On the forecasting ability of ARFIMA models when infrequent breaks occur

VASCO J. GABRIEL* AND LUIS F. MARTINS†

*Department of Economics, University of Minho, Portugal
and
University of Surrey, UK
E-mail: vjgabriel@eeg.uminho.pt

†Department of Economics, Penn State University, USA and Department of Quantitative Methods, ISCTE, Portugal

Summary Recent research has focused on the links between long memory and structural breaks, stressing the memory properties that may arise in models with parameter changes. In this paper, we question the implications of this result for forecasting. We contribute to this research by comparing the forecasting abilities of long memory and Markov switching models. Two approaches are employed: the Monte Carlo study and an empirical comparison, using the quarterly Consumer Price inflation rate in Portugal in the period 1968–1998. Although long memory models may capture some in-sample features of the data, we find that their forecasting performance is relatively poor when shifts occur in the series, compared to simple linear and Markov switching models.

Keywords: *Long Memory*, *Regime switching*, *Forecasting*.

1. INTRODUCTION

There has been a considerable interest in long memory and structural change in time series, as witnessed by the remarkable growth of the theoretical and empirical research on these issues over the last years. However, only recently have econometricians begun to consider the relationships between the two seemingly distinct phenomena. Indeed, Granger and Teräsvirta (1999), Granger and Hyung (1999), Diebold and Inoue (2001) or Gourieroux and Jasiak (2001) show analytically and via Monte Carlo that models with regime changes may exhibit long memory properties.

What are the implications of these results for forecasting? Despite 'spurious long memory' effects due to regime shifts, will an ARFIMA specification still be an effective tool for forecasting? Diebold and Inoue (2001, p. 157) suggest that '[e]ven if the "truth" is structural change, long memory may be a very convenient shorthand description, which may remain very useful for tasks such as prediction'. Thus, we investigate whether a long memory approach will be 'robust' to structural breaks in a time series, in terms of providing good forecasts for financial and macroeconomic data. Judging by the way predictions are constructed for long memory models (i.e. taking into account the information of distant lags), one may anticipate that ARFIMA models would experience difficulties in forecasting future immediate regime changes, unless the switching is transitory.

The question of the relative forecast performance of long memory and structural change models has not (to our knowledge) been addressed yet. Near-observational equivalence does not necessarily imply similar forecasting abilities. This problem may also be seen as a variant of the issue of whether to use forecasts from trend-stationary or difference-stationary models, which arises from the fact that these models are practically indistinguishable in small samples. See Clements and Hendry (2000) for a recent discussion on this issue.

We compare the univariate forecast accuracy of one type of regime switching model, the Markov Switching (MS) model, with that of fractionally integrated ARMA (ARFIMA) models. Other models with parameter shifts could have been considered (e.g. TAR, STAR or STOPBREAK models), but we stress the MS specification, since it is a widely used approach to model changes in time series. Our analysis is conducted by means of Monte Carlo simulations and empirically, by investigating the ability of the two methods to forecast the inflation rate in Portugal. It is interesting to use inflation rates for this comparison, since we may find different means and variances for different periods in these series, but we also may use long memory models to account for their persistence.

Concerning the Monte Carlo experiment, in the first set of simulations we use the empirical estimates as parameter values for the data generating process (DGP). Subsequently, we refine the experiment by extending the simulations in Clements and Hendry (1998), including long memory and MS models and evaluating their forecast accuracy under different DGPs. Obviously, by focusing on univariate methods we are simplifying our analysis, mainly for expositional simplicity. Nevertheless, this may be viewed as a first approximation to more evolved forecasting practices, since univariate forecasts are usually taken as benchmarks for later comparisons. See Stock and Watson (1999) for a recent discussion on inflation forecasting.

In a related study, Clements and Krolzig (1998) claim that, although non-linear models (including the MS model) may be superior in capturing some features of the data, their forecast performance is not superior to more simple linear time series models. Moreover, Clements and Hendry (1998) argue that some types of linear models may be robust to structural breaks, in terms of their ability to circumvent forecast failure. These authors compared the prediction accuracy of several linear models when the DGP produced a single change in the mean.

Notwithstanding this, none of these works considered the more general linear ARFIMA model. Given the potential confusion between long memory and regime shifts that may arise in many empirical situations, it is of obvious interest to assess how long memory models behave in terms of forecasting when time series suffer regime shifts. Therefore, our paper may be viewed as the implementation of the ideas in Diebold and Inoue (2001), *inter alia*, to forecasting problems and as a complement to the studies of Clements and Hendry (1998) and Clements and Krolzig (1998).

The paper proceeds as follows. In Section 2, we briefly review modelling and forecasting with ARFIMA and Markov switching models, introducing definitions and notation, and consider why parameter shifts may cause the appearance of long memory characteristics in a given time series. Section 3 discusses empirical aspects of our example, including a forecasting exercise, complemented by Monte Carlo analysis. The next section presents the results of further Monte Carlo simulations. Finally, Section 5 provides some discussion and conclusions.

2. LONG MEMORY AND REGIME SWITCHING MODELS

2.1. Fractional ARIMA models

Long memory in time series econometrics has been the subject of many studies, and recent surveys of the literature may be found in Baillie (1996). Fractional integration, as in Granger and Joyeux (1980), for example, aims to circumvent some of the limitations of integer analysis of ARIMA models. A fractionally integrated ARMA process y_t may be represented by

$$\Phi(L)(1-L)^d y_t = \Theta(L)\varepsilon_t, \qquad \varepsilon_t \sim \text{i.i.d.}(0, \sigma^2), \tag{1}$$

where d is a parameter that assumes a non-integer value in the difference operator, $(1-L)^d$. The fractional differencing operator is defined by the binomial expansion

$$(1-L)^d = \sum_{i=0}^{\infty} \binom{d}{i} (-L)^i, \tag{2}$$

or $(1-L)^d = 1 - dL + d(d-1)/2!L^2 - d(d-1)(d-2)/3!L^3 + \cdots$, for $d > -1$. The process is stationary and invertible if the roots of the autoregressive polynomial of order p, $\Phi\ (L) = 1 - \phi_1 L - \cdots - \phi_p L^p$, and of the moving-average part of order q, $\Theta(L) = 1 + \theta_1 L + \cdots + \theta_q L^q$, lie outside the unit circle, with $|d| < 0.5$. Obviously, the ARFIMA model generalizes the traditional ARIMA representation with integer values for d.

Long memory is usually defined in the time domain, characterized by a hyperbolically decaying autocorrelation function, with $\rho_y(k) = ak^{2d-1}$ as $k \to \infty$, or alternatively, in the frequency domain, where in the lowest frequencies the spectrum is $f_y(\omega) \sim c\omega^{-2d}$, when $\omega \to 0$. It is also noted that a process is $I(d)$ (for $d > 0$) if the variance of the partial sum process $S_T = \sum_{t=1}^{T} y_t$ is of order $O(T^{2d+1})$ as $T \to \infty$. The process y_t exhibits long memory for $d \in (0, 1)$, being covariance-stationary if $d < 0.5$ and still mean-reverting if $d < 1$. This contrasts with stationary, $I(0)$, ARMA, or 'short memory', processes, where dependence tends to be dissipated geometrically with time, meaning that shocks have a temporary effect in the process. In its turn, $I(1)$ processes are not mean-reverting, wherefore shocks have permanent effects. Fractional ARMA models are, thus, an intermediate and flexible form of analysing time series.

Several methods have been proposed to estimate the parameter d and the remaining parameters of the ARFIMA specification, either in the time or in the frequency domain. See Geweke and Porter-Hudak (1983, hereafter GPH), Fox and Taqqu (1986) and Sowell (1992), among others, and Baillie (1996) for comparisons and discussion of small sample properties.

Concerning prediction from ARFIMA processes, this is usually carried out by using an infinite autoregressive representation of (1), written as $\Pi(L)y_t = \varepsilon_t$, or

$$y_t = \sum_{j=1}^{\infty} \pi_j y_{t-j} + \varepsilon_t, \tag{3}$$

where $\prod\ (L) = (1 - \pi_1 L - \pi_2 L^2 - \cdots) = \Phi\ (L)(1-L)^d \Theta(L)^{-1}$. In terms of practical implementation, this form needs truncation after k lags, but there is no obvious way of doing it. This truncation problem will also be related to the forecast horizon considered in predictions (see Crato and Ray 1996). From (3) it is clear that the forecasting rule will pick up the influence of distant lags, thus capturing their persistent influence. However, if a shift in the process occurs, this means that pre-shift lags will also have some weight on the prediction, which may cause some biases for post-shift horizons.

2.2. *Markov switching models*

The importance of non-linearities (along with structural changes) in economic series has often been debated in the literature. The discussion was further intensified since Hamilton (1989) proposed his autoregressive Markov switching model to analyse US GNP growth rate. It offers a powerful and flexible instrument to characterize macroeconomic fluctuations, by accommodating asymmetries and changes in the behaviour of economic time series. Several extensions and generalizations have been presented, see Kim and Nelson (1999), *inter alia*, for a survey.

Consider, for simplicity, the first-order autoregressive Markov switching model with two regimes, MS(2)-AR(1),

$$y_t - \mu(s_t) = \phi[y_{t-1} - \mu(s_{t-1})] + \sigma(s_t)\varepsilon_t, \tag{4}$$

where $\varepsilon_t \sim$ n.i.d.(0, 1). Here, s_t is a binary random variable on $S = \{1, 2\}$, indicating the unobserved regime or state driving the process at date t. To complete the specification of the model, it is postulated that $\{s_t\}$ is a stationary first-order Markov chain in S with transition matrix $P = (p_{ij})$, where

$$p_{ij} = \Pr(s_t = j|s_{t-1} = i), \qquad i, j \in S. \tag{5}$$

Furthermore, it is assumed that $\{s_t\}$ is independent of $\{\varepsilon_t\}$. Therefore, the mean $\mu(s_t)$ and the variance $\sigma^2(s_t)$ of the innovation ε_t switch between two states according to an unobserved Markov chain. It is also possible to consider a more general specification, where the dynamic components, namely the autoregressive coefficients, are allowed to depend on s_t.

Estimation of the parameters of the model, $\theta = \{\mu(s_t), \sigma^2(s_t), \phi, p_{ij}\}$, is carried out by maximizing the likelihood function of the MS-AR model. It involves recursive computation of probabilities about the unobserved regimes and obtaining $\hat{\theta}$ that maximizes the log-likelihood function. This may be achieved through numerical optimization or using the EM procedure (see Hamilton 1994; Kim and Nelson 1999).

In terms of forecasting, the MS specification allows us to obtain forecasts in an easy fashion. To construct forecasts for the regime probabilities conditional on past values of y_t (Y_t), consider now the general case of an N-state Markov chain and let P denote the matrix of transition probabilities for the N states and let

$$\hat{\lambda}_t' = [p(s_t = 1|Y_t)\ p(s_t = 2|Y_t) \cdots p(s_t = N|Y_t)] \tag{6}$$

be the vector containing the inference about the current state (the filtered probabilities). The optimal h-step-ahead of prediction for the probabilities of the unobserved state conditional on information available at date t is given by $\hat{\lambda}_{t+h|t}' = \hat{\lambda}_t' P^h$ or,

$$\Pr(s_{t+h} = j|Y_t) = \sum_{i=1}^{N} \Pr(s_{t+h} = j|s_t = i)\Pr(s_t = i|Y_t). \tag{7}$$

On the other hand, to construct forecasts for the observed series $\{y_t\}$, we calculate the conditional expectation $E(y_{t+h|t})$ as

$$E(y_{t+h|t}) = \sum_{j=1}^{N} \Pr(s_{t+h} = j|Y_t)E(y_{t+h}|Y_t, s_{t+h} = j), \tag{8}$$

meaning that the forecast for each regime is multiplied by the corresponding probability that the process will be in that regime and the sum of these products will form the forecast for y_{t+h}. For the simple MS(2)-AR(1) model in (4), $E(y_{t+h}|Y_t, s_{t+h} = j) = \mu(s_{t+h}) + \phi[y_{t+h-1} - \mu(s_{t+h-1})]$, so we have

$$\hat{y}_{t+h|t} = \hat{\mu}(s_{t+h|t}) + \phi[\hat{y}_{t+h-1|t} - \hat{\mu}(s_{t+h-1|t})], \tag{9}$$

where $\hat{\mu}(s_{t+h|t}) = \sum_{j=1}^{2} \hat{\mu}_j \Pr(s_{t+h} = j|Y_t)$. However, as the regimes become unpredictable (implying that $\Pr(s_t|s_{t-1}) = \Pr(s_t)$), the forecasting rule will become linear, since then $\hat{\mu}(s_{t+h|t}) = \tilde{\mu}$, the unconditional mean of y_t, and thus,

$$\hat{y}_{t+h|t} = \tilde{\mu}(1 - \phi^h) + \phi^h y_t, \tag{10}$$

which means that forecasts will be, in essence, similar to those of linear models (see Clements and Krolzig 1998). Of course, this recursion could be easily extended to more complicated models (see Hamilton 1994; Clements and Krolzig 1998).

2.3. Long memory in markov switching models

As mentioned in the Introduction, some recent papers deal with the relationship between long memory and regime shifts, namely stochastic regime switching. These authors analysed several cases with stochastic parameter shifts, by looking at the behaviour of the autocorrelations of the processes (or by deducing the rate of growth of the variance of partial sums of the processes), showing that they may be described asymptotically as an $I(d)$ process. The key idea behind this result is the following: as the frequency of regime switching decreases (i.e. as p_{11} and p_{22} approach unity in the Markov switching case), the process will closely resemble a fractionally integrated series. Moreover, the size of the parameter shifts will also be a factor to take into account because larger magnitudes of breaks will introduce more persistence in the series.

This can be easily verified by considering an example with the simple two-regimes first-order autoregressive Markov switching model in (4). The corresponding population autocorrelation function at lag k is given by

$$\rho_k = \frac{\pi_1\pi_2(\mu_1 - \mu_2)^2 vec(P^k)v_1 + \phi^k\pi'(I_k - \phi^2 B)^{-1}\sigma_s}{(\pi_1\mu_1 + \pi_2\mu_2)^2 + \left(\pi_1\sigma_1^2 + \pi_2\sigma_2^2\right)(1 + \phi^2)^{-1}}, \tag{11}$$

where π_j represents the ergodic probability of staying in regime j $(j = 1, 2)$, $\pi = [\pi_1, \pi_2]$ $v_1 = [\pi_2, -\pi_2, \pi_1, -\pi_1]$, $\sigma_s = [\sigma_1^2, \sigma_2^2]$, μ_j and σ_j^2 are the state dependent means and variances, I_k is a k-dimensional identity matrix and B is the matrix of transition probabilities for the 'time reversed' Markov chain (see, Timmermann 2000, Propositions 2 and 4). Setting $\sigma_1^2 = \sigma_2^2 = 1$, $p_{11} = p_{22} \in \{0.95, 0.98, 0.99\}$, $\mu_1 = 1$ and considering distinct values for μ_2 (i.e., different magnitudes of shifts) and ϕ, we calculated the autocorrelation function up to $k = 50$. From the results presented in Table 1, it is possible to observe the following. Firstly, there is a positive relationship between the persistence of the process and the transition probabilities, as well as between the size of the shift

Table 1. Autocorrelation function of an autoregressive Markov switching-mean model.

	$p_{11} = p_{22} = 0.95$								
	$\phi = 0.2$			$\phi = 0.5$			$\phi = 0.9$		
μ_2	2	5	10	2	5	10	2	5	10
$k = 1$	0.336	0.755	0.866	0.563	0.80	0.875	0.90	0.90	0.90
10	0.068	0.277	0.332	0.056	0.262	0.327	0.349	0.349	0.349
20	0.235	0.097	0.116	0.019	0.091	0.114	0.122	0.122	0.122
50	0.001	0.004	0.005	0.001	0.004	0.005	0.005	0.005	0.005
	$p_{11} = p_{22} = 0.98$								
	$\phi = 0.2$			$\phi = 0.5$			$\phi = 0.9$		
μ_2	2	5	10	2	5	10	2	5	10
$k = 1$	0.347	0.803	0.923	0.573	0.845	0.932	0.903	0.926	0.947
10	0.129	0.528	0.632	0.106	0.499	0.624	0.363	0.485	0.60
20	0.086	0.351	0.42	0.07	0.332	0.415	0.136	0.259	0.376
50	0.025	0.103	0.124	0.021	0.097	0.122	0.011	0.059	0.104
	$p_{11} = p_{22} = 0.99$								
	$\phi = 0.2$			$\phi = 0.5$			$\phi = 0.9$		
μ_2	2	5	10	2	5	10	2	5	10
$k = 1$	0.351	0.819	0.942	0.576	0.86	0.95	0.904	0.935	0.963
10	0.158	0.648	0.777	0.13	0.613	0.767	0.37	0.551	0.72
20	0.129	0.53	0.635	0.105	0.501	0.626	0.146	0.357	0.555
50	0.07	0.289	0.346	0.058	0.273	0.342	0.021	0.16	0.29

and the autoregressive parameter. Secondly, the rate of decay of the autocorrelations slows down as the latter parameters increase.[1] Even after 50 lags, the autocorrelations are non-negligible. This means that a stationary $I(0)$ process as this Markov switching-mean model generates substantial persistence and, in certain cases (such as those of large permanent changes), may be easily confused with a random walk (see, Timmermann 2000 (Section 6); Nunes *et al.* 1997.

Furthermore, accounting for shifts in the process has the effect of reducing the estimated fractional integration parameter, $\hat{d}$, according to Bos *et al.* (1998) and Granger and Hyung (1999), which indicates that spurious long memory may arise due to neglected shifts. However, Granger

[1]Note that, in our example, the autocorrelations do not depend on the break size when the persistence of the regime is equal to the autoregressive parameter (see columns in Table 1 corresponding to $p_{ij} = 0.95$ and $\rho = 0.9$).

and Hyung (1999) argue that a 'spurious break'-type phenomenon (see Nunes *et al.* 1997) may appear when trying to estimate the number of breaks of an $I(d)$ process with no breaks. For instance, using a Schwarz–Bayesian criterion approach to estimate the number of breaks will lead asymptotically to an infinite number of breaks being estimated, except for $d = 0$, where the correct number of breaks (none) is consistently estimated. Therefore, these results seem to point that the issue 'long memory vs. structural breaks' is just an intermediate form of the controversy 'unit roots vs. structural breaks'.

An interesting feature of the way optimal prediction rules are constructed from MS models is that the prediction can be decomposed into linear and non-linear contributions to the forecast. The contribution of the MS structure depends on the magnitude of the regime shifts and on the persistence of the regimes, given by $p_{11} + p_{22} - 1$ (see Clements and Krolzig 1998, pp. 70–71). Thus, for small breaks and less persistent regimes, a forecast from an MS model will be generated in a way that will resemble a linear prediction rule. On the other hand, it is expected that an MS model will perform better when the regimes are more persistent and for larger breaks. Note, however, that these same factors that favour prediction from MS models are central for the result that an MS process will display long memory properties. Hence, this adds relevance to our study, since it is interesting to assess if the empirical similarities between the two models will continue to hold in terms of forecasting.

3. THE INFLATION RATE IN PORTUGAL

3.1. Empirical analysis

We use MS and ARFIMA specifications to model the empirical path of the inflation rate in Portugal, and evaluate their forecast performance in a simple out-of-sample forecast comparison. This is carried out on a data set of seasonally unadjusted quarterly observations of Consumer Price inflation for the period 1968:1–1998:4. The series is constructed by taking first-differences and logs of the CPI. It is evident from Figure 1 that the series displays seasonality and clear changes in the mean and variance. For simplicity, we will abstract from the problems posed by seasonality and work with unadjusted data, thus concentrating on the other features of the data.[2] In this period, several major events in Portugal led to changes in economic policy and substantial fluctuations in the inflation rate: oil shocks, the democratic Revolution with the subsequent loss of its colonies (1974, 1975), two agreements with the International Monetary Fund (1978 and 1983), the entry in the European Economic Community (in 1986) and, later, in the European Monetary System (in 1992), among others.

Indeed, prior knowledge about the economic conditions in distinct periods and observation of the series supports the hypothesis of different regimes. On the other hand, these events led to an increased persistence in the inflation rate in Portugal, when compared to other European countries. In fact, the series shows the typical behaviour of a series with long memory, with a very slow return to a low inflation regime after a large shock, so one may expect a high estimate for the order of integration.

[2] We considered different methods to account for seasonality, but the results of our subsequent analysis did not change qualitatively.

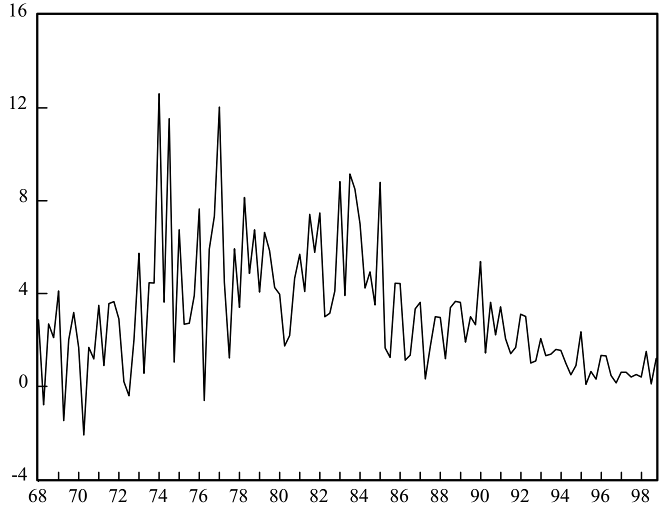

Figure 1. Quarterly CPI inflation rate in Portugal, 1968–1998.

Long memory models have been successfully applied to model inflation rates in several industrialized countries. Hassler and Wolters (1995) found evidence that many inflation rates are neither $I(0)$ nor $I(1)$, having estimated a fractional order of integration of approximately 0.5. Bos *et al.* (1998) consider long memory and level shifts to explain the behaviour of US inflation rate. See, also, Ooms and Doornik (1999) for an application to US and UK inflation rates, including forecasting, as well as Baillie *et al.* (1996).

In turn, MS models are particularly suitable to analyse some of the dynamic features of inflation rates, namely by capturing the apparent changes in mean and variance. Regime shifts in inflation rates have been studied utilizing a variety of specifications with MS. Garcia and Perron (1996) explored the possibility of more than two regimes in the inflation rate process. Evans and Wachtel (1993) and Kim (1994), for example, used richer specifications of the basic MS model to study the link between inflation and uncertainty, accounting for possible changing (conditional) heteroskedasticity of inflation rates. We will not, however, consider these models in our analysis.

In Table 2, we present some tests concerning the properties of the data. Different unit root tests (ADF, Phillips-Perron and DF-GLS as in Elliott *et al.* 1996) and the KPSS stationarity test are computed, and they do not agree on whether there is a unit root in the inflation rate or not. However, both types of tests are known to have their performance affected by the presence of breaks. Furthermore, when testing for structural change using the procedures defined in Andrews (1993), there is clear evidence of breaks in the series.

On the other hand, the estimation of the order of integration d also allows us to test whether the series is $I(0)$ or $I(1)$. We have adopted the frequency domain estimator of Fox and Taqqu (1986) throughout the paper.[3] Looking at the estimates of d and respective standard errors (see Table 3), it can easily be seen that both the $I(0)$ and $I(1)$ hypothesis are rejected.[4] Therefore, it is

[3]We tried different procedures, such as the GPH estimator and the exact maximum likelihood method of Sowell (1992), but the one we adopted seemed to do better in the subsequent forecasting exercises.

[4]The results are for the period 1968:1–1998:4, that is, retaining four observations for prediction. Holding back 16 observations does not change substantially the previous results, so they are not shown.

Table 2. Unit roots, stationarity and structural changes tests for the inflation rate in Portugal.[a]

ADF	−2.222
PP-Z_α	−75.673**
PP-Z_t	−7.618**
KPSS	0.761**
DF-GLS	−2.229
sup-F	422.431**
avg-F	40.295**
exp-F	285.884**

[a]The lag length for the ADF and DF-GLS tests is selected according to a t-test downwards selection procedure, by setting the maximum lag equal to 8 and then testing downwards until a significant last lag is found, at the 5% level. For the Phillips–Perron and KPSS tests, the long run variance is estimated by means of a quadratic spectral kernel with an automatically selected bandwidth estimator
*−5% significant statistic.
**−1% significant statistic.

Table 3. Estimation results for the inflation rate in Portugal (1968:1–1998:4).[a]

	(0, *d*, 0)	(1, *d*, 0)	*MS*(2)	*MS*(2)−*AR*(4)	*MS*(3)−*AR*(2)
d	0.4003 (0.055)	0.466			
ϕ		−0.125 (0.094)			
σ^2	2.565 (0.163)	1.986 (0.152)			
μ_1			1.776 (0.171)	1.335 (0.586)	0.883 (0.126)
μ_2			5.227 (0.42)	4.854 (0.845)	2.33 (0.203)
μ_3					5.251 (0.377)
$\sigma_1{}^2$			2.025 (0.349)	0.582 (0.121)	0.405 (0.121)
$\sigma_2{}^2$			7.96 (1.599)	7.953 (1.465)	2.104 (0.476)
$\sigma_3{}^2$					7.875 (1.625)
p_{11}			0.989 (0.012)	0.97 (0.024)	0.952 (0.04)
p_{22}			0.973 (0.022)	0.973 (0.021)	0.969 (0.024)
p_{33}					0.972 (0.031)

[a]Standard errors in brackets.

difficult to clearly state how the process behaves in the considered sample period. Note that $\hat{d}$ in the ARFIMA (0, d, 0) is less than, but close to, 0.5, which is consistent with the evidence provided in Hassler and Wolters (1995) for the inflation rates of other countries.[5] However, introducing an autoregressive component induces an increase in the estimated d.

Regarding the estimation of MS models, we present in Table 3 results for three distinct specifications: the simple MS model, the widely used MS(2)-AR(4) model and the three-regime model proposed by Garcia and Perron (1996) for the inflation rate. Each model clearly points

[5]The estimates of d range between approximately 0.3 and 0.7 using other estimation methods.

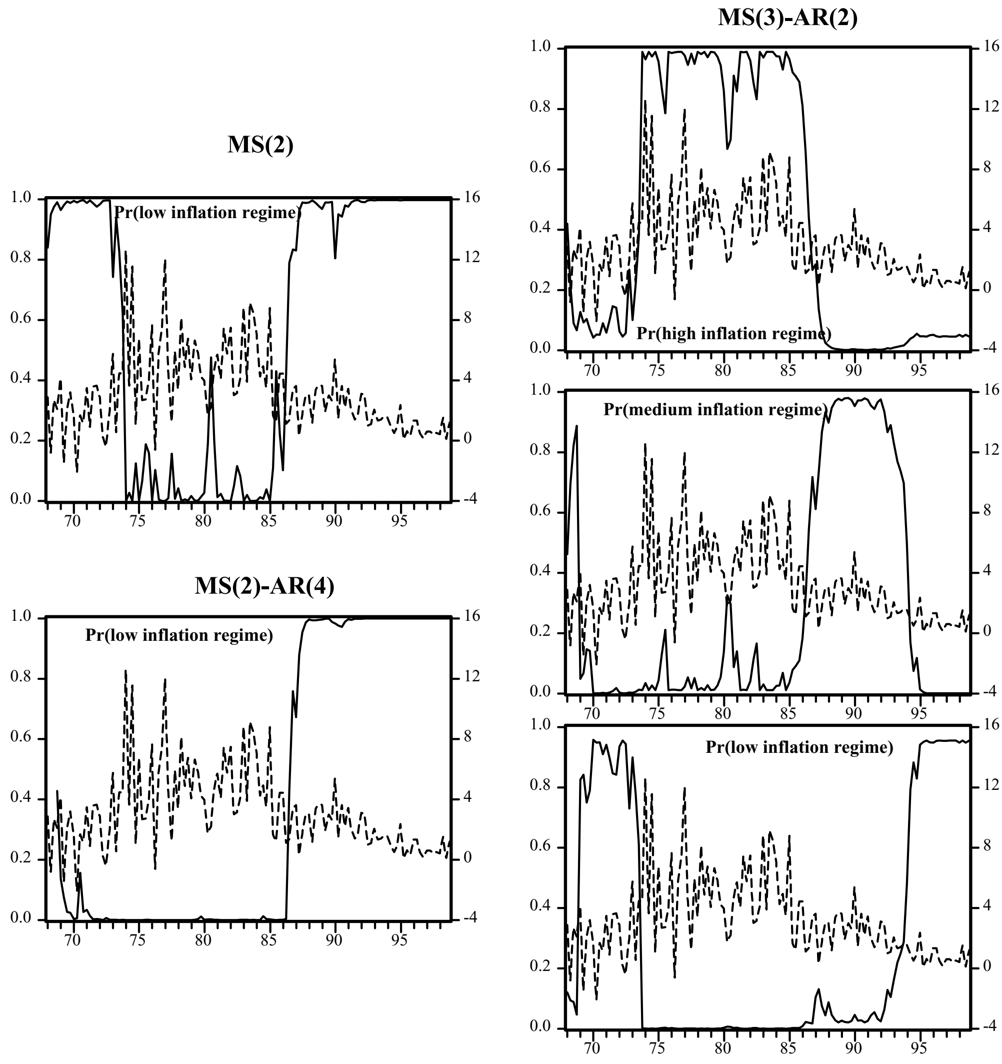

Figure 2. Inflation regimes.

to different means and variances within the sample period.[6] Moreover, the estimated transition probabilities are quite large, indicating that the regimes are very persistent. Therefore, it is not surprising to find evidence of long memory in the series, considering the results in Diebold and Inoue (2001), *inter alia*. Figure 2 displays the regime classification (based on filtered probabilities) for each model. We see that the last period is one of low inflation and that the 2-regime models

[6]One could test the specification of the MS models using the tests proposed in Hansen (1992), for example, but since that is not our main concern, we disregarded that matter.

Table 4. Forecasting performance for the inflation rate in Portugal.

Forecast period	4		16	
Models	FMSE	FMAE	FMSE	FMAE
(0, d, 0)	0.358	0.57	0.398	0.479
(1, d, 0)	0.281	0.483	0.349	0.444
RW	0.725	0.668	0.541	0.577
IMA	0.338	0.529	0.301	0.495
MS(2)-AR(4)	0.242	0.437	0.557	0.636
MS(3)-AR(2)	0.561	0.611	0.248	0.394
MS(2)	1.832	1.227	1.554	1.149

coincide in the dating of the shift (around 1986), whereas the 3-regime model interprets this switch as a change to a medium-inflation period, later followed by a low-inflation one.

Turning to the forecast comparison, we undertake a simple forecasting exercise with a shorter forecasting horizon (4 periods) and a longer one (16 periods), measuring the forecast mean-squared error (FMSE), as well as the forecast mean absolute error (FMAE). For comparison purposes, we consider different types and classes of models. Besides those mentioned above, we also include the random walk (RW) model and an integrated moving-average (IMA) model. The latter model was found to be one of the most robust forecasting devices by Clements and Hendry (1998) in their study. Prediction for the ARFIMAs from (3) was conducted with $k = 10$.

From Table 4, we observe that no single model dominates the others, with the MS(2)-AR(4) predicting better for a 4-period forecast horizon, while the MS(3)-AR(2) does well for 16-steps forecasts. It is interesting to highlight the performance of the ARFIMA (1, d, 0) model, which ranks second for both the shorter and longer horizons. The simplest ARFIMA (0, d, 0) also works well, ranking fourth for each prediction period. Using different lags for the prediction rule of the ARFIMA models did not alter the results substantially, since the π_j's from (3) approach zero very quickly. The good performance of fractional models may be explained by the fact that the last observations, as well as those of the prediction horizon, are relatively stable, that is, it is all taking place in the same ‘regime’. Curiously, the worst model was the simple MS model, perhaps meaning that extra (autoregressive) parameters are needed to account for the dynamics in the series.

In order to circumvent the specificity of these results, in the next section we design a simple Monte Carlo study by taking empirical models of the inflation rate as the DGP. Although an artificial DGP may be useful in this context, it is preferable to use more empirically meaningful estimated models, even if these only offer a poor approximation to the true DGP. This practice also permits controlling for sampling variability of a one-shot type of forecast comparison as in this section, with the empirical example.

3.2. A simple Monte Carlo experiment

For our initial results, we base our DGP on the estimated baseline MS(2) model (third column of Table 3) because it provides a simple, yet rough, description of the data, by estimating changes in mean and variance. Furthermore, we consider a second DGP where we restrict the break points to be those obtained from observing the filtered regime probabilities for the baseline MS model.

Table 5. Monte Carlo FMSE for the empirical MS(2) DGP.

h	(0, d, 0)	(1, d, 0)	RW	IMA	MS(2)
1	10.184	9.416	12.063	7.202	9.011
2	11.682	10.354	12.844	7.524	9.223
3	13.494	11.756	13.418	7.930	9.142
4	14.438	12.519	13.571	8.041	8.907
5	15.637	13.738	14.390	8.144	8.989
6	16.950	14.803	14.182	8.545	8.650
7	17.537	15.244	13.972	8.165	8.596
8	18.634	16.339	15.361	8.687	8.774
9	20.018	17.681	15.879	9.195	8.672
10	20.371	17.995	15.300	8.876	8.818
11	21.183	18.763	15.284	9.359	8.769
12	22.511	20.063	16.315	9.635	9.123
13	22.483	20.146	16.195	9.147	9.263
14	23.264	21.020	16.835	9.821	9.458
15	23.669	21.488	17.000	10.124	9.479
16	25.585	22.733	20.584	14.022	9.788
Average	18.540	16.504	15.178	9.026	9.041

We also consider a smaller value for the variance of the last regime, which is in accordance with what is observed in the series. This DGP is given by $y_t = \mu_t + \sigma_t \varepsilon_t$, with

$$\mu_t = \begin{cases} \mu_1 = 1.8, & \sigma_1^2 = 2, \quad t \le 24 \\ \mu_2 = 5.2, & \sigma_2^2 = 8, \quad 24 < t \le 74 \\ \mu_3 = 1.8, & \sigma_3^2 = 1, \quad t > 74 \end{cases}. \tag{12}$$

While DGP (12) is not truly an MS process (there is no Markov chain behind it), it may be viewed as one with fixed break points. Using GAUSS software, we generate 5000 series of 128 observations, retaining 16 observations for forecasts comparisons.

The results (see Tables 5 and 6) are quite similar, although the forecasts errors are larger in the first case, since the break points are unknown. As expected, for both DGPs the MS model does relatively well, because it is the closest to the specified DGP. The IMA model performs slightly better, which, however, is not surprising, given the results in Clements and Hendry (1998). As for the ARFIMA models, although they provide reasonable forecasts for shorter periods, their performance quickly deteriorates as the forecast horizon increases.

It may be argued that the previous comparison is unfair to ARFIMA models, because these models are trying to fit a different DGP model. Thus, in a second set of simulations, we take the estimated (0, d, 0) and (1, d, 0) ARFIMA models of the inflation rate in Table 3 as the DGP. The corresponding results are shown in Table 7. Surprisingly, we observe that the simple MS model provides better predictions, followed by the (0, d, 0) and IMA models, while the ARFIMA (1, d, 0) is the worst, even when it is the true DGP. Of course, for other plausible DGPs, the results and the ranking could be different. Therefore, in the next section, we refine our empirically based analysis by considering a more complete set of the Monte Carlo experiments.

Table 6. Monte Carlo FMSE for the empirical DGP (12).[a]

h	(0, d, 0)	(1, d, 0)	RW	IMA	$MS(2)$
1	1.455	1.334	2.037	1.130	1.188
2	1.608	1.408	2.033	1.122	1.146
3	1.818	1.532	2.076	1.150	1.156
4	1.956	1.608	2.082	1.129	1.127
5	2.129	1.729	2.114	1.129	1.133
6	2.299	1.857	2.146	1.154	1.139
7	2.410	1.928	2.154	1.120	1.168
8	2.554	2.040	2.131	1.100	1.231
9	2.632	2.108	2.189	1.114	1.185
10	2.879	2.321	2.213	1.139	1.179
11	3.015	2.431	2.236	1.137	1.226
12	3.104	2.511	2.229	1.094	1.227
13	3.204	2.610	2.274	1.106	1.217
14	3.299	2.700	2.246	1.098	1.234
15	3.323	2.741	2.297	1.107	1.295
16	3.555	2.964	2.383	1.140	1.246
Average	2.577	2.114	2.178	1.123	1.194

[a]From the 5000 replications, the following results were obtained for the main parameters. Mean $d = 0.388$ (SE = 0.053); mean $p_{11} = 0.983$ (SE = 0.021); mean $p_{22} = 0.968$ (SE = 0.031).

Table 7. Monte Carlo FMSE for the empirical ARFIMA DGP.[a]

Forecast period	(0, d, 0), $d = 0.4003$		(1, d, 0), $d = 0.466$, $\phi = -0.125$	
Models	FMSE	FMAE	FMSE	FMAE
Models	FMSE	FMAE	FMSE	FMAE
(0, d, 0)	1.264	0.895	1.315	0.914
(1, d, 0)	1.737	1.047	2.634	1.267
RW	1.816	1.075	1.744	1.054
IMA	1.357	0.928	1.296	0.908
$MS(2)$	0.966	0.772	1.075	0.815

[a]The results are averages for $h = 16$.

4. FURTHER MONTE CARLO ANALYSIS

The results from the previous section suggest that ARFIMA models do not perform well when the true model has parameter shifts, while the simple MS model does reasonably well. More surprisingly, forecasts from the latter are superior to those from long memory models, even when the DGP is an ARFIMA. Furthermore, the idea that a linear model as the IMA may be robust to breaks, as explained in Clements and Hendry (1998), finds echo in these simulations. However, as with all Monte Carlo experiments, there is an inevitable specificity concerning the DGPs and the obtained results. In order to compare the relative merits of long memory and MS models in a more general setting, a set of simple Monte Carlo simulations is carried out. We stress what is essential to our case, that is, magnitude and frequency of parameter switching, as discussed in the previous section.

Hence, in the first stage we base our simulations on the DGP studied by Clements and Hendry (1998). These authors compared the prediction accuracy of several linear models with a simple deterministic switching-mean DGP, having concluded that some types of linear models are robust to structural breaks, in terms of forecast failure. Thus, we extend their study by analysing how long memory models behave when time series suffer regime shifts. This is potentially interesting given the near-observational equivalence between long memory and parameter shifts. We begin by considering the simple switching-mean process

$$y_t = \mu_t + \varepsilon_t, \qquad t = 1, \ldots, T \tag{13}$$

where we assume that $\varepsilon_t \sim$ n.i.d.(0, 1) and μ_t evolves as

$$\mu_t = \begin{cases} \mu_1, & t \leq \tau \\ \mu_2, & t > \tau \end{cases}, \tag{14}$$

where τ is an exogenously fixed break point. In our experiments, μ_1 is always 1, while we allow μ_2 to take on different values, in this case $\mu_2 \in \{2, 5, 10\}$. The case $\mu_2 = 10$ corresponds to the DGP analysed in Clements and Hendry (1998), but we also wish to consider other empirically relevant shift magnitudes. Obviously, pronounced breaks as the latter may be detected even by visual inspection, and we may expect fractional models to perform worse. However, the other cases are empirically plausible, as we have seen in the empirical example. For simplicity, the variance is kept constant and we let $\tau = T/2$, generating $T = 100$ plus $h = 16$ random observations in each replication, where the last h observations are held back for the forecast simulation.

We also specify a Markov switching DGP

$$y_t = \mu(s_t) + \sigma(s_t)\varepsilon_t, \tag{15}$$

where μ depends on a stationary first-order Markov chain $\{s_t\}$, independent of $\{\varepsilon_t\}$. The values for μ_2 are taken from $\{2, 5\}$, and in our simulations, the values of the transition probabilities are taken from $(p_{11}, p_{22}) \in \{(0.95, 0.95), (0.99, 0.99)\}$. We should stress that this is the type of parameter setting under which long memory and MS specifications are likely to be confused, as explained before. For this specific DGP, we consider a sample size of 200 observations, given the persistence in the regimes we are considering, and restrict the variances to be equal in the two regimes.

As in the previous section, we also simulate data from a long memory DGP, although in a more general setting. In fact, we simulate data from ARFIMA (0, d, 0) and (1, d, 0) models for a range of values of ϕ and d in the region of stationarity and mean-reversion, $\{-0.75, -0.25, 0.25, 0.75\}$ and $\{-0.49, -0.25, 0.25, 0.49, 0.75\}$, respectively. We tried different values for the parameters, but the results were not significantly distinct, and these values illustrate the question we are addressing.

Finally, another interesting situation that merits attention is when structural change occurs in the forecasting period. It may be of interest to see how different models are robust in terms of 'adapting' their forecasts to a change outside the sample period, especially if we wish to assess '*ex ante*' forecast accuracy. In the previous cases, the models were estimated with the information about the first break. In this case, it may be that, although the forecasts are constructed without the information about the second break, some models may still be robust to the second shift. Thus, we modify the previous DGP by assuming that

$$\mu_t = \begin{cases} \mu_1, & t \le \tau \\ \mu_2, & \tau < t \le T + h/2, \\ \mu_3, & t > T + h/2 \end{cases} \tag{16}$$

which introduces a second break in the middle of the forecasting period. We focus on the empirically more plausible values for μ_2, i.e. (2, 5). When $\mu_2 = 2$, we let $\mu_3 = (1, 3)$, and when $\mu_2 = 5$, μ_3 is allowed to take the values (1, 9).

In all experiments, the number of replications was 5000 and the criterion used for comparisons is the forecast mean-squared error. In each replication, we fit a simple Markov switching-mean model, ARFIMA (0, d, 0) and ARFIMA (1, d, 0) models, a RW and an IMA model, and compute the respective forecasts. We tried different specifications for the ARFIMA models, but in general the ones considered here worked better in terms of forecasting.

Tables 8–10 show the results of the simulations for the four DGPs under study. Considering the results in Table 8 for DGP (14), we observe that the ARFIMA specifications are not, in general, robust predictors. Although their ability to forecast for shorter periods is reasonable, it rapidly deteriorates, a result that was already seen in Section 3.2. Even when the break is relatively small ($\mu_2 = 2$), long memory models offer disappointing forecasts. This contrasts with the results for the IMA model, in that they do not depend on h. In turn, the MS approach is generally superior to the ARFIMAs, and occasionally better than the IMA, especially for shorter forecasting periods. Moreover, for larger shift magnitudes, one gets higher estimates for d, as predicted in Section 2.3. That also leads to a decrease in the predictive ability for all models, with the exception of the ARFIMA (1, d, 0).

The above comments also apply to the MS DGP (Table 9), except that in this case, forecasts from the IMA model also suffer when h increases. Furthermore, we observe that less frequent switching improves the performance of all models. Curiously, the average $\hat{d}$ decreases slightly in this situation (see notes on Table 3), although the estimates are not significantly different for $p_{ij} = 0.95$ and $p_{ij} = 0.99 (i = j)$. Thus, the IMA model is still the best, while the ARFIMAs improve their relative performance in this DGP. The MS becomes relatively more inaccurate when the shift is larger, which is in contradiction to what might be expected (recall Section 2.3).

Considering the results for the long memory DGPs in Table 10, to some extent these confirm the conclusions of Section 3.2. Overall, the ARFIMA (0, d, 0) is the most well-balanced model, but, for some regions of the parameter space and for both DGPs, the simple MS model offers the best predictions, namely when $d = 0.25$ and $d = 0.49$, which is probably the most common interval for d in empirical applications. However, when the DGP is an ARFIMA (1, d, 0), for negative and/or high ϕ the MS performance worsens. On the other hand, the IMA model performed consistently well, while ARFIMA (1, d, 0) does the opposite, even when it is the true DGP. This is probably explained by the same sort of identification problem in estimating d and ϕ (especially when we let ϕ increase), which has already been documented in the literature (see, e.g. Pérez and Ruiz 2001). We also report averages of estimated transition probabilities for the MS model, with corresponding standard deviations. It is interesting to note that, as persistence increases in the DGP (either d or ϕ get larger), on average a more persistent MS model is fitted. Moreover, the transition probabilities are estimated with increasing precision. This means that a type of 'spurious switching' is occurring when the true process is a long memory one.

As for the DGP in (16), an upwards shift in the mean will worsen the predictive ability of the ARFIMA models, when compared to a 'reverting' shift (see lower part of Table 8). In this last case, the ARFIMAs are to be preferred to the other models, but are clearly worse in the

Table 8(a). Monte Carlo FMSE from DGP (14) and (16) with $\mu_2 = 2$.[a]

	(0, d, 0)		(1, d, 0)		RW		IMA		MS(2)	
h					NB					
1	1.540		1.384		2.000		1.069		0.974	
2	1.815		1.523		2.052		1.075		0.979	
3	2.082		1.687		2.050		1.072		1.015	
4	2.300		1.829		2.027		1.053		1.085	
5	2.534		2.013		2.053		1.062		1.020	
6	2.751		2.188		2.117		1.084		1.134	
7	2.908		2.312		2.123		1.060		1.094	
8	3.087		2.458		2.211		1.055		1.116	
9	3.194		2.547		2.183		1.051		1.145	
10	3.479		2.804		2.239		1.069		1.089	
11	3.617		2.936		2.216		1.069		1.149	
12	3.770		3.086		2.285		1.091		1.285	
13	3.957		3.266		2.341		1.065		1.243	
14	4.110		3.422		2.287		1.046		1.318	
15	4.115		3.460		2.330		1.076		1.236	
16	4.256		3.612		2.372		1.067		1.375	
Average	3.094		2.532		2.180		1.067		1.141	
	$\mu_3 = 1$	$\mu_3 = 3$	$\mu_3 = 1$	$\mu_3 = 3$	$\mu_3 = 1$	$\mu_3 = 3$	$\mu_3 = 1$	$\mu_3 = 3$	$\mu_3 = 1$	$\mu_3 = 3$
9	1.242	7.146	1.089	6.004	3.438	2.926	2.089	2.013	14.828	1.794
10	1.348	7.611	1.159	6.449	3.455	3.023	2.047	2.091	14.559	1.865
11	1.396	7.839	1.191	6.681	3.473	2.958	2.068	2.071	14.420	1.959
12	1.462	8.077	1.242	6.929	3.569	3.001	2.096	2.086	13.936	1.949
13	1.528	8.386	1.283	7.249	3.602	3.079	2.026	2.103	13.319	1.791
14	1.581	8.640	1.321	7.522	3.538	3.037	1.976	2.116	13.296	1.887
15	1.601	8.629	1.356	7.564	3.676	2.983	2.082	2.071	13.419	1.823
16	1.656	8.857	1.403	7.822	3.706	3.038	2.039	2.095	12.780	1.842
Average	1.464	8.148	1.256	7.028	3.557	3.006	2.053	2.081	13.820	1.864

[a]The values in the row 'Average' represent the means of each column. The reported FMSE's are obtained considering the 5000 replications. NB represents 'no break' in the forecasting period. From the 5000 replications the following results were obtained for the main parameters. Mean $d = 0.357$ (SE $= 0.050$); mean $p_{11} = 0.989$ (SE $= 0.010$); mean $p_{22} = 0.989$ (SE $= 0.014$).

first situation. This again is not surprising, since the fractional models will incorporate pre-shift information in their predictions.

5. CONCLUSION

Forecasting is a difficult task, which becomes even more complicated in a rapidly changing world, where structural changes may occur. Recent studies have focused on this issue, and the aim of this paper is to provide further insight to the problem. Given that economic time series usually display

Table 8(b). Monte Carlo FMSE from DGP (14) and (16) with $\mu_2 = 5$.

	(0, *d*, 0)		(1, *d*, 0)		RW		IMA		MS(2)	
h					NB					
1	1.596		1.443		2.003		1.210		0.975	
2	1.792		1.529		2.060		1.224		0.987	
3	2.101		1.584		2.065		1.218		1.033	
4	2.447		1.672		2.052		1.191		1.117	
5	2.874		1.832		2.091		1.203		1.085	
6	3.336		2.014		2.173		1.230		1.218	
7	3.771		2.164		2.203		1.201		1.192	
8	4.267		2.358		2.314		1.216		1.251	
9	4.682		2.490		2.327		1.197		1.281	
10	5.365		2.820		2.396		1.222		1.282	
11	5.866		3.022		2.413		1.210		1.383	
12	6.447		3.293		2.520		1.239		1.561	
13	7.112		3.607		2.599		1.218		1.593	
14	7.717		3.889		2.574		1.187		1.707	
15	8.129		4.081		2.694		1.212		1.639	
16	8.789		4.414		2.769		1.207		1.857	
Average	4.768		2.638		2.328		1.211		1.322	
	$\mu_3 = 1$	$\mu_3 = 9$	$\mu_3 = 1$	$\mu_3 = 9$	$\mu_3 = 1$	$\mu_3 = 9$	$\mu_3 = 1$	$\mu_3 = 9$	$\mu_3 = 1$	$\mu_3 = 9$
9	5.691	35.673	9.677	27.303	21.532	15.121	17.460	16.934	17.414	18.069
10	5.022	37.708	9.007	28.632	21.685	15.108	17.243	17.201	16.966	18.859
11	4.533	39.198	8.552	29.492	22.109	14.716	17.314	17.105	16.601	19.363
12	4.115	40.779	8.127	30.458	22.566	14.474	17.368	17.110	15.997	19.988
13	3.631	42.592	7.569	31.646	22.796	14.401	17.174	17.261	15.173	20.632
14	3.182	44.252	7.044	32.734	22.970	14.178	17.016	17.357	15.007	20.929
15	3.012	45.246	6.877	33.285	23.717	13.670	17.345	17.079	15.008	20.786
16	2.694	46.884	6.432	34.396	23.983	13.555	17.205	17.209	14.208	21.776
Average	3.985	41.542	7.910	30.993	22.670	14.403	17.266	17.157	15.797	20.050

[a]See notes on Table 8a. From the 5000 replications the following results were obtained for the main parameters: mean $d = 0.665$ (SE = 0.042); mean $p_{11} = 0.989$ (SE = 0.005); mean $p_{22} = 0.989$ (SE = 0.004).

high persistence and signs of structural breaks, it is natural to compare distinct modelling and forecasting methodologies, which try to address the different features of the data. By looking at the forecast performance of ARFIMA, MS and simple linear models, we tried to assess whether these approaches are flexible enough to cope with changes in parameters.

Although long memory models may capture some in-sample features of the data, we found that when shifts occur in the series we considered, their forecast performance is relatively poor when compared to MS models. It seems, therefore, that forecasting from 'spuriously' fitted long memory models does not carry any gains in terms of common forecast accuracy measures, even in parameter settings where we could expect the two types of model not to be easily distinguished. This result is related to the way that predictions are constructed for ARFIMA processes (as pointed out in

Table 8(c). Monte Carlo FMSE from DGP (14) and (16) with $\mu_2 = 10$.[a]

	(0, d, 0)	(1, d, 0)	RW	IMA	MS(2)
h			NB		
1	1.741	1.689	2.012	1.420	0.980
2	1.768	1.726	2.089	1.443	1.014
3	1.866	1.723	2.127	1.432	1.104
4	2.005	1.734	2.159	1.396	1.210
5	2.228	1.819	2.256	1.410	1.258
6	2.516	1.932	2.414	1.438	1.451
7	2.787	2.013	2.535	1.411	1.482
8	3.131	2.142	2.748	1.443	1.662
9	3.412	2.209	2.898	1.412	1.759
10	3.934	2.424	3.067	1.442	1.902
11	4.309	2.543	3.235	1.415	2.119
12	4.802	2.738	3.501	1.447	2.459
13	5.376	2.980	3.719	1.443	2.687
14	5.889	3.169	3.851	1.401	2.934
15	6.306	3.317	4.219	1.417	2.960
16	6.967	3.560	4.475	1.421	3.456
Average	3.689	2.357	2.956	1.424	1.902

[a]See notes on Table 8a. From the 5000 replications the following results were obtained for the main parameters. Mean $d = 0.862$ (SE = 0.038); mean $p_{11} = 0.990$ (SE = 0.001); mean$p_{22} = 0.990$ (SE = 0.001).

Section 2.1), despite the theoretical and empirical similarities that may exist between fractional processes and series with breaks. By attributing weight to distant lags when forming forecasts, ARFIMA models will in general be slow to react to shifts in the series. On the other hand, we also found a 'spurious switching' phenomenon, where MS models may capture long memory behaviour. However, in some cases, this misspecification may be robust in terms of forecasting.

Moreover, our findings, in a more general framework, are in accordance with what Clements and Hendry (1998) and Clements and Krolzig (1998) claim, that is, that simple linear time series models remain useful tools for prediction. Indeed, our experiments allow us to conclude that the IMA model is the best predictor for most of the DGPs under study.

Obviously, the results in our paper are specific to the empirical data and the Monte Carlo design we have chosen. It would be useful to look at other situations and data, for instance financial data, where both long memory and structural change models are commonly used. On the other hand, it would also be interesting to analyse how these results would carry over other forecast settings, namely multivariate forecasting.

ACKNOWLEDGEMENTS

Financial support from the Fundação para a Ciência e Tecnologia and from the Research Unit in Economic Policy-NIPE is gratefully acknowledged. The authors thank an anonymous referee, the journal editor, Ron Smith, Haris Psaradakis, Martin Sola and participants at the 20th International Symposium on Forecasting, Lisbon, for helpful comments. The usual disclaimer applies.

Table 9(a). Monte Carlo FMSE from the MS DGP (15) with $T = 200, \mu_2 = 2$.[a]

	(0, d, 0)		(1, d, 0)		RW		IMA		MS(2)	
h	0.95	0.99	0.95	0.99	0.95	0.99	0.95	0.99	0.95	0.99
1	1.461	1.424	1.379	1.323	2.021	1.976	1.205	1.100	1.503	1.638
2	1.674	1.642	1.505	1.438	2.112	2.018	1.253	1.150	1.507	1.564
3	1.843	1.822	1.631	1.570	2.158	2.046	1.273	1.160	1.569	1.620
4	1.968	1.969	1.725	1.674	2.214	2.077	1.286	1.155	1.589	1.622
5	2.164	2.161	1.882	1.823	2.261	2.085	1.311	1.151	1.670	1.656
6	2.218	2.210	1.929	1.862	2.324	2.116	1.305	1.129	1.660	1.588
7	2.340	2.333	2.033	1.967	2.342	2.111	1.320	1.152	1.676	1.621
8	2.349	2.332	2.045	1.967	2.323	2.065	1.278	1.100	1.660	1.548
9	2.560	2.548	2.233	2.167	2.377	2.142	1.318	1.157	1.772	1.601
10	2.697	2.712	2.367	2.328	2.442	2.246	1.358	1.206	1.847	1.670
11	2.740	2.721	2.408	2.336	2.452	2.216	1.352	1.193	1.868	1.641
12	2.894	2.865	2.557	2.473	2.495	2.226	1.373	1.178	1.939	1.654
13	2.953	2.962	2.628	2.583	2.504	2.254	1.390	1.228	2.003	1.692
14	3.053	3.018	2.736	2.648	2.595	2.296	1.415	1.210	2.083	1.690
15	3.090	3.107	2.786	2.751	2.541	2.270	1.404	1.232	2.090	1.735
16	3.147	3.151	2.862	2.832	2.706	2.709	1.466	1.501	1.925	1.297
Average	2.446	2.436	2.169	2.108	2.366	2.178	1.331	1.187	1.772	1.614

[a]The values in the row Average represent the means of each column. The reported FMSE's are obtained considering the 5000 replications. The notation 0.95 and 0.99 in the second row represents $(p_{11}, p_{22}) = (0.95, 0.95)$ and $(p_{11}, p_{22}) = (0.99, 0.99)$, respectively. From the 5000 replications the following results were obtained: for 0.95, mean $d = 0.334$ (SE $= 0.040$); for 0.99, mean $d = 0.319$ (SE $= 0.058$).

REFERENCES

Andrews, D. W. K. (1993). Tests for parameter instability and structural changes with unknown change points. *Econometrica 61*, 821–56.

Baillie, R. T. (1996). Long memory processes and fractional integration in econometrics. *Journal of Econometrics 73*, 5–59.

Baillie, R. T., C. F. Chung and M. A. Tieslau (1996). Analysing the inflation rate by the fractionally integrated ARFIMA-GARCH model. *Journal of Applied Econometrics 11*, 23–40.

Bos, C. S., P. H. Franses and M. Ooms (1998). Long memory and level-shifts: Re-analyzing inflation rates. Econometrics Institute Report 9811, Erasmus University Rotterdam.

Clements, M. P. and D. F. Hendry (1998). Forecasting economic processes. *International Journal of Forecasting 14*, 111–31.

Clements, M. P. and D. F. Hendry (2000). Forecasting with difference-stationary and trend-stationary models. *Econometrics Journal 4*, 1–19.

Clements, M. P. and H.-M. Krolzig (1998). A comparison of the forecast performance of markov-switching and threshold autoregressive models of US GNP. *Econometrics Journal 1*, 47–75.

Crato, N. and B. K. Ray (1996). Model selection and forecasting for long-range dependent processes. *Journal of Forecasting 15*, 107–25.

Diebold, F. X. and A. Inoue (2001). Long memory and structural change. *Journal of Econometrics 105*, 131–59.

Table 9(b). Monte Carlo FMSE from the MS DGP (15) with $T = 200, \mu_2 = 5$.[a]

	(0, d, 0)		(1, d, 0)		RW		IMA		MS(2)	
h	0.95	0.99	0.95	0.99	0.95	0.99	0.95	0.99	0.95	0.99
1	2.362	1.692	2.399	1.594	2.808	2.118	2.350	1.474	8.238	10.405
2	3.057	2.048	3.129	1.863	3.614	2.304	3.036	1.682	8.245	9.938
3	3.603	2.389	3.713	2.081	4.270	2.469	3.587	1.800	8.100	9.958
4	4.169	2.822	4.280	2.377	4.969	2.692	4.186	1.978	7.925	9.808
5	4.675	3.238	4.786	2.634	5.490	2.818	4.623	2.103	7.985	9.955
6	5.052	3.495	5.186	2.782	6.037	2.975	5.069	2.166	7.790	9.792
7	5.523	3.901	5.635	3.041	6.576	3.143	5.539	2.366	7.863	9.900
8	5.790	4.103	5.927	3.144	6.933	3.192	5.816	2.366	7.753	9.383
9	6.216	4.647	6.312	3.547	7.285	3.501	6.095	2.644	7.901	9.462
10	6.568	5.130	6.636	3.891	7.567	3.751	6.307	2.847	8.035	9.427
11	6.804	5.310	6.894	3.979	7.876	3.859	6.525	2.916	8.073	9.088
12	7.224	5.710	7.300	4.212	8.242	3.944	6.826	2.949	8.233	9.184
13	7.390	6.124	7.443	4.499	8.370	4.108	6.898	3.115	8.339	9.002
14	7.791	6.477	7.862	4.742	8.842	4.329	7.230	3.242	8.610	9.168
15	7.899	6.854	7.951	5.014	8.903	4.445	7.270	3.368	8.565	9.083
16	7.840	7.079	7.958	5.146	9.471	4.583	7.010	3.602	8.208	9.048
Average	5.748	4.438	5.838	3.409	6.723	3.389	5.522	2.538	8.116	9.537

[a]See notes on Table 9a. From the 5000 replications the following results were obtained: for 0.95, mean $d = 0.636$ (SE = 0.052); for 0.99, mean $d = 0.570$ (SE = 0.122).

Elliott, G., T. J. Rothenberg and J. H. Stock (1996). Efficient tests for an autoregressive unit root. *Econometrica 64*, 813–36.

Evans, M. and P. Wachtel (1993). Inflation regimes and the sources of inflation uncertainty. *Journal of Money, Credit and Banking 25*, 475–511.

Fox, R. and M. Taqqu (1986). Large-sample properties of parameter estimates for strongly dependent stationary gaussian time series. *Annals of Statistics 14*, 517–32.

Garcia, R. and P. Perron (1996). An analysis of the real interest rate under regime shifts. *Review of Economics and Statistics 78*, 111–25.

Geweke, J. and S. Porter-Hudak (1983). The estimation and application of long memory time series models. *Journal of Time Series Analysis 4*, 221–38.

Gourieroux, C. and J. Jasiak (2001). Memory and infrequent breaks. *Economics Letters 70*, 29–41.

Granger, C. W. J. and N. Hyung (1999). Occasional structural breaks and long memory, UCSD Discussion Paper 99-14.

Granger, C. W. J. and R. Joyeux (1980). An introduction to long-memory time series models and fractional differencing. *Journal of Time Series Analysis 1*, 15–39.

Granger, C. W. J. and T. Teräsvirta (1999). A simple nonlinear time series model with misleading linear properties. *Economics Letters 62*, 161–65.

Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica 57*, 357–84.

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

Table 10. Monte Carlo FMSE from the ARFIMA DGPs with $T = 200$.[a]

	(0, d, 0)	(1, d, 0)	RW	IMA	MS(2)	$\hat{p}_{11}$, $\hat{p}_{22}$
			(0, d, 0)			
$d = -0.49$	1.231	1.287	2.806	1.288	1.530	0.233, 0.265 (0.200,0.210)
$d = -0.25$	1.057	1.131	2.365	1.090	1.472	0.286, 0.274 (0.256,0.252)
$d = 0.25$	1.073	1.179	1.889	1.171	1.125	0.569, 0.571 (0.293,0.293)
$d = 0.49$	1.273	1.891	1.802	1.439	1.165	0.940, 0.939 (0.038,0.040)
$d = 0.75$	1.792	2.386	1.943	1.837	3.072	0.974, 0.973 (0.021,0.021)
			(1, d, 0)			
$(d, \phi) = (0.25,-0.75)$	1.784	1.861	3.738	1.839	2.119	0.096, 0.086 (0.077,0.068)
$(d, \phi) = (0.25,-0.25)$	1.028	1.052	1.903	1.041	1.350	0.352, 0.410 (0.383,0.367)
$(d, \phi) = (0.25,0.25)$	1.249	1.595	1.947	1.548	0.958	0.880, 0.854 (0.090,0.145)
$(d, \phi) = (0.25,0.75)$	2.642	3.710	2.742	2.751	2.459	0.945, 0.942 (0.026,0.026)
$(d, \phi) = (0.49,-0.75)$	1.661	1.753	2.884	1.573	1.936	0.078, 0.082 (0.067,0.075)
$(d, \phi) = (0.49,-0.25)$	1.167	1.190	1.744	1.186	1.065	0.884, 0.845 (0.206,0.274)
$(d, \phi) = (0.49,0.25)$	1.671	2.392	2.074	1.944	1.631	0.942, 0.943 (0.034,0.034)
$(d, \phi) = (0.49,0.75)$	3.847	7.632	3.741	3.642	5.045	0.968, 0.969 (0.019,0.018)

[a]The results are averages for $h = 4$. Average standard deviations for estimated transition probabilities are presented in parentheses.

Hassler, U. and J. Wolters (1995). Long memory in inflation rates: international evidence. *Journal of Business and Economic Statistics 13*, 37–45.

Kim, C.-J. (1994). Unobserved-component time series models with markov-switching heteroskedasticity: Changes in regime and the link between inflation and inflation uncertainty. *Journal of Business and Economic Statistics 12*, 157–79.

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

Nunes, L. C., P. Newbold and C. M. Kuan (1997). Testing for unit roots with breaks: evidence on the great crash and the unit root hypothesis reconsidered. *Oxford Bulletin of Economics and Statistics*, *59*, 435–48.

Ooms, M. and J. Doornik (1999). Inference and forecasting for fractional autoregressive integrated moving average models with an application to US and UK inflation. Econometrics Institute Report 9947-A, Erasmus University Rotterdam.

Pérez, A. and E. Ruiz (2001). Finite sample properties of a QML estimator of stochastic volatility models with long memory. *Economics Letters 70*, 157–64.

Sowell, F. (1992). Maximum likelihood estimation of stationary univariate fractionally integrated time series models. *Journal of Econometrics 53*, 165–88.

Stock, J. H. and M. W. Watson (1999). Forecasting inflation. *Journal of Monetary Economics 44*, 293–335.

Timmermann, A. (2000). Moments of markov switching models. *Journal of Econometrics 96*, 75–111.