Journal of Forecasting
J. Forecast. **24**, 1–16 (2005)
Published online in Wiley InterScience (www.interscience.wiley.com). **DOI**: 10.1002/for.937

Forecasting Time Series with Long Memory and Level Shifts

NAMWON HYUNG[1] AND PHILIP HANS FRANSES[2]*
[1] *University of Seoul, Korea*
[2] *Erasmus University Rotterdam, The Netherlands*

ABSTRACT
It is well known that some economic time series can be described by models which allow for either long memory or for occasional level shifts. In this paper we propose to examine the relative merits of these models by introducing a new model, which jointly captures the two features. We discuss representation and estimation. Using simulations, we demonstrate its forecasting ability, relative to the one-feature models, both in terms of point forecasts and interval forecasts. We illustrate the model for daily S&P500 volatility. Copyright © 2005 John Wiley & Sons, Ltd.

KEY WORDS long memory; level shifts; forecasting; stock returns

INTRODUCTION

There appear to be several economic time series, which can be characterized by either long memory or by occasional level shifts. Examples of these are inflation rates and certain financial volatility series, see, for example Bos *et al*. (1999). On the theoretical side, studies like Granger and Hyung (2001) and Diebold and Inoue (2001) have shown that apparent long memory can also be caused by neglected occasional level shifts, while other studies indicate the reversal effects. Hence, the empirical results for certain economic series should not come as a surprise, as models for long memory and models for occasional level shifts seem to be able to pick up the same features of empirical data.

In this paper we also consider the two types of models, although our approach differs from all previous studies in two respects. First, we not only consider in-sample fit, but also report on the results of an extensive simulation experiment concerning out-of-sample forecasting. This forecasting exercise includes point and interval forecasts. Second, we do not consider the two mentioned models separately, as we propose a joint model, which we call the FI-BREAK model.

The outline of the paper is as follows. In the next section, we discuss the representation of the various models and parameter estimation. In the third section, we outline the salient features of our simulation design and report our results for out-of-sample forecasts. We also compare the effects of misspecification on in-sample parameter estimates. In the fourth section, we examine if the

* Correspondence to: Philip Hans Franses, Econometric Institute, Erasmus University Rotterdam, PO Box 1738, NL-3000 DR Rotterdam, The Netherlands. E-mail: franses@few.eur.nt

simulation results carry though to empirical data, where we focus on volatilities of stock returns. A final section concludes.

THE FI-BREAK MODEL

We exploit the possibility that the occasional structural break model (or simply BREAK model) and I(d) model can be summarized into one single model. One motivation for this joint model is that both individual models can capture a long memory component to some extent, and hence a joint model would be able to capture all long memory components. To construct such a joint model for a time series y_t, one can think of

$$\begin{aligned} (1-L)^d y_t &= m_t + \varepsilon_t \\ m_t &= m_{t-1} + q_t \eta_t \end{aligned} \tag{1}$$

where q_t follows an i.i.d. binominal distribution, that is

$$q_t = \begin{cases} 1, & \text{with probability } p \\ 0, & \text{with probability } 1-p \end{cases} \tag{2}$$

For simplicity, we assume $\varepsilon_t \sim$ i.i.d. $(0, \sigma_\varepsilon^2)$ and $\eta_t \sim$ i.i.d. $(0, \sigma_\eta^2)$. Equation (1) can be rewritten as $y_t = \tilde{m}_t + u_t$, where $(1 - L)^d \tilde{m}_t = m_t$ and $(1 - L)^d u_t = \varepsilon_t$. The time series y_t can be decomposed into a long memory break component $\tilde{m}_t$ and a long memory component u_t.

Another related model is the STOPBREAK model of Engle and Smith (1999), that is

$$\begin{aligned} y_t &= m_t + \varepsilon_t \\ m_t &= m_{t-1} + q_{t-1}\varepsilon_{t-1} \end{aligned} \tag{3}$$

where the function q_t is specified as

$$q_t = \frac{(\varepsilon_t + \ldots + \varepsilon_{t-s+1})^2}{\gamma + (\varepsilon_t + \ldots + \varepsilon_{t-s+1})^2}$$

for some value $s > 0$, see Smith (2000) for further details. This model includes an endogenous smooth transition function to indicate structural breaks, and it can be seen as a contender to the discrete break model.

Note that the model in (1) generalizes the model put forward in van Dijk *et al*. (2002), which is

$$\begin{aligned} (1-L)^d y_t &= x_t \\ x_t &= \phi_1(L)x_{t-1}(1-G(s_t;\ \gamma,\ c)) + \phi_2(L)x_{t-1}G(s_t;\ \gamma,\ c) + \varepsilon_t \end{aligned} \tag{4}$$

where the transition function $G(\cdot)$ is assumed to be the logistic function

$$G(s_t;\ \gamma,\ c) = [1 + \exp\{-\gamma(s_t - c)/\sigma_{st}\}]^{-1}$$

where $\gamma > 0$, s_t is the transition variable and σ_{st} is the standard deviation of s_t. This fractionally integrated smooth transition autoregressive model allows for only two different regimes corresponding to $G(\cdot) = 0$ and $G(\cdot) = 1$.

Based on the discussion above, we decide to focus on the following representation of a FI-BREAK model, that is

$$\begin{aligned}\alpha(L)(1-L)^d y_t &= m_t + \varepsilon_t \\ m_t &= m_{t-1} + q_{t-1}\varepsilon_{t-1} \\ q_t &= \frac{(\varepsilon_t + \ldots + \varepsilon_{t-s+1})^2}{\gamma + (\varepsilon_t + \ldots + \varepsilon_{t-s+1})^2}\end{aligned} \tag{5}$$

where $\alpha(L) = (1 - \alpha_1 L - \ldots - \alpha_p L^p)$. We assume that m_0 is fixed and known. This term determines the unconditional mean of the process.

This general model can be seen to nest several related models by imposing certain parameter restrictions.

I. When $d = 0$ and $\gamma \to \infty$, the model becomes an AR(p) model. Indeed, as $\gamma \to \infty$, $q_t = 0$ for all t, which implies that $m_t = m_0$ for all t. Furthermore, if additionally $d = 0$, one has $\alpha(L)y_t = m_0 + \varepsilon_t$.
II. When $0 < d < 1$ and $\gamma \to \infty$, the model becomes an ARFI(p, d) model, that is, $\alpha(L)(1 - L)^d y_t = m_0 + \varepsilon_t$.
III. When $d = 0$ and $0 < \gamma < \infty$, the model is the familiar STOPBREAK model. If $\gamma < \infty$, this process contains an endogenous smooth break.
IV. When $0 < d < 1$ and $0 < \gamma < \infty$, the FI-BREAK model combines an I(d) model and a break model.

We summarize the various results in Table I. In this paper, we consider only AR type dynamics, for estimation convenience. Note that there are also other parameter combinations, such as $\gamma = 0$ and $d = 1$, but we choose to consider only the above models I, II, III and IV. Indeed, if we know the degree of integration of a series (1 or 2), we could take proper differences and return to one of the models above.

We can rewrite the FI-BREAK model (5) as

$$\sum_{j=0}^{\infty} \pi_j \Delta y_{t-j} = \varepsilon_t - \theta_{t-1}\varepsilon_{t-1}$$

Table I. Parameter space and models

	$d = 0$	$0 < d < 1$	$d = 1$
$\gamma \to 0$	ARI(p, 1)	ARFI(p, $1 + d$)	ARI(p, 2)
$0 < \gamma < \infty$	STOPBREAK	FI-BREAK	Integrated BREAK
$\gamma \to \infty$	AR(p)	ARFI(p, d)	ARI(p, 1)

where $\theta_{t-1} = 1 - q_{t-1}$ and $\sum_{j=0}^{\infty} \pi_j L^j = (1 - \alpha_1 L - \ldots - \alpha_p L^p)(1-L)^d$. Using Theorem 1 in Engle and Smith (1999), we can show that this nonlinear moving average process is invertible with probability 1 if prob($q_t > 0$) > 0 and if

$$q_t \left[1 + \frac{\varepsilon_t}{q_t} \frac{\partial q_t}{\partial \varepsilon_t} \middle| \varepsilon_t \right] < 2$$

with probability 1. Hence, we can estimate the model parameters using the AML method of Beran (1995), see also van Dijk *et al*. (2002). Along similar lines, the approximate time-domain maximum likelihood (AML) estimator for the FI-BREAK model is consistent and asymptotically normal. Unreported simulation results support this.

SIMULATION

In this section we rely on extensive Monte Carlo simulations to examine the relative performance of the FI-BREAK model. We consider in-sample fit and out-of-sample forecasting.

Design

We simulate nine different types of time series. In all cases the simulated noise, ε_t and η_t, is generated from the standard Gaussian distribution with variances $\sigma^2_\varepsilon = 1$ and σ^2_η.

I. Fractionally integrated process

$$(1-L)^d y_t = \mu + \varepsilon_t$$

DGP(1): $d = 0.1$. This is persistent fractional noise with all autocorrelation coefficients positive. The autocorrelation decays at a slow hyperbolic rate but the long memory properties are not very prominent.
DGP(2): $d = 0.4$. Now the series has prominent long memory properties.
DGP(3): $d = 0.7$. This implies non-stationarity ($d > 0.5$), but the limiting value of the impulse response function is equal to 0, such that shocks do not have permanent effects.

II. Break process

$$y_t = m_t + \varepsilon_t, \quad m_t = m_{t-1} + q_t \eta_t$$

where q_t follows an i.i.d. binominal distribution, that is, $q_t = 1$ with probability p, $q_t = 0$ otherwise.

DGP(4): $p = 0.01$, $\sigma^2_\eta = 0.1$. The expected number of breaks in 300 observations is 3.
DGP(5): $p = 0.01$, $\sigma^2_\eta = 0.5$. The expected number of breaks in 300 observations is 3 with larger size of the breaks than DGP(4).
DGP(6): $p = 0.03$, $\sigma^2_\eta = 0.5$. The expected number of breaks in 300 observations is 9.

III. Mixture of fractional integration and break process

$$(1-L)^d y_t = m_t + \varepsilon_t, \quad m_t = m_{t-1} + q_t \eta_t$$

where q_t follows an i.i.d. binominal distribution, that is, $q_t = 1$ with probability $p = 0.01$ and $q_t = 0$ otherwise.

DGP(7): $\sigma_\eta^2 = 0.5$, $d = 0.1$. A time series with a weak long memory component but with clear (visual) breaks.
DGP(8): $\sigma_\eta^2 = 0.5$, $d = 0.4$. Time series with clear breaks and evident long memory.
DGP(9): $\sigma_\eta^2 = 0.1$, $d = 0.4$. Data with evident long memory but with weak evidence of breaks.

For each model, we generate 1000 time series, each having length $T = 324$, and we estimate model parameters for the sample size $T = 300$, which would be equivalent to 25 years for monthly data. In each case, we use 24 additional data points for assessing out-of-sample forecasting performance. This allows us to see how modelling the long memory characteristic might aid long-term forecasting.

We use the following estimation procedures for the AR, ARFI model, BREAK model and FI-BREAK model:

(1) The AR model by the least squares method, that is

$$\alpha(L)y_t = \mu + \varepsilon_t$$

where $\alpha(L) = 1 - \alpha_1 L - \ldots - \alpha_p L^p$.

(2) The ARFI model by the AML method, that is, Beran's (1995) approximate time-domain maximum likelihood procedure. The model reads as

$$\alpha(L)(1-L)^d y_t = \mu + \varepsilon_t$$

and prior to estimating d the sample mean is subtracted first.

(3) The break model using Bai and Perron's (1998) method, that is, sequential estimation of the break points. The model is given by

$$\alpha(L)y_t = \mu_t + \varepsilon_t$$

where $\mu_t = \mu_r$ for $r \in (k_{r-1}, k_r]$; $r = 1, \ldots, R+1$, such that $0 = k_0 < k_1 < \ldots < k_R < k_{R+1} = T$. The number of breaks R is treated as an unknown variable. We identify R break points with dynamic components in a parametric model, by introducing lagged dependent variables so as to have an autoregressive model. We use this model instead of the STOPBREAK model, as this break model signals breaks better in empirical analysis of possibly long memory data.

(4) Finally, our FI-BREAK model, for which we also use the AML method. We set $s = 12$ in (5), allowing that the permanent breaks do not all occur in one period, but that it may take time for a permanent shock to filter through, for example, one year for macro data.

For each series, final models are selected using AIC or BIC. It might be that higher-order AR models do a reasonable job of capturing the long memory property. Crato and Ray (1996) find that AIC generally selects lower-order AR models ($p \leq 4$) for an ARFIMA process, even when the amount of persistence is strong and p is allowed to range up to 20.

For out-of-sample forecasting from the ARFI model, we consider the 'naïve' method, see Crato and Ray (1996). A similar technique is used for the FI-BREAK model. We assume that the current level will not change in the future, and then we construct forecasts by using the 'naïve' method to solve out the fractional integration part. To compare the forecasting performance of the selected models, we compute the MSFE statistic of one-step and multi-step ahead forecasts. For the latter, we take aboard all forecast error up to and including the horizon h. We also examine the statistical performance of the interval forecasts. Christoffersen (1998) suggests that a 'good' interval forecast should have correct conditional coverage. As the scheme of our simulation study ensures independence between forecast errors, we only check the correct unconditional coverage in the simulations.

Results

Table II presents the most salient estimation results for the simulated series. The entries are values averaged over 1000 replications with sample size 300, as we save 24 observations for out-of-sample experiments. The first and second numbers of each entry are results for the models selected by AIC and BIC, respectively. The values in parentheses are the standard deviations for the simulated data. The numbers of R in the FI-BREAK model are calculated by Σq_t.

When the DGP is fractionally integrated without breaks as DGP(1), DGP(2) and DGP(3), spurious breaks appear to get detected by Bai and Perron's (1998) method. Simulation results show a positive relation between the average number of estimated breaks and the value of d in DGPs.

The reverse phenomenon can be observed as well. We can observe a clear upward bias for DGP(4), DGP(5) and DGP(6) in the estimation of d of the ARFI model, which is of course due to neglected level shifts. Note that the FI-BREAK model appears to be quite successful in filtering out the break components when estimating d. When there are breaks as for DGP(4), DGP(5) and DGP(6), the estimated values of d are very close to zero, although the estimated values of d in the ARFI model deviate from zero significantly.

When the time series contain a break and long memory components like for DGP(7) and DGP(9), the FI-BREAK model is successful for the estimation of d. For DGP(8), that is the DGP with strong long memory but with a weak break component, the estimation of d in the FI-BREAK model is poor, but still much better than in the ARFI model. However, our unreported simulation results for larger samples indicate that the estimation method yields more reliable estimates of d for the FI-BREAK model.

Finally, at least for these DGPs, we find that BIC leads to models that perform considerably better than when the AIC is used. This confirms the results reported in Crato and Ray (1996).

Tables III to V summarize the root mean squared forecast errors (RMSFE) for various horizons. In Table III we present the forecasting performance of three models relative to the ARFI model. The best model for these I(d) data is again the ARFI model, as we observe for all models that the ratios (except for one) are higher than 0. However, the FI-BREAK model seems to lead to the smallest differences in forecast errors.

The results in Table IV suggest that the FI-BREAK model works better than the BREAK model, even though data were generated by a break process. Clearly, AR and ARFI models suffer from bias and consequently produce poor forecasts. In the presence of structural breaks, the BREAK model

Table II. In-sample estimation results of d and R

	DGP(1)		DGP(2)		DGP(3)	
	$d = 0.1$	$R = 0$	$d = 0.4$	$R = 0$	$d = 0.7$	$R = 0$
ARFI	0.038 (0.137)	–	0.291 (0.231)	–	0.585 (0.263)	–
	0.077 (0.073)		0.362 (0.130)		0.651 (0.170)	
BREAK	–	0.41 (0.66)	–	1.95 (2.01)	–	1.48 (1.64)
		0.33 (0.59)		1.08 (1.14)		1.10 (1.07)
FI-BREAK	0.071 (0.062)	0.15 (0.78)	0.308 (0.153)	0.39 (1.85)	0.471 (0.302)	1.24 (3.87)
	0.081 (0.052)	0.17 (0.99)	0.368 (0.087)	0.23 (1.27)	0.607 (0.215)	0.72 (2.76)
	DGP(4)		DGP(5)		DGP(6)	
	$d = 0$	$R = 3$	$d = 0$	$R = 3$	$d = 0$	$R = 9$
ARFI	0.069 (0.143)	–	0.217 (0.202)	–	0.412 (0.169)	–
	0.049 (0.095)		0.176 (0.171)		0.349 (0.175)	
BREAK	–	0.74 (0.72)	–	1.31 (0.99)	–	2.56 (1.21)
		0.65 (0.67)		1.23 (0.93)		2.43 (1.18)
FI-BREAK	0.043 (0.066)	1.25 (1.96)	0.070 (0.099)	3.33 (3.66)	0.142 (0.152)	6.49 (6.22)
	0.027 (0.038)	1.38 (2.05)	0.047 (0.054)	3.77 (3.91)	0.094 (0.092)	7.73 (6.27)
	DGP(7)		DGP(8)		DGP(9)	
	$d = 0.1$	$R = 3$	$d = 0.4$	$R = 3$	$d = 0.4$	$R = 3$
ARFI	0.366 (0.192)	–	0.799 (0.148)	–	0.627 (0.202)	–
	0.322 (0.177)		0.787 (0.144)		0.619 (0.182)	
BREAK	–	2.06 (1.15)	–	3.89 (3.23)	–	4.17 (2.43)
		1.96 (1.15)		2.06 (2.17)		2.74 (1.88)
FI-BREAK	0.140 (0.128)	4.71 (4.64)	0.570 (0.337)	5.41 (10.74)	0.408 (0.298)	4.97 (8.29)
	0.131 (0.082)	5.21 (4.91)	0.538 (0.336)	8.18 (17.03)	0.394 (0.276)	6.60 (10.67)

Note: The entries are values averaged over 1000 replications with sample size 300. The first and second numbers of each entry are results from the model selected by AIC and BIC. The values in parentheses are the standard deviation for the simulated data. The numbers of R in the FI-BREAK model are calculated by Σq_t.

Table III. Root mean squared forecast errors relative to an ARFI model (as a percentage)

DGP		h-Step	Model		
			BREAK	FI-BREAK	AR
DGP(1)	AIC	1	2.4	–0.3	0.6
$d = 0.1$		6	14.9	1.2	1.4
		12	29.1	2.2	0.9
		24	49.9	5.1	1.0
	BIC	1	3.2	0.0	1.7
		6	14.1	1.3	1.8
		12	26.8	2.7	2.2
		24	45.1	4.1	2.6
DGP(2)	AIC	1	8.5	0.3	2.0
$d = 0.4$		6	23.2	2.4	6.9
		12	32.7	4.3	9.7
		24	45.0	6.8	14.1
	BIC	1	7.6	0.2	3.6
		6	23.5	2.7	13.7
		12	35.0	5.1	21.6
		24	44.7	7.8	27.3
DGP(3)	AIC	1	8.1	0.9	2.1
$d = 0.7$		6	18.7	4.5	7.9
		12	27.2	7.5	12.2
		24	33.9	10.8	20.2
	BIC	1	4.3	0.9	2.9
		6	16.5	3.1	10.9
		12	26.8	5.3	21.2
		24	33.6	8.4	36.7

Note: The entries are values averaged over 1000 replications. The entries are calculated by $\{(\text{RMSFE of a model})/(\text{RMSFE of ARFI}) - 1\} \times 100$.

can give better forecasts than linear models as it can adjust its forecasts immediately when the structural changes are detected. As the size of breaks gets larger, the worst predictions are made by the ARFI model and the AR model as these models cannot adjust to such breaks.

In Table V, we examine the case where the data are generated to have both fractional integration and break components. When d is close to zero with breaks like DGP(7), the BREAK model is almost as good as the FI-BREAK model. Otherwise, except for this case, none of the models can produce better forecasts than the FI-BREAK model itself. In sum, these simulation results suggest that it is worthwhile to consider the FI-BREAK model for forecasting, also if the data have only breaks or only long memory.

Finally, Table VI gives the results of the interval evaluation exercise. We only test for the adequacy of unconditional coverage, as our simulation entails independence between forecast errors by construction. From independently generated series, we calculate only one set of 1- to 24-step forecasts. The coverage rates for the ARFI model are approximately correct, although sometimes too narrow when the DGPs contain breaks. The coverage rates for the BREAK model are too narrow

Table IV. Root mean squared forecast errors relative to a BREAK model (as a percentage)

DGP		h-Step	Model		
			ARFI	FI-BREAK	AR
DGP(4)	AIC	1	–0.1	–2.0	1.5
$R = 3$		6	0.5	–6.7	7.2
		12	4.2	–9.6	15.3
		24	8.1	–12.9	22.9
	BIC	1	0.1	–2.2	2.3
		6	5.6	–4.5	20.1
		12	14.8	–5.5	39.7
		24	21.0	–8.8	51.5
DGP(5)	AIC	1	0.8	–1.9	3.7
$R = 3$		6	15.3	–0.7	26.2
		12	33.2	3.1	54.6
		24	48.9	5.7	80.4
	BIC	1	3.8	–0.2	12.0
		6	24.5	0.1	64.6
		12	49.7	3.7	119.0
		24	71.4	5.2	155.1
DGP(6)	AIC	1	–2.0	–4.1	0.4
$R = 9$		6	2.3	–7.6	5.9
		12	13.2	–7.1	15.6
		24	24.2	–2.6	30.5
	BIC	1	–1.4	–4.8	3.1
		6	10.5	–7.3	36.2
		12	25.7	–9.0	65.8
		24	38.8	–6.5	88.7

Note: The entries are values averaged over 1000 replications. The entries are calculated by $\{(\text{RMSFE of a model})/(\text{RMSFE of BREAK}) - 1\} \times 100$.

for multi-step forecasts. The FI-BREAK model shows the smallest amount of cases where the unconditional coverage probabilities lie outside the acceptance region.

To summarize, if one believes that a time series might have long memory or perhaps have structural breaks, one can start with fitting the joint FI-BREAK model. Our simulation results show that the parameter estimates in this model can suggest whether submodels are more appropriate. In terms of forecasting, we find that the joint model leads to good forecasts, even when the DGP is one of the submodels.

AN EMPIRICAL ILLUSTRATION

In this section, we compare the estimation and forecast performance of AR, ARFI, BREAK and FI-BREAK models for volatilities (log of squared returns) of the S&P 500 series from January 4, 1928 to August 30, 1991, amounting to 17,054 daily observations. We split the total sample into

Table V. Root mean squared forecast errors relative to a FI-BREAK model (as a percentage)

DGP		h-Step	Model		
			ARFI	BREAK	AR
DGP(7)	AIC	1	6.8	2.9	9.7
$d = 0.1$, $R = 3$		6	15.3	5.9	22.9
		12	21.5	5.1	30.1
		24	34.4	1.2	46.2
	BIC	1	9.0	3.5	17.1
		6	20.9	3.9	53.4
		12	31.4	3.7	79.4
		24	48.6	−0.5	112.0
DGP(8)	AIC	1	2.4	15.5	4.4
$d = 0.4$, $R = 3$		6	11.3	33.5	11.9
		12	19.3	37.0	14.1
		24	34.1	41.9	19.5
	BIC	1	1.7	5.2	3.7
		6	11.2	22.1	17.7
		12	21.2	31.3	26.6
		24	37.6	41.9	38.6
DGP(9)	AIC	1	2.0	9.2	5.1
$d = 0.4$, $R = 3$		6	7.8	27.2	13.1
		12	15.7	29.2	19.3
		24	33.3	30.8	30.9
	BIC	1	3.5	6.9	7.9
		6	11.2	26.3	29.6
		12	20.4	31.7	51.7
		24	38.6	38.5	86.0

Note: The entries are values averaged over 1000 replications. The entries are calculated by {(RMSFE of a model)/(RMSFE of FI-BREAK) − 1} × 100.

Table VI. Coverage rates of interval forecasts

DGP	Coverage	h-Step	Model			
			AR	ARFI	BREAK	FI-BREAK
DGP(1)	75%	1	0.764	0.770	0.756	0.768
		6	0.749	0.758	0.743	0.754
		12	0.750	0.750	0.724	0.747
		24	0.738	0.740	0.727	0.734
	95%	1	0.932*	0.941	0.927*	0.938
		6	0.947	0.942	0.938	0.944
		12	0.951	0.952	0.946	0.952
		24	0.942	0.940	0.933*	0.937
DGP(2)	75%	1	0.761	0.767	0.738	0.765
		6	0.736	0.750	0.686*	0.734
		12	0.738	0.740	0.681*	0.725
		24	0.724	0.744	0.660*	0.739

Table VI. *continued*

DGP	Coverage	*h*-Step	Model			
			AR	ARFI	BREAK	FI-BREAK
	95%	1	0.937	0.940	0.926*	0.940
		6	0.939	0.938	0.906*	0.931*
		12	0.941	0.949	0.908*	0.948
		24	0.932	0.944	0.899*	0.942
DGP(3)	75%	1	0.750	0.764	0.738	0.768
		6	0.736	0.747	0.645*	0.725
		12	0.723	0.741	0.620*	0.715*
		24	0.668*	0.724	0.554*	0.696*
	95%	1	0.937	0.936	0.921*	0.938
		6	0.931*	0.933*	0.877*	0.923*
		12	0.925*	0.941	0.857*	0.931*
		24	0.883*	0.935*	0.798*	0.918*
DGP(4)	75%	1	0.743	0.745	0.743	0.753
		6	0.760	0.768	0.754	0.769
		12	0.718*	0.724	0.717*	0.717*
		24	0.744	0.744	0.751	0.755
	95%	1	0.950	0.951	0.947	0.950
		6	0.953	0.957	0.955	0.958
		12	0.961	0.966*	0.962	0.961
		24	0.940	0.943	0.939	0.945
DGP(5)	75%	1	0.730	0.745	0.731	0.744
		6	0.745	0.766	0.755	0.769
		12	0.712*	0.711*	0.711*	0.720*
		24	0.719*	0.728	0.731	0.747
	95%	1	0.941	0.941	0.943	0.947
		6	0.951	0.951	0.944	0.955
		12	0.949	0.955	0.945	0.951
		24	0.922*	0.936	0.928*	0.930*
DGP(6)	75%	1	0.729	0.741	0.726	0.750
		6	0.752	0.766	0.729	0.753
		12	0.719*	0.726	0.684*	0.720*
		24	0.704*	0.704*	0.659*	0.700*
	95%	1	0.941	0.948	0.934*	0.948
		6	0.940	0.944	0.927*	0.944
		12	0.927*	0.938	0.920*	0.939
		24	0.920*	0.924*	0.892*	0.910*
DGP(7)	75%	1	0.734	0.745	0.718*	0.754
		6	0.720*	0.739	0.700*	0.732
		12	0.700*	0.727	0.666*	0.703*
		24	0.662*	0.693*	0.682*	0.715*
	95%	1	0.934*	0.940	0.935*	0.944
		6	0.929*	0.941	0.929*	0.940
		12	0.930*	0.930*	0.904*	0.925*
		24	0.902*	0.917*	0.916*	0.921*

Table VI. *continued*

DGP	Coverage	h-Step	Model			
			AR	ARFI	BREAK	FI-BREAK
DGP(8)	75%	1	0.748	0.772	0.742	0.760
		6	0.756	0.774	0.701*	0.762
		12	0.746	0.755	0.684*	0.743
		24	0.710*	0.704*	0.585*	0.710*
	95%	1	0.947	0.949	0.934*	0.941
		6	0.940	0.944	0.891*	0.942
		12	0.949	0.952	0.871*	0.913*
		24	0.911*	0.927*	0.787*	0.883*
DGP(9)	75%	1	0.741	0.767	0.724	0.761
		6	0.755	0.778*	0.657*	0.758
		12	0.734	0.751	0.617*	0.713*
		24	0.680*	0.720*	0.550*	0.709*
	95%	1	0.949	0.951	0.935*	0.948
		6	0.940	0.936	0.888*	0.935*
		12	0.943	0.952	0.847*	0.928*
		24	0.908*	0.932*	0.808*	0.906*

Note: The entries are the empirical coverage probabilities of the 75% and 95% theoretical forecast intervals. We use 1000 independent forecasts for calculating the coverage probabilities. Asterisks indicate cases where the unconditional coverage probabilities lie outside the 95% acceptance region.

Table VII. In-sample comparison of various models

Period	ARFI		BREAK		FI-BREAK		
	d	Log-likelihood	R	Log-likelihood	d	R	Log-likelihood
1 1928–34	0.278	−2094.8	3	−2084.6	0.115	14.0	−2102.4
2 1934–40	0.169	−2050.3	2	−2038.8	0.001	11.3	−2055.9
3 1941–47	0.123	−2056.4	4	−2026.3	0.060	17.8	−2054.2
4 1947–53	0.089	−2102.6	3	−2083.6	0.057	6.4	−2102.0
5 1954–60	0.069	−2024.6	2	−2014.3	0.051	3.2	−2025.1
6 1960–66	0.061	−2135.8	1	−2132.0	0.061	0.0	−2135.8
7 1967–73	0.159	−2149.4	3	−2132.4	0.140	1.6	−2147.9
8 1973–79	0.334	−2147.6	2	−2143.4	0.067	3.9	−2159.1
9 1980–86	0.059	−2200.1	2	−2187.8	0.000	11.4	−2196.0
10 1986–91	0.133	−2240.8	4	−2213.1	0.000	5.8	−2241.0

Note: The estimation results are based on the first 1000 observations of each subsample. The values of R in FI-BREAK model are calculated by Σq_t. The data are the logs of squared returns of S&P 500 series from January 4, 1928 to August 30, 1991 with 17,054 daily observations. We split the total sample into 10 subperiods.

Table VIII. Root mean squared forecasts errors, relative to the FI-BREAK model

Period	ARFI						BREAK					
	$h = 1$	5	10	20	120	240	$h = 1$	5	10	20	120	240
1	0.998	1.008	1.028	1.044	1.064	1.340	1.011	1.062^{*+}	1.083	1.082	0.957	0.785
2	1.006*	1.020	1.051	1.115	1.178	1.080	1.003	1.014	1.012	1.005	1.025	0.999*
3	1.005	1.024	1.045	1.076	1.479^{*+}	1.515^{*+}	1.013^{+}	1.043*	1.061	1.086	1.300*	1.245^{+}
4	1.000	1.001	0.998	0.985	0.871	0.971	1.008	1.047	1.084	1.115	1.060	0.756
5	0.999*	0.997	0.995	0.998	1.103	1.037	1.010	1.043	1.087	1.165	1.566^{*+}	1.674^{+}
6	0.999	0.999	1.008	1.000	0.909	0.681^{*+}	1.017^{+}	1.006	1.005	1.020	1.139	1.220^{+}
7	0.998*	0.982	0.971	0.942	0.867^{*+}	0.940	1.012	1.056	1.087*	1.106	1.133	1.023*
8	1.008	1.024	1.044	1.078	1.583	1.520*	1.012*	1.072^{+}	1.115^{+}	1.167	1.392^{*+}	1.400^{*+}
9	1.008	1.030	1.061	1.098	0.917	0.687^{*+}	1.006	1.036	1.082	1.140	1.145*	1.105
10	1.004	1.006	1.005	0.990	0.841	0.739	1.007	1.039^{*+}	1.083^{*+}	1.143^{*+}	1.266	1.284^{*+}

Note: The entries are calculated by (RMSFE of a model)/(RMSFE of FI-BREAK). * (+) denotes that the null of equal forecast accuracy is rejected (5% level) by the Diebold and Mariano test when the loss function has absolute (squared) loss.

Table IX. Out-of-sample forecastability: coverage rates of 75% interval forecasts

Period	AR						ARFI					
	$h = 1$	5	10	20	120	240	$h = 1$	5	10	20	120	240
1	72.7	73.5	72.2	69.0*	65.3*	64.3*	75.1	76.0	75.2	75.2	68.6*	65.7*
2	68.6*	68.1*	67.6*	67.4*	64.2*	65.3*	70.3*	71.2	69.9*	69.9*	63.4*	65.6*
3	78.5	79.7*	79.7*	80.1*	80.3*	82.7*	76.2	78.3	78.7	79.2	80.3*	82.7*
4	74.1	74.0	74.5	74.7	73.1	75.6	76.4	75.5	75.1	75.4	73.7	74.6
5	69.7*	70.0*	69.6*	69.6*	68.6*	68.5*	70.2*	70.9*	71.2*	70.4*	68.1*	68.2*
6	74.6	74.6	74.6	74.2	72.4*	75.8	76.2	77.4	76.5	75.8	72.3	75.3
7	74.3*	75.9	75.2	75.6	74.4	70.3*	75.8	76.5	75.9	76.6	74.1	70.3*
8	79.0*	80.5*	81.7*	81.4*	81.2*	78.8*	80.8*	81.5*	80.7*	81.4*	80.6*	78.4
9	77.8	78.1	77.8	77.9	76.7	79.6	78.3	78.8	78.7	78.4	76.7	79.8
10	83.0	83.0	82.9	82.6	82.3	81.6*	81.5	83.1*	82.9*	82.7	82.0	80.7*
Period	BREAK						FI-BREAK					
	$h = 1$	5	10	20	120	240	$h = 1$	5	10	20	120	240
1	74.8	74.7	74.8	74.5	71.2	69.2*	75.0	76.0	75.4	75.3	73.0	72.7
2	70.1*	70.3*	71.1	71.1	66.7*	66.0*	71.2	70.4*	71.1*	71.0	67.1*	66.3*
3	74.7	75.2	75.4	75.8	75.3	79.5	74.5	75.6	75.5	76.0	77.5	81.3*
4	74.5	73.6	74.5	74.7	74.3	75.3	77.4	76.8	77.5	77.0	75.0	76.1
5	70.3*	70.8*	69.2*	68.1*	69.6*	69.2*	72.0*	72.1	70.9*	70.7*	70.5	69.9
6	75.5	75.5	74.0	73.1	69.4*	68.6*	77.5	77.7	76.3	75.4	70.8*	71.1
7	78.1	75.9	76.2	75.9	73.4	68.3*	75.8	76.4	75.8	75.9	75.5	69.8*
8	79.5*	78.8	79.5*	80.1*	78.9	78.4	80.6*	80.2*	79.9*	79.3*	80.1*	78.1
9	77.8	78.4	77.4	78.4	75.3	76.7	78.7	79.3	79.3	79.3	75.7	77.8
10	81.3	81.1	80.8	79.9	78.4	76.4	81.6	81.6*	81.4*	81.5*	81.4	78.1

Note: The entries are the empirical coverage probabilities of the 75% theoretical forecast interval. Asterisks indicate models with conditional coverage probabilities outside the 95% acceptance region.

10 subperiods. We remove all zero returns which are due to holidays, and this leads to different sample sizes for each subsample.

For the in-sample period, we estimate the parameters with AR lags selected by BIC. We fix the AR order p of the FI-BREAK model to be the same as that of the BREAK model.

Table VII reports on parameter estimates, that is, the d in the ARFI and FI-BREAK models, and the R in the BREAK and FI-BREAK models. We estimate parameters using the first 1000 observations of each subsample. The values of R in the FI-BREAK model are calculated by Σq_t. The numbers of breaks in the BREAK model appear to vary from 1 to 4. The values of d in the ARFI model range from 0.059 in the period 1986–1991 up to a value of 0.334 in the period 1973–1979. However, the values of d in the FI-BREAK model range from 0.000 to 0.140. Hence, by including a break component, the long memory parameter decreases rapidly.

The root mean squared forecast errors are recorded in Table VIII. Forecast horizons include 20, 60, 120 and 240 trading days, which approximately match with 1 month, 1 quarter, 6 months and 1 year. The Diebold and Mariano (1995) statistic has p-values which show strong evidence that the FI-BREAK model is superior to the BREAK model, but we obtain mixed results when we compare the ARFI and FI-BREAK models. Table IX reports on the empirical conditional coverage probabilities of the three models, as well as of the linear AR model. The conditional coverage is approximately correct for the short (1-, 5-, 10-, 20-step) horizons for all models. The coverage rates for the FI-BREAK model are more often correct for faraway horizons.

CONCLUSION

In this paper we introduced a model that can jointly capture structural breaks and long memory. We examined our FI-BREAK model to see if S&P 500 volatility has long memory, level shifts or both. We found that level shifts are dominant, but that long memory did not always disappear. We used a range of forecast evaluation techniques to investigate the relative performance of the FI-BREAK model to a BREAK model or an I(d) model, and found that the new joint model would be preferable for these data.

The simulation results in this paper indicated that the long memory parameter in the FI-BREAK model can be consistently estimated, also in the case that there is an unknown break component. Also, our simulations indicated that the forecast performance of the FI-BREAK model is better than for the closest competitive models.

ACKNOWLEDGEMENTS

The research of Hyung was supported by a research grant from the University of Seoul (2002).

REFERENCES

Bai JS, Perron P. 1998. Estimating and testing linear models with multiple structural changes. *Econometrica* **66**: 47–78.

Beran J. 1995. Maximum likelihood estimation of the differencing parameter for invertible short and long memory autoregressive integrated moving average models. *Journal of the Royal Statistical Society, Series B* **57**: 659–672.

Bos C, Franses PH, Ooms M. 1999. Long memory and level shifts: re-analyzing inflation rates. *Empirical Economics* **24**: 427–449.
Christoffersen PE. 1998. Evaluating interval forecasts. *International Economic Review* **39**: 841–862.
Crato N, Ray BK. 1996. Model selection and forecasting for long-range dependent processes. *Journal of Forecasting* **15**: 107–125.
Diebold FX, Inoue A. 2001. Long memory and regime switching. *Journal of Econometrics* **105**: 131–159.
Diebold FX, Mariano RS. 1995. Comparing predictive accuracy. *Journal of Business and Economic Statistics* **13**: 253–263.
Engle RF, Smith AD. 1999. Stochastic permanent breaks. *Review of Economics and Statistics* **81**: 553–574.
Granger CWJ, Hyung N. 2001. Occasional structural breaks and long memory. UCSD Working Paper.
Smith AD. 2000. Forecasting inflation in the presence of structural breaks. UC Davis Working Paper.
van Dijk DJC, Franses PH, Paap R. 2002. A nonlinear long memory model for US unemployment. *Journal of Econometrics* **110**: 135–165.

Authors' biographies:
Namwon Hyung holds a PhD degree from the University of California in San Diego. At present he is an Assistant Professor at the Department of Economics of the University of Seoul. His research interests are time series analysis and forecasting.

Philip Hans Franses is Professor of Applied Econometrics and Professor of Marketing Research, both at the Erasmus University Rotterdam. Currently, he is the director of the Econometric Institute. His research interests include marketing research models and techniques.

Authors' addresses:
Namwon Hyung, Department of Economics, University of Seoul, Seoul, South Korea.

Philip Hans Franses, Econometric Institute, Erasmus University Rotterdam, NL-3000 DR Rotterdam, The Netherlands.

