frac{1}{T}\sum_{t=1}^{T}\hat{\beta}_t^2}}}, \\
\rho &= \frac{\text{Cov}(\beta_t^*, \hat{\beta}_t)}{\sqrt{\text{Var}(\beta_t^*)\text{Var}(\hat{\beta}_t)}},
\end{aligned}
$$

where β_t^* is the true signal and $\hat{\beta}_t$ is the estimated signal. The correlation coefficient, the RMSFE statistics, and the Theil inequality coefficient are computed for each replication and then averaged. The performance of the FLS filter is assessed under two scenarios.

The first one is denoted as the "unfeasible" case and concerns the best attainable performance of the filter. In this first case the optimal signal estimate is obtained by selecting the best process, among all the ones computed by changing the value of the penalization parameter, on the basis of RMSFE minimization or correlation coefficient maximization. In addition to the absolute performance, the FLS filter has also been evaluated by comparing its performance with the one provided by the WK filter (Harvey 1998), for the unfeasible case as well. The latter corresponds to the case in which all the relevant parameters are assumed to be known, i.e. the fractional differencing parameter, the signal to noise ratio and the break process. Then, the performance of the FLS approach, relative to the WK filter, has been assessed by computing ratios for the average RMSFE and ρ statistics. The comparison with the WK filter is of particular interest, since it yields the minimum mean square error estimator of the signal component under Gaussianity, and the minimum mean square error estimator within the class of linear estimators when Gaussianity does not hold.

The second scenario, on the other hand, is denoted as the "feasible" case. In this latter case the optimal estimate of the signal component is selected according to serial correlation tests performed on the residuals, i.e. the Box–Pierce test, carried out using a serial correlation order $p = T^{0.5}$, and the Bartlett white noise test (see Priestley 1981, pp 479–483), for various significance levels, i.e. 1, 5, 10, and 20%. In addition to contrast the feasible FLS filter with its unfeasible version, the comparison concerns the Kalman filter in its feasible version.

In both the feasible and unfeasible cases the scalar model has been considered for implementation, assuming unitary weighting, i.e. $M_t = D_t = 1$. The results of the Monte Carlo exercise are reported in Tables 1, 2, 3, 4 and 5.

3.1 Long memory

3.1.1 The unfeasible case

As is shown in Table 1, when only long memory is considered, the FLS filter provides in general a superior performance to the WK filter for the unfeasible case. In fact, only for the case $d = 0.20$, the FLS filter shows a larger RMSFE than the WK filter, albeit the RMSFE gap is never larger than 10%. Yet, also for this latter case, the two approaches show the same performance in terms of correlation coefficients. The two filtering methods provide a very close performance for the $d = 0.40$ case as well, with the FLS approach yielding an improvement in terms of correlation coefficient of up to 10% relative to the WK filter.

On the other hand, the superior performance of the FLS filter can be clearly noted for the nonstationary long memory case, with noticeable reductions in the RMSFE (from 20 to 40%) and improvements in the correlation coefficient (up to 20%). While the relative performance of the filters tends to be independent of the value of the signal to noise ratio, their absolute performance is negatively affected as the inverse signal to noise ratio increases. Yet, the performance of both filters is very satisfactory in all the cases, apart from the $d = 20$ case with $(s/n)^{-1} > 5$. In fact, the Theil inequality coefficient tends to be low and below 0.5 in almost all the cases: out of 45 possible cases, the IC coefficient is below 0.3 in 50% of the cases and below 0.5 in 90% of the cases. Finally, it can be noted that the absolute performance of the FLS filter tends to improve as the degree of persistence of the process increases.

Similar results hold for the long memory plus structural change case, albeit, by construction, the results are biased in favor of the WK filter, since in this latter case the break process is taken as known, while, when the FLS filter is employed, the break process is estimated jointly with the long memory component.[2] Yet, the performance of the two filters is never very different. For instance, concerning the RMSFE statistic,

[2] It should be noted that the FLS filter, differently from the WK filter, can handle any kind of nonstationarities, both stochastic (of any order) and deterministic. Deterministic nonstationarities are, on the other hand, not allowed by the WK filter.

Table 1 Monte Carlo results, absolute and relative performance of the FLS filter, unfeasible case

Long memory

FLS filter, relative (WK filter) performance

RMSFE						ρ					
$(s\backslash n)^{-1}\backslash d$	0.2	0.4	0.6	0.8	1	$(s\backslash n)^{-1}\backslash d$	0.2	0.4	0.6	0.8	1
10	1.077	0.998	0.773	0.802	0.597	10	1.004	1.083	1.226	1.130	1.209
7	1.086	1.005	0.777	0.797	0.593	7	1.002	1.042	1.177	1.104	1.173
5	1.093	1.009	0.777	0.800	0.590	5	0.998	1.020	1.138	1.088	1.151
3	1.104	1.018	0.785	0.792	0.582	3	0.997	1.005	1.098	1.073	1.121
2	1.099	1.018	0.792	0.789	0.575	2	0.997	1.002	1.076	1.062	1.108
1	1.072	1.010	0.796	0.782	0.567	1	0.998	1.002	1.053	1.049	1.082
0.5	1.045	0.996	0.803	0.785	0.565	0.5	0.999	1.003	1.037	1.036	1.061
0.25	1.026	0.991	0.827	0.809	0.574	0.25	1.000	1.002	1.022	1.022	1.042
0.125	1.012	0.995	0.859	0.841	0.597	0.125	1.000	1.001	1.011	1.012	1.026

FLS filter, absolute performance

IC						ρ					
$(s\backslash n)^{-1}\backslash d$	0.2	0.4	0.6	0.8	1	$(s\backslash n)^{-1}\backslash d$	0.2	0.4	0.6	0.8	1
10	0.605	0.506	0.318	0.361	0.275	10	0.318	0.515	0.828	0.762	0.870
7	0.577	0.481	0.305	0.330	0.251	7	0.375	0.556	0.838	0.797	0.888
5	0.552	0.458	0.288	0.310	0.230	5	0.423	0.592	0.852	0.820	0.905
3	0.500	0.421	0.268	0.278	0.200	3	0.518	0.653	0.867	0.854	0.927
2	0.458	0.388	0.252	0.255	0.181	2	0.592	0.703	0.880	0.876	0.939
1	0.379	0.328	0.224	0.218	0.151	1	0.718	0.787	0.903	0.908	0.957
0.5	0.302	0.269	0.195	0.187	0.127	0.5	0.820	0.856	0.926	0.932	0.969
0.25	0.229	0.212	0.164	0.158	0.106	0.25	0.896	0.911	0.947	0.951	0.978
0.125	0.168	0.161	0.133	0.129	0.088	0.125	0.943	0.949	0.965	0.967	0.985

Long memory and structural change

FLS filter, relative (WK filter) performance

RMSFE						ρ					
$(s\backslash n)^{-1}\backslash d$	0.2	0.4	0.6	0.8	1	$(s\backslash n)^{-1}\backslash d$	0.2	0.4	0.6	0.8	1
10	1.182	1.125	1.049	0.915	0.831	10	0.953	0.979	0.986	1.010	1.024
7	1.126	1.060	0.981	0.830	0.765	7	0.951	0.987	1.001	1.023	1.046
5	1.097	1.014	0.933	0.756	0.690	5	0.934	0.998	1.023	1.044	1.080
3	1.093	1.011	0.925	0.737	0.675	3	0.950	0.999	1.022	1.038	1.070
2	1.085	0.999	0.909	0.98	0.638	2	0.959	1.003	1.030	1.048	1.082
1	1.061	0.993	0.893	0.661	0.595	1	0.986	1.004	1.030	1.052	1.079
0.5	1.037	0.989	0.893	0.652	0.586	0.5	0.996	1.003	1.020	1.040	1.060
0.25	1.025	0.988	0.902	0.670	0.589	0.25	0.999	1.001	1.013	1.027	1.042
0.125	1.010	0.994	0.925	0.710	0.617	0.125	1.000	1.001	1.006	1.015	1.026

Table 1 continued

FLS filter, absolute performance											
IC						ρ					
$(s\backslash n)^{-1}\ \backslash d$	0.2	0.4	0.6	0.8	1	$(s\backslash n)^{-1}\ \backslash d$	0.2	0.4	0.6	0.8	1
10	0.265	0.216	0.225	0.153	0.191	10	0.851	0.901	0.891	0.953	0.921
7	0.308	0.243	0.253	0.158	0.201	7	0.802	0.877	0.862	0.950	0.915
5	0.376	0.287	0.295	0.171	0.217	5	0.708	0.834	0.819	0.943	0.904
3	0.345	0.262	0.269	0.151	0.187	3	0.749	0.858	0.848	0.955	0.928
2	0.366	0.285	0.283	0.154	0.184	2	0.720	0.836	0.836	0.954	0.933
1	0.344	0.282	0.267	0.149	0.159	1	0.760	0.841	0.859	0.957	0.951
0.5	0.273	0.234	0.225	0.126	0.132	0.5	0849	0.889	0.899	0.969	0.966
0.25	0.208	0.186	0.184	0.107	0.109	0.25	0.912	0.930	0.931	0.977	0.976
0.125	0.154	0.143	0.147	0.090	0.091	0.125	0.952	0.959	0.956	0.984	0.983

The Table reports the ratio of the root mean square forecast error (RMSFE) and the correlation coefficient (ρ) for the FLS filter, relative to the WK filter, in the unfeasible case. In the table also the inequality coefficient (IC) and the correlation coefficient for the unfeasible FLS filter have been reported. Both the pure long memory and the long memory plus structural change case are considered. The inverse signal to noise ratio ranges from 0.125 to 10, while the fractional differencing parameter from 0.20 to 1

for the $d = 0.20$ case the performance gap tends to be close to 10% in favor of the WK filter, while the two filters perform closely for the $d = 0.40$ case, with the FLS filter performing best for $(s/n)^{-1} < 2$. Starting with the nonstationary long memory case ($d > 0.5$), the superiority of the FLS filter becomes clear-cut, with the performance gap obtaining values even close to 40% in favor of the FLS filter. On the other hand, concerning the ρ statistic, the performance gap tends to be smaller, in the range 5–10% in both directions, with the WK filter performing best when persistence is low and the FLS filter performing best when persistence is noticeable. Coherently with the findings for the pure long memory case, the absolute performance of the FLS filter in the long memory plus structural change case is also extremely satisfactory, even outperforming, in terms of IC and ρ statistics, its own performance in the pure long memory case. The finding is coherent with the fact that the performance of the FLS filter increases with the degree of persistence, and unaccounted breaks, for any degree of fractional differencing, do add persistence to the overall process.[3] The improvement is always noticeable when observational noise does characterize the series ($(s/n)^{-1} > 1$), with the IC statistics showing reductions in the range 10–50% and the ρ coefficient increasing in the range 10–150%.

3.1.2 The feasible case

As shown in Table 2, the proposed rule for the selection of the penalization parameter, i.e. the lowest value which yields not serially correlated residuals, yields a

[3] It is well known that neglected structural change may lead to the detection of a unit root in the autoregressive representation of a process. The local level model estimated by FLS, by assuming unit root persistence, albeit misspecified, may therefore be a suitable approximating model.

Table 2 Monte Carlo results, performance of the feasible FLS filter relative to the unfeasible FLS filter

Box–Pierce test (1%)

Long memory

RMSFE						ρ					
$(s\backslash n)^{-1}\backslash d$	0.2	0.4	0.6	0.8	1	$(s\backslash n)^{-1}\backslash d$	0.2	0.4	0.6	0.8	1
10	1.086	1.102	1.290	1.158	1.303	10	0.649	0.959	0.905	0.954	0.930
7	1.041	1.054	1.198	1.108	1.244	7	0.635	0.978	0.937	0.972	0.953
5	1.017	1.025	1.144	1.066	1.168	5	0.645	0.982	0.960	0.983	0.971
3	1.014	1.013	1.067	1.029	1.111	3	0.637	0.966	0.984	0.993	0.985
2	1.050	1.032	1.031	1.020	1.069	2	0.641	0.943	0.993	0.995	0.992
1	1.196	1.117	1.018	1.019	1.032	1	0.635	0.900	0.995	0.996	0.997
0.5	1.443	1.281	1.057	1.045	1.019	0.5	0.637	0.867	0.989	0.993	0.999
0.25	1.833	1.504	1.137	1.101	1.026	0.25	0.643	0.860	0.982	0.989	0.999
0.125	2.435	1.870	1.271	1.215	1.048	0.125	0.650	0.853	0.977	0.984	0.999

Long memory and structural change

RMSFE						ρ					
$(s\backslash n)^{-1}\backslash d$	0.2	0.4	0.6	0.8	1	$(s\backslash n)^{-1}\backslash d$	0.2	0.4	0.6	0.8	1
10	1.028	1.045	1.065	1.108	1.093	10	0.993	0.992	0.987	0.990	0.987
7	1.018	1.030	1.056	1.100	1.083	7	0.993	0.993	0.988	0.990	0.987
5	1.011	1.025	1.042	1.097	1.085	5	0.992	0.993	0.990	0.989	0.986
3	1.016	1.012	1.017	1.051	1.051	3	0.980	0.996	0.995	0.995	0.993
2	1.045	1.017	1.014	1.034	1.036	2	0.925	0.991	0.993	0.997	0.995
1	1.185	1.072	1.041	1.019	1.023	1	0.790	0.964	0.983	0.998	0.998
0.5	1.443	1.196	1.100	1.023	1.021	0.5	0.740	0.938	0.974	0.999	0.999
0.25	1.849	1.430	1.217	1.043	1.028	0.25	0.705	0.914	0.964	0.998	0.999
0.125	2.487	1.823	1.420	1.082	1.038	0.125	0.683	0.892	0.953	0.997	0.999

Bartlett test (1%)

Long memory

RMSFE						ρ					
$(s\backslash n)^{-1}\backslash d$	0.2	0.4	0.6	0.8	1	$(s\backslash n)^{-1}\backslash d$	0.2	0.4	0.6	0.8	1
10	1.100	1.068	1.046	1.407	1.239	10	0.805	0.979	0.974	0.904	0.946
7	1.051	1.039	1.021	1.256	1.209	7	0.809	0.982	0.977	0.943	0.959
5	1.030	1.014	1.009	1.188	1.153	5	0.784	0.977	0.976	0.962	0.973
3	1.008	1.005	1.008	1.098	1.083	3	0.759	0.962	0.974	0.983	0.987
2	1.034	1.002	1.022	1.053	1.053	2	0.737	0.937	0.968	0.992	0.993
1	1.165	1.105	1.065	1.013	1.015	1	0.696	0.889	0.960	0.998	0.999
0.5	1.419	1.286	1.155	1.010	1.012	0.5	0.662	0.848	0.946	0.999	0.999
0.25	1.820	1.579	1.326	1.042	1.025	0.25	0.643	0.818	0.931	0.997	0.999
0.125	2.433	2.038	1.608	1.086	1.044	0.125	0.633	0.801	0.919	0.995	0.999

Table 2 continued

Long memory and structural change

RMSFE						ρ					
$(s\backslash n)^{-1}\ \backslash d$	0.2	0.4	0.6	0.8	1	$(s\backslash n)^{-1}\ \backslash d$	0.2	0.4	0.6	0.8	1
10	1.029	1.043	1.060	1.113	1.104	10	0.991	0.991	0.988	0.983	0.987
7	1.014	1.033	1.046	1.124	1.118	7	0.994	0.992	0.989	0.979	0.985
5	1.008	1.024	1.039	1.190	1.159	5	0.991	0.993	0.989	0.968	0.979
3	1.015	1.005	1.011	1.105	1.086	3	0.969	0.997	0.996	0.986	0.991
2	1.049	1.011	1.006	1.095	1.090	2	0.884	0.990	0.995	0.989	0.992
1	1.188	1.071	1.029	1.055	1.058	1	0.730	0.958	0.985	0.995	0.997
0.5	1.440	1.202	1.082	1.013	1.019	0.5	0.702	0.929	0.975	0.999	0.999
0.25	1.836	1.411	1.194	1.010	1.011	0.25	0.686	0.911	0.962	0.999	1.000
0.125	2.435	1.763	1.395	1.035	1.025	0.125	0.682	0.894	0.949	0.999	1.000

The Table reports the ratio of the root mean square forecast error (RMSFE) and the correlation coefficient (ρ) for the feasible FLS filter, relative to the unfeasible FLS filter. Both the pure long memory and the long memory plus structural change case are considered. The inverse signal to noise ratio ranges from 0.125 to 10, while the fractional differencing parameter from 0.20 to 1

very satisfactory outcome, relative to the performance of the unfeasible version of the filter.[4]

In particular, when the Box–Pierce test is employed, with lag order set to $T^{0.5}$ and significance level set to 10%, the filter performs very satisfactorily. Moreover, when $(s/n)^{-1}$ is large, i.e. when the process is indeed affected by observational noise, the RMSFE for the feasible case gets very close to the minimal RMSFE value attained in its unfeasible version. The worst performance is shown for the $d = 0.20$ case with $(s/n)^{-1} < 0.5$, and for the $d = 0.40$ case with $(s/n)^{-1} < 0.25$, i.e. for the cases in which the noise is almost negligible and the signal is covariance stationary. In this latter case it seems that the rule based on the 10% significance level leads to a too large penalization parameter, and, therefore, to discarding too much of the process in signal reconstruction. The selection of the 1% significance level seems to lead to some improvement for the two latter cases only.

Moreover, the use of the Bartlett test, still selecting a 10% significance level, leads to some improvement over the Box–Pierce test for the $d \leq 0.40$ case, for any degree of noise. In addition, for the ρ coefficient the increase in performance is very strong for the $d = 0.20$ case, i.e. up to 60%. Hence, the Bartlett test would seem to be preferred to the Box–Pierce test for the weak long memory case. In the other cases the use of the Box–Pierce test can lead to similarly satisfactory results ($d = 0.40$) or even superior results ($d > 0.40$).

On the other hand, for the long memory plus structural change case the two serial correlation tests perform similarly well both in terms of RMSFE and ρ statistics, also improving on their performance in the pure long memory case. In particular, the Box–Pierce based selection procedure shows a very strong increment in the ρ coeffi-

[4] For reasons of space in the table only the results for the 1% significance level case are reported. A full set of results is available upon request from the author.

cient for the $d = 0.20$ case (in the range 20–100%), while the increase associated with the Bartlett test is more modest (in the range 10–20%). Changes in the RMSFE statistics are less noticeable, and restricted to the case of very large $(s/n)^{-1}$ statistics (for the Bartlett test based selection procedure the improvement in the RMSFE statistics is in the range 5–20%). Selecting a lower significance level (1%) leads again to some improvement in terms of both RMSFE and ρ statistics for the $d \leq 0.40$ case for both Bartlett and Box–Pierce tests based selection rules, particularly when $(s/n)^{-1}$ is small.

3.2 Short memory

3.2.1 The unfeasible case

As shown in Table 3, also for the short memory case the FLS and WK filters, in the unfeasible case, show a similar performance, with the FLS filter performing slightly better than the WK filter for the $(s/n)^{-1} < 1$ case. In fact, the ratios of the relevant statistics are very close to unity in all the cases. Moreover, in absolute terms, the performance of the FLS filter tends to improve as the degree of persistence of the series increases. Coherent with this latter result, the FLS filter performs better in the short memory plus structural break case than in the short memory case only, particularly in the presence of sizable noise, i.e. $(s/n)^{-1} > 2$, with an increase in the ρ statistic in the range 10–150%. Similarly, the reduction in the IC coefficient falls in the range 10–50%. In particular, in terms of IC statistic, figures are below 0.30 in about 30 and 60% of the cases for the short memory and short memory plus structural change cases, respectively, and always below 0.40 and 0.60, respectively, in the remaining cases. On the other hand, in terms of ρ statistic, figures are in the range 0.32–0.36 and 0.70–0.95 for the short memory and short memory plus structural change cases, respectively, with the performance of both filters worsening as the degree of noise increases.

In addition, the FLS filter performs better in the long memory cases than in the short memory cases, both in relative and in absolute terms. In relative terms, the increase in performance is up to 20% in terms of RMSFE statistic and up to 10% in terms of ρ statistic In absolute terms figures are 30 and 25% for the IC and ρ statistics, respectively. It should be noted however that the differences between the long memory and the short memory cases tend to be small for the structural breaks case. Moreover, as already found for the long memory case, the absolute performance of the FLS filter, as the one of the WK filter, tends to worsen as the degree of noise increase.

3.2.2 The feasible case

As shown in Table 4, the proposed rule for the selection of the penalization parameter, i.e. the lowest value which yields not serially correlated residuals, yields a very satisfactory outcome, relative to the performance of the unfeasible version of the filter also for the short memory case. In general the selection of the 1% significance level yields better results than the 10% level for both the Box–Pierce and the Bartlett tests, with the two test based criteria yielding a similar performance in all the cases as well. As already found for the unfeasible case, the FLS filter tends to perform better as the

Table 3 Monte Carlo results, absolute and relative performance of the FLS filter, unfeasible case

Short memory
FLS filter, relative (WK filter) performance

RMSFE					ρ				
$(s\backslash n)^{-1} \setminus \pi$	0.2	0.4	0.6	0.8	$(s\backslash n)^{-1} \setminus \pi$	0.2	0.4	0.6	0.8
10	1.041	1.035	1.031	1.007	10	1.008	1.010	1.023	1.006
7	1.036	1.032	1.029	1.001	7	1.005	1.010	1.025	1.005
5	1.043	1.036	1.033	1.013	5	1.003	1.008	1.022	1.002
3	1.057	1.046	1.041	1.026	3	1.002	1.006	1.016	1.000
2	1.066	1.059	1.046	1.008	2	1.001	1.004	1.009	1.003
1	1.069	1.045	1.020	0.919	1	1.000	1.000	1.001	1.018
0.5	0.997	0.976	0.947	0.821	0.5	1.001	1.002	1.006	1.034
0.25	0.961	0.944	0.916	0.789	0.25	1.001	1.003	1.007	1.029
0.125	0.964	0.952	0.931	0.830	0.125	1.001	1.001	1.004	1.014

FLS filter, absolute performance

IC					ρ				
$(s\backslash n)^{-1} \setminus \pi$	0.2	0.4	0.6	0.8	$(s\backslash n)^{-1} \setminus \pi$	0.2	0.4	0.6	0.8
10	0.613	0.593	0.568	0.486	10	0.315	0.342	0.380	0.538
7	0.584	0.565	0.540	0.451	7	0.369	0.395	0.435	0.598
5	0.555	0.537	0.508	0.420	5	0.423	0.446	0.494	0.652
3	0.504	0.486	0.461	0.373	3	0.515	0.539	0.582	0.727
2	0.459	0.443	0.417	0.363	2	0.593	0.614	0.654	0.779
1	0.379	0.368	0.347	0.280	1	0.718	0.732	0.760	0.847
0.5	0.301	0.293	0.279	0.230	0.5	0.821	0.829	0.845	0.896
0.25	0.228	0.225	0.217	0.185	0.25	0.896	0.899	0.907	0.933
0.125	0.169	0.167	0.163	0.145	0.125	0.943	0.945	0.947	0.958

Short memory and structural change
FLS filter, relative (WK filter) performance

RMSFE					ρ				
$(s\backslash n)^{-1} \setminus \pi$	0.2	0.4	0.6	0.8	$(s\backslash n)^{-1} \setminus \pi$	0.2	0.4	0.6	0.8
10	1.166	1.130	1.099	1.125	10	0.952	0.964	0.974	0.973
7	1.111	1.071	1.047	1.072	7	0.947	0.971	0.985	0.978
5	1.073	1.043	1.026	1.048	5	0.937	0.977	1.003	0.980
3	1.074	1.038	1.027	1.053	3	0.957	0.986	0.996	0.984
2	1.076	1.048	1.035	1.027	2	0.964	0.981	0.991	0.993
1	1.072	1.042	0.999	0.938	1	0.985	0.994	1.004	1.013
0.5	0.998	0.971	0.908	0.839	0.5	0.999	1.003	1.015	1.028
0.25	0.959	0.940	0.880	0.806	0.25	1.002	1.004	1.013	1.023
0.125	0.962	0.948	0.911	0.843	0.125	1.001	1.002	1.006	1.011

Table 3 continued

FLS filter, absolute performance									
IC					ρ				
$(s\backslash n)^{-1} \setminus \pi$	0.2	0.4	0.6	0.8	$(s\backslash n)^{-1} \setminus \pi$	0.2	0.4	0.6	0.8
10	0.276	0.268	0.262	0.237	10	0.838	0.846	0.851	0.878
7	0.327	0.312	0.304	0.264	7	0.776	0.792	0.801	0.851
5	0.399	0.380	0.363	0.304	5	0.670	0.696	0.720	0.806
3	0.363	0.345	0.322	0.267	3	0.720	0.747	0.779	0.851
2	0.379	0.363	0.335	0.277	2	0.700	0.725	0.766	0.843
1	0.353	0.340	0.312	0.263	1	0.749	0.765	0.803	0.863
0.5	0.278	0.270	0.253	0.214	0.5	0.844	0.852	0.871	0.909
0.25	0.210	0.208	0.198	0.171	0.25	0.910	0.912	0.921	0.941
0.125	0.155	0.154	0.151	0.134	0.125	0.951	0.952	0.954	0.964

The Table reports the ratio of the root mean square forecast error (RMSFE) and the correlation coefficient (ρ) for the FLS filter, relative to the WK filter, in the unfeasible case. In the table also the inequality coefficient (IC) and the correlation coefficient for the unfeasible FLS filter have been reported. Both the pure short memory and the short memory plus structural change case are considered. The inverse signal to noise ratio ranges from 0.125 to 10, while the autoregressive parameter from 0.20 to 0.80

degree of persistence increases. With reference to the 1% significance level only and the short memory case, as the autoregressive parameter π increases from 0.20 to 0.80, the (relative) correlation coefficient almost doubles, getting very close to the unitary value, i.e. the performance of the feasible filter becomes very similar to that of the unfeasible filter. On the other hand, for the short memory plus structural change case similar changes in performance can be found only for the $(s/n)^{-1} < 2$ cases. Overall the feasible FLS filter performs best in the short memory plus structural change case, particularly in the presence of sizable noise $(s/n)^{-1} > 1$. For instance, for the Box–Pierce test based criterion (1% significance level), in the short memory case the ρ statistic is in the range 0.42 to 0.98, while the range for the short memory plus structural change case is 0.54 to 1.

Hence, on the basis of the above results, it can be concluded that the feasible FLS filter tends to perform better in the long memory case than in the short memory case, with the performance gap being much more noticeable under structural stability. In fact, by comparing the performance of the Box–Pierce based selection rule, choosing a 1% significance level, it is found that the minimum (relative) ρ statistic is about 0.64 in the long memory case, while the figure is 0.42 in the short memory case. On the other hand, figures for the structural instability cases are 0.68 and 0.54, respectively.

3.3 FLS filtering versus Kalman filtering

In order to compare the performance of the proposed filtering strategy with an alternative estimation method, the Monte Carlo exercise has been repeated by considering a feasible version of the local level model, estimated by means of the Kalman filter using the prediction error Gaussian likelihood function approach. The settings of the exercise are the same as the ones employed in the previous sections.

Table 4 Monte Carlo results, performance of the feasible FLS filter relative to the unfeasible FLS filter

Box–Pierce test (1%)
Short memory

RMSFE					ρ				
$(s\backslash n)^{-1}\ \backslash\pi$	0.2	0.4	0.6	0.8	$(s\backslash n)^{-1}\ \backslash\pi$	0.2	0.4	0.6	0.8
10	1.091	1.081	1.067	1.048	10	0.508	0.661	0.787	0.942
7	1.048	1.038	1.029	1.024	7	0.527	0.657	0.776	0.950
5	1.019	1.015	1.012	1.017	5	0.531	0.652	0.774	0.956
3	1.018	1.024	1.036	1.023	3	0.528	0.650	0.756	0.963
2	1.068	1.083	1.090	1.035	2	0.535	0.614	0.744	0.969
1	1.242	1.275	1.234	1.061	1	0.523	0.554	0.738	0.974
0.5	1.540	1.597	1.387	1.106	0.5	0.493	0.481	0.789	0.975
0.25	2.005	1.942	1.524	1.197	0.25	0.472	0.547	0.857	0.972
0.125	2.765	2.262	1.809	1.325	0.125	0.415	0.705	0.881	0.971

Short memory and structural change

RMSFE					ρ				
$(s\backslash n)^{-1}\ \backslash\pi$	0.2	0.4	0.6	0.8	$(s\backslash n)^{-1}\ \backslash\pi$	0.2	0.4	0.6	0.8
10	1.027	1.026	1.031	1.021	10	0.992	0.993	0.992	0.995
7	1.019	1.017	1.020	1.018	7	0.992	0.994	0.991	0.994
5	1.012	1.010	1.016	1.017	5	0.989	0.987	0.976	0.989
3	1.022	1.029	1.060	1.026	3	0.961	0.963	0.948	0.988
2	1.067	1.066	1.082	1.035	2	0.866	0.901	0.917	0.983
1	1.247	1.214	1.110	1.064	1	0.663	0.766	0.922	0.978
0.5	1.554	1.468	1.202	1.118	0.5	0.612	0.735	0.931	0.976
0.25	2.039	1.830	1.351	1.207	0.25	0.568	0.727	0.934	0.975
0.125	2.774	2.376	1.580	1.353	0.125	0.542	0.732	0.935	0.973

Bartlett test (1%)

Short memory

RMSFE					ρ				
$(s\backslash n)^{-1}\ \backslash d$	0.2	0.4	0.6	0.8	$(s\backslash n)^{-1}\ \backslash d$	0.2	0.4	0.6	0.8
10	1.095	1.079	1.059	1.050	10	0.516	0.683	0.813	0.955
7	1.061	1.035	1.016	1.018	7	0.531	0.696	0.830	0.960
5	1.027	1.010	1.008	1.007	5	0.533	0.710	0.847	0.966
3	1.018	1.016	1.023	1.012	3	0.539	0.716	0.865	0.967
2	1.070	1.057	1.051	1.027	2	0.551	0.721	0.883	0.969
1	1.242	1.195	1.117	1.057	1	0.559	0.725	0.904	0.971
0.5	1.521	1.421	1.251	1.118	0.5	0.564	0.725	0.904	0.968
0.25	1.912	1.768	1.460	1.243	0.25	0.586	0.724	0.902	0.960
0.125	2.417	2.309	1.708	1.448	0.125	0.620	0.721	0.906	0.954

Table 4 continued

Short memory and structural change									
RMSFE					ρ				
$(s\backslash n)^{-1} \backslash\pi$	0.2	0.4	0.6	0.8	$(s\backslash n)^{-1} \backslash\pi$	0.2	0.4	0.6	0.8
10	1.024	1.013	1.024	1.013	10	0.993	0.997	0.995	0.997
7	1.014	1.006	1.014	1.008	7	0.995	0.997	0.996	0.997
5	1.008	1.008	1.005	1.007	5	0.989	0.986	0.991	0.993
3	1.017	1.025	1.002	1.018	3	0.959	0.972	0.978	0.990
2	1.057	1.056	1.052	1.033	2	0.847	0.930	0.947	0.980
1	1.216	1.156	1.125	1.071	1	0.655	0.856	0.913	0.971
0.5	1.489	1.339	1.274	1.153	0.5	0.630	0.842	0.901	0.966
0.25	1.934	1.649	1.508	1.288	0.25	0.612	0.827	0.892	0.961
0.125	2.603	2.145	1.841	1.524	0.125	0.599	0.814	0.887	0.954

The Table reports the ratio of the root mean square forecast error (RMSFE) and the correlation coefficient (ρ) for the feasible FLS filter, relative to the unfeasible FLS filter. Both the pure short memory and the short memory plus structural change case are considered. The inverse signal to noise ratio ranges from 0.125 to 10, while the autoregressive parameter from 0.20 to 0.80

As is shown in Table 5, the results are qualitatively similar across specifications, pointing to a superior performance of the FLS approach, relatively to the Kalman filter approach, both when observational noise is sizable, i.e. $(s/n)^{-1} > 1$, and as the persistence of the process increases. Hence, in general the FLS filter tends to perform better than the Kalman filter for the long memory case and for the structural break cases. The superior performance of the FLS filter is surely noticeable when noisiness is sizable ($(s/n)^{-1} > 1$). For instance, for the long memory case the Kalman filter leads to an increase in RMSFE in the range 10–360%, while for the long memory plus structural change case the increase in the statistic is in the range 15–290%. For these two latter cases the Kalman filter shows an inferior performance to the FLS filter in 84% of the cases. A similar picture holds for the short memory cases, with the FLS filter outperforming the Kalman filter in over 70% of the cases. Similar conclusions, still for the $(s/n)^{-1} > 1$ case, can also be drawn on the basis of the correlation coefficient, albeit according to this latter statistic, particularly for the low persistence case ($d = 0.20$ and $\pi = 0.20$) and under structural stability, the Kalman filter does seem to provide a comparable, if not slightly superior, performance than the FLS filter.

Given the Gaussian setting of the exercise, overall the different performance of the two approaches can be explained on the basis of the different strategy followed for the selection of the signal-noise ratio, suggesting that the proposed approach is not only very satisfactory in general, but also superior to alternative methods when filtering is actually needed, i.e. when measurement error is sizable.[5]

[5] As already clarified in the methodological section, the key difference between the two filters in the Gaussian setting amounts to the fact that the signal-noise ratio is estimated by means of ML estimation when the prediction error decomposition is employed, as it is usual when Kalman filtering is carried out, while a data driven selection approach for the signal-noise ratio is employed for the FLS filter.

Table 5 Monte Carlo results, relative performance of the Kalman filter, feasible case

Long memory
Kalman filter, relative (FLS filter) performance

RMSFE						ρ					
$(s\backslash n)^{-1}\backslash d$	0.2	0.4	0.6	0.8	1	$(s\backslash n)^{-1}\backslash d$	0.2	0.4	0.6	0.8	1
10	2.823	3.225	4.102	4.001	4.577	10	1.467	0.583	0.218	0.226	0.187
7	2.497	2.901	3.849	3.775	4.337	7	1.521	0.656	0.287	0.293	0.254
5	2.193	2.600	3.584	3.497	4.239	5	1.549	0.737	0.373	0.378	0.339
3	1.786	2.138	3.148	3.084	3.933	3	1.558	0.870	0.552	0.563	0.512
2	1.501	1.800	2.787	2.710	3.684	2	1.525	0.962	0.717	0.726	0.678
1	1.099	1.291	2.194	2.125	3.161	1	1.386	1.049	0.963	0.962	0.927
0.5	0.798	0.909	1.670	1.583	2.630	0.5	1.159	1.008	1.071	1.074	1.058
0.25	0.575	0.695	1.242	1.129	2.093	0.25	0.901	0.872	1.054	1.064	1.077
0.125	0.417	0.583	0.899	0.826	1.607	0.125	0.682	0.737	0.989	1.009	1.048

Long memory and structural change
Kalman filter, relative (FLS filter) performance

IC						ρ					
$(s\backslash n)^{-1}\backslash d$	0.2	0.4	0.6	0.8	1	$(s\backslash n)^{-1}\backslash d$	0.2	0.4	0.6	0.8	1
10	2.710	3.069	3.398	3.758	3.809	10	0.669	0.619	0.569	0.647	0.636
7	2.469	2.873	3.208	3.722	3.872	7	0.696	0.619	0.561	0.648	0.630
5	2.229	2.648	3.048	3.687	3.854	5	0.758	0.629	0.566	0.628	0.618
3	1.816	2.214	2.573	3.311	3.548	3	0.860	0.729	0.660	0.727	0.703
2	1.546	1.911	2.260	3.093	3.428	2	0.988	0.784	0.720	0.753	0.724
1	1.146	1.435	1.705	2.573	3.082	1	1.232	0.914	0.853	0.829	0.791
0.5	0.831	1.033	1.237	2.003	2.589	0.5	1.338	1.017	0.961	0.918	0.884
0.25	0.594	0.744	0.897	1.477	2.104	0.25	1.414	1.076	1.023	0.973	0.946
0.125	0.424	0.569	0.724	1.117	1.654	0.125	1.463	1.111	1.044	0.996	0.981

Short memory
Kalman filter, relative (FLS filter) performance

RMSFE					ρ				
$(s\backslash n)^{-1}\backslash\pi$	0.2	0.4	0.6	0.8	$(s\backslash n)^{-1}\backslash\pi$	0.2	0.4	0.6	0.8
10	2.586	3.013	3.443	4.056	10	0.701	0.510	0.435	0.342
7	2.424	2.886	3.344	4.072	7	0.839	0.620	0.528	0.399
5	2.224	2.700	3.182	4.000	5	1.093	0.756	0.647	0.460
3	1.833	2.231	2.596	3.388	3	1.266	0.900	0.789	0.561
2	1.529	1.837	2.151	3.007	2	1.496	1.127	0.939	0.648
1	1.097	1.241	1.419	2.231	1	1.783	1.532	1.153	0.814
0.5	0.783	0.808	0.919	1.586	0.5	1.950	1.925	1.206	0.925
0.25	0.551	0.575	0.761	1.234	0.25	2.081	1.768	1.139	0.977
0.125	0.382	0.458	0.770	1.326	0.125	2.387	1.398	1.095	0.975

Table 5 continued

Short memory and structural change

Kalman filter, relative (FLS filter) performance

IC					ρ				
$(s\backslash n)^{-1}\ \backslash\pi$	0.2	0.4	0.6	0.8	$(s\backslash n)^{-1}\ \backslash\pi$	0.2	0.4	0.6	0.8
10	2.607	2.704	2.805	2.888	10	0.682	0.667	0.661	0.662
7	2.326	2.453	2.554	2.697	7	0.725	0.699	0.693	0.675
5	2.095	2.202	2.319	2.503	5	0.808	0.769	0.751	0.691
3	1.684	1.795	1.887	2.124	3	0.919	0.873	0.856	0.784
2	1.431	1.533	1.660	1.883	2	1.086	1.002	0.939	0.832
1	1.035	1.138	1.293	1.443	1	1.489	1.254	1.015	0.930
0.5	0.732	0.823	0.935	1.033	0.5	1.627	1.342	1.065	1.009
0.25	0.513	0.591	0.766	0.822	0.25	1.762	1.373	1.070	1.029
0.125	0.365	0.432	0.769	0.871	0.125	1.850	1.366	1.050	1.015

The Table reports the ratio of the root mean square forecast error (RMSFE) and the correlation coefficient (ρ)for the Kalman filter, relative to the FLS filter, in the feasible case (Box–Pierce test, 1% significance level). Both the pure long/short memory and the long/short memory plus structural change cases are considered. The inverse signal to noise ratio ranges from 0.125 to 10, the fractional differencing parameter from 0.20 to 1, and the autoregressive parameter from 0.20 to 0.80

4 Empirical application

Monthly time series data for four exchange rates, i.e. the euro/US$ rate, the yen/US$ rate, the GBP£/US$ rate, and the Canadian $/US$ rate, over the period 1980:1–2006:6, have been investigated. Monthly noisy (log) volatility proxies for the above variables have been constructed as the (log) absolute value of the innovations of the various exchange rate log returns, obtained from the estimation of a standard VAR model for the 4 variables in the data set, with lag length set to two lags on the basis of misspecification tests and the AIC criterion. Noise filtering has then been carried out using the FLS approach described in the methodological section. Following the Monte Carlo results, the Box–Pierce test optimal filtering based procedure has been employed, setting the significance level of the test to 1%. In Fig. 1 the actual and filtered exchange rate volatility series are reported. As is shown in the figure, the estimated signal components are not affected by the large noise component affecting the data, well capturing the trend behavior in the series. By comparing the descriptive statistics reported in Table 6, it can also be noted that filtering does not alter the unconditional expectation of the series, leading, on the other hand, to a large reduction in the standard deviation of the series, in the range 170–360%. Coherent with this latter finding, as shown by the reduction in the spread between the minimum and maximum values and in the excess kurtosis of the series, the range of variation of the series is sizably decreased by filtering, while changes in skewness are in general much less noticeable.

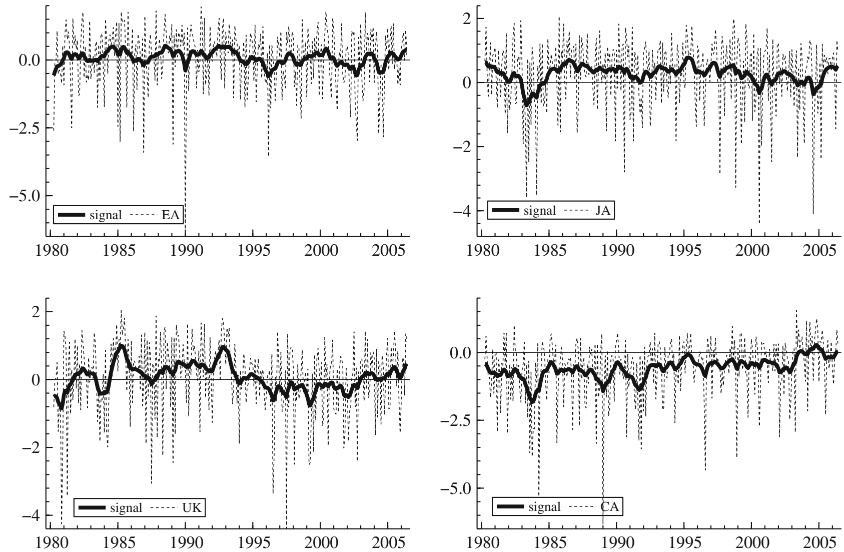

Fig. 1 Actual and filtered (signal) exchange rate log volatility series (e/US$: EA, yen/US$: JA, GBP£/US$: UK, Canadian $/US$: CA)

Table 6 Actual and filtered data, descriptive statistics

	ACTUAL				FILTERED			
	EA	JP	UK	CA	EA	JP	UK	CA
Mean	0.082	0.271	0.064	−0.584	0.082	0.271	0.064	−0.584
SD	0.246	0.267	0.372	0.371	1.128	1.091	1.067	1.110
SKW	−0.469	−1.062	0.169	−0.559	-1.255	−1.296	−1.028	−1.485
KURT	−0.176	1.359	−0.156	0.977	2.931	2.173	1.531	3.266
Min	−0.598	−0.713	−0.855	−1.833	-6.237	−4.426	−4.389	−6.337
Max	0.521	0.777	0.999	0.276	1.957	2.072	2.033	1.543

The Table reports descriptive statistics for the actual and filtered exchange rate volatility series (€ /US$: EA, yen/US$: JA, GBP£/US$: UK, Canadian $/US$: CA). Statistics refer to the mean (MEAN), standard deviation (S.D.), skewness (SKW), excess kurtosis (KURT), and the minimum (MIN) and maximum (MAX) values, respectively

5 Conclusions

In the paper a new noise filtering approach, based on FLS estimation of an unobserved component local level model, has been introduced. The proposed FLS filter has been found to perform well in Monte Carlo analysis, independently of the persistence properties of the data and the size of the signal to noise ratio, outperforming in general even the WK filter, which, theoretically, is a minimum mean square estimator. Moreover, a key advantage of the proposed filter, relatively to available competitors, is that any persistence property of the data can be handled, without any pretesting, being computationally fast and not demanding, and easy to be implemented as well.

References

Arino MA, Marmol F (2004) A permanent-transitory decomposition for ARFIMA processes. J Stat Plan Inference 124:87–97

Baillie RT (1996) Long memory processes and fractional integration in econometrics. J Econom 73:5–59

Beltratti A, Morana C (2006) Breaks and persistence: macroeconomic causes of stock market volatility. J Econom 131:151–177

Beveridge S, Nelson CR (1981) A new approach to decomposition of economic time series into permanent and transitory components with particular attention to measurement of the business cycle. J Monet Econ 7:151–174

Engle RF, Granger CWJ (1987) Co-integration and error correction representation. Estimation Test Econom 55:251–276

Gonzalo J, Granger C (1995) Estimation of common long-memory components in cointegrated systems. J Bus Econ Stat 13(1):27–35

Harvey AC (1998) Long memory in stochastic volatility. In: Knight J, Satchell S (eds) Forecasting volatility in financial markets. Butterworth-Heineman, Oxford, pp 307–320

Harvey AC (1989) Forecasting, structural time series models and the Kalman filter. Cambrdige University Press, Cambridge

Kasa K (1992) Common stochastic trends in international stock markets. J Monet Econ 29:95–124

Kalaba R, Tesfatsion L (1989) Time-varying linear regression via flexible least squares. Comput Math Appl 17:1215–1245

Kalaba R, Tesfatsion L (1990a) An organizing principle for dynamic estimation. J Optim Theory Appl 64(3):445–470

Kalaba R, Tesfatsion L (1990b) Flexible least squares for approximately linear systems. IEEE Trans Syst Man Cybern 20:978–989

Kladroba A (2005) Flexible least squares estimation of state space models: an alternative to Kalman filtering? University of Duisburg-Essen, Faculty of Economics, Working Paper, no.149, Essen

Marinucci D, Robinson PM (2001) Semiparametric fractional cointegration analysis. J Econom 105: 225–247

Morana C (2002) Common persistent factors in inflation and excess nominal money growth and a new measure of core inflation. Stud Nonlinear Dyn Econom 6(3):art.3–5

Morana C (2004) Frequency domain principal components estimation of fractionally. Cointegrated Process Appl Econ Lett 11:837–842

Morana C (2006) A small scale macroeconometric model for Euro-12 area. Econ Model 23(3):391–426

Morana C (2007) Multivariate modelling of long memory processes with common components. Comput Stat Data Anal 52:919–934

Priestley MB (1981) Spectral analysis and time series. Academic Press, London

Robinson PM, Yajima Y (2002) Determination of cointegrating rank in fractional systems. J Econom 106(2):217–241