NIPE

Documentos de Trabalho

Working Paper Series

The Forecast Performance of Long Memory and Markov Switching Models

Vasco J. Gabriel e Luís F. Martins

NIPE WP 2 / 2000

NÚCLEO DE INVESTIGAÇÃO EM POLÍTICAS ECONÓMICAS

UNIVERSIDADE DO MINHO

The Forecast Performance of Long Memory and Markov Switching Models

Vasco J. Gabriel[*] and Luis F. Martins[†]

June 2000

Abstract

Recent research has focused on the links between long memory and structural change, stressing the long memory properties that may arise in models with parameter changes. In this paper, we contribute to this research by comparing the forecasting abilities of long memory and Markov switching models. Two approaches are employed: a 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, when shifts occur in the series considered, their forecast performance is relatively poor, when compared with simple linear and Markov switching models. Moreover, our findings, in a more general framework, are in accordance with the works of Clements and Hendry (1998) and Clements and Krolzig (1998), reinforcing the idea that simple linear time series models remain useful tools for prediction.

Key Words: Long Memory; Structural change; Forecasting

JEL Classification: C12; C22; C52

[*] Department of Economics, University of Minho (email: vjgabriel@eeg.uminho.pt). Financial support from the Sub-Programa Ciência e Tecnologia do Segundo Quadro Comunitário de Apoio, grant number PRAXIS XXI/BD/16141/98 and from the Research Unit in Economic Policy - NIPE is gratefully acknowledged.

[†] Instituto Superior de Ciências do Trabalho e da Empresa, UNIDE, Lisbon (email:luis.martins@iscte.pt)

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, Diebold and Inoue (1999), Granger and Teräsvirta (1999) and Granger and Hyung (1999) show analytically and via Monte Carlo that models with structural change may exhibit long memory properties. Earlier, Hidalgo and Robinson (1996) developed a test for structural change in a long memory context, which has not been much used in practice (see, however, the work of Bos, Franses and Ooms, 1998).

What are the implications of these results for forecasting? Since long memory and structural breaks may be hard to distinguish in practice, we investigate whether a long memory approach will be "robust" to structural breaks in a time series, in terms of providing useful forecasts for financial and macroeconomic data. The question of the relative forecast performance of long memory and structural change models has not (to our knowledge) been addressed yet.

Therefore, in this paper we compare the univariate forecast accuracy of one type of structural change model, the Markov Switching (MS) model, and fractionally integrated ARMA (ARFIMA) models . We conduct our analysis 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. Other structural change models may have been considered, but we stress the MS specification, since it is a widely used approach to model changes in parameters.

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 data generating process (DGP) produced a single change in the mean.

Nevertheless, none of these works considered the more general linear ARFIMA model. Since long memory and regime shifts are intimately related and may easily be confused in many empirical situations, it is of obvious interest to assess how long memory models behave in terms of forecasting when time series su¤er regime shifts[1]. Therefore, our paper may be viewed as the implementation of the ideas in Diebold and Inoue (1999), inter alia, to forecasting problems and as a complement to the studies of Clements and Hendry (1998) and Clements and Krolzig (1998).

In the Monte Carlo experiment, we extend the simulations in Clements and Hendry (1998) by including long memory and MS models and evaluating their forecast accuracy under di¤erent DGP's with parameter changes. Concerning our empirical illustration, we use MS and ARFIMA speci…cations to model the empirical path of the in‡ation rate in Portugal and then 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 in‡ation for the period 1968:1-1998:4. We re…ne our early Monte Carlo study with further simulations, using empirical estimates as parameter values for the DGP. Obviously, by focusing on univariate methods we are simplifying our analysis, mainly for expositional simplicity. Nevertheless, this may be viewed as a …rst approximation to more evolved forecasting practices, since univariate forecasts are usually taken as benchmarks for later comparisons[2].

The paper proceeds as follows. In section 2, we brie‡y review modelling and forecasting with ARFIMA and Markov switching models, introducing de…nitions and notation, and consider why structural breaks may cause the appearance of long memory characteristics in a given time series. The next section presents a set of Monte Carlo simulations and results. Section 4 discusses empirical aspects of our example, including a forecasting exercise, complemented by Monte Carlo analysis in the next section. Finally, section 6 provides some discussion and conclusions. All unreported results are available upon request.

2 Long Memory and Structural Change 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), for example. Fractional integration, as in Granger (1980) and Granger and Joyeux (1980), aims to circumvent some of the limitations of integer analysis

[1] Although "spurious long memory" may arise in this situation (as stressed by Granger and Teräsvirta, 1999 and Granger and Hyung, 1999), an ARFIMA speci…cation may still be a useful tool for forecasting.

[2] See Stock and Watson (1999) for a recent discussion on forecasting in‡ation.

of ARIMA models. A fractionally integrated ARMA process y_t may be represented by

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

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 $(1-L)^d = 1 - dL + d(d-1)/2!L^2 - d(d-1)(d-2)/3!L^3 + \ldots$, for $d > -1$. The process is stationary and invertible if the roots of the autoregressive polynomial of order p, $\Phi(L) = 1 - \phi_1 L - \ldots - \phi_p L^p$; and of the moving-average part of order q, $\Theta(L) = 1 + \theta_1 L + \ldots + \theta_q L^q$; 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 traditionally 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 analyzing 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 $\Pi(L) = (1 - \pi_1 L - \pi_2 L^2 - \ldots) = \Phi(L)(1-L)^d \Theta(L)^{-1}$. This form obviously needs truncation after k lags, but unfortunately there is no solution on how to proceed in this case. The truncation problem will also be related to the forecast horizon considered in the predictions (see Crato and Ray, 1996).

2.2 Markov Switching Models

The importance of non-linearities (along with structural changes) in economic series has always been debated in the literature. The discussion was further intensified since Hamilton (1989) proposed his autoregressive Markov switching model to analyze US GNP growth rate. It offers a powerful and flexible instrument to characterize macroeconomic fluctuations, by accommodating asymmetries and changes in the behavior of economic time series. Several extensions and generalizations have been presented, see Kim and Nelson (1998), 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 in $S = \{1;2\}$, 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 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 and Kim and Nelson, 1998, for more details).

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

$$\hat{\xi}'_t = \begin{bmatrix} p(s_t = 1|Y_t) & p(s_t = 2|Y_t) & \cdots & p(s_t = N|Y_t) \end{bmatrix} \tag{6}$$

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

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

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

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

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), the second term in the summation is $\mu(s_{t+h}) + \phi[y_{t+h-1} - \mu(s_{t+h-1})]$, so we have

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

where $\hat{\mu}(s_{t+h|t}) = \sum_{j=1}^{2} \mu_j \Pr(s_{t+h} = j|Y_t) = \bar{\mu}$ and, thus,

$$\hat{y}_{t+h|t} = \hat{\mu}(1 - \phi^h) + \phi^h y_t. \tag{9}$$

Of course, this recursion could be easily extended to more complicated models (see Hamilton, 1994).

2.3 Long Memory in Markov Switching Models

As mentioned in the introduction, some recent papers deal with the relationship between long memory and structural change, namely stochastic regime switching. Diebold and Inoue (1999), Granger and Hyung (1999) and Granger and Teräsvirta (1999) analyzed several cases with stochastic parameter shifts, by looking at the behavior 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 (that is, as p_{11} and p_{22} approach unity in the Markov switching case), the process will closely resemble a fractionally integrated series. Complementarily, the size of the structural breaks will also be a factor, since a similar effect will arise for larger magnitudes of breaks.

This can be easily verified by considering an example with the simple two-regimes first-order autoregressive Markov switching model in (4), $y_t - \mu(s_t) = \phi[y_{t-1} - \mu(s_{t-1})] + \sigma(s_t)\varepsilon_t$. The corresponding population autocorrelation function is given by

$$\rho_{T,k} = \frac{\pi_1\pi_2(\mu_1 - \mu_2)^2 \mathrm{vec}(P^k)v_1 + \phi^k\pi'(I_k - \phi^2 B)^{-1}\sigma_s}{(\pi_1\mu_1 + \pi_2\mu_2)^2 + (\pi_1\sigma_1^2 + \pi_2\sigma_2^2)(1+\phi^2)^{-1}}, \tag{10}$$

where π_j represents the ergodic probability of staying in regime j $(j = 1,2)$; $\pi = [\pi_1, \pi_2]$ $v_1 = [\pi_2, -\pi_2, \pi_1, -\pi_1]$, $\sigma_s = \left[\sigma_1^2, \sigma_2^2\right]$, 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, Proposition 2 and 4). Setting $\sigma_1^2 = \sigma_2^2 = 1$, $p_{11} = p_{22} = (0.95, 0.98, 0.99)$, $\mu_1 = 1$ 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 that the rate of decay of the autocorrelations slows down as the transition probabilities, the size of the shift and the autoregressive parameter increase. 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[3], may be easily confused with a random walk. See Timmermann (2000, Section 6) and Nunes et al. (1997).

Furthermore, accounting for eventual shifts in the process has the effect of reducing the estimated fractional integration parameter, $\hat{d}$, according to Bos, Franses and Ooms (1998) and Granger and Hyung (1999), which indicates that long memory may arise due to neglected shifts. However, Granger and Hyung (1999) argue that a "spurious break"-type of phenomenon[4] 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 it 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 a MS model will be generated in a way that will resemble a linear prediction rule. On the other hand, it is expected that a MS model will perform better when the regimes are more persistent and for larger breaks. Note, however, that these same factors that favor prediction from MS models are central for the result that a MS process will display long memory

[3] Such as the case of large permanent changes.

[4] See Nunes et al (1997).

properties. Hence, this adds relevance to our study, since it will be interesting to assess if the empirical similarities between the two models will continue to hold in terms of forecasting.

3 Monte Carlo Study

In order to compare the relative merits of long memory and MS models, we designed a set of simple Monte Carlo simulations. As with all Monte Carlo experiments, there is always an inevitable specificity concerning the DGP's and the obtained results. However, we stress what is essential to our case, that is, magnitude and frequency of parameter switching, as discussed in the previous section. To simplify, the variance is kept constant.

Hence, we base our simulations on the DGP studied by Clements and Hendry (1998). We consider the simple switching-mean process

$$y_t = \mu_t + \varepsilon_t; \qquad t = 1; \ldots; T \tag{11}$$

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

$$\mu_t = \begin{cases} \mu_1; & t \leq \tau \\ \mu_2; & t > \tau \end{cases}; \tag{12}$$

where τ is a 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 $\mu_2 = (2; 5; 10)$: The case $\mu_2 = 10$ corresponds to the DGP analyzed in Clements and Hendry (1998), but we also wish to consider other empirically relevant shift magnitudes. For simplicity, we let $\tau = T/2$ and we generate $T = 100$ plus $h = 16$ random observations in each replication, where the last h observations are held back for the forecast simulation.

Another interesting situation that merits attention is when structural change occurs in the forecasting period. It is of great interest to see how different models may be "robust", in terms of adapting their forecasts to a change outside the sample period. Thus, we modify the previous DGP by assuming that

$$\mu_t = \begin{cases} \mu_1; & t \leq \tau \\ \mu_2; & \tau < t \leq T + h/2 \\ \mu_3; & t > T + h/2 \end{cases}; \tag{13}$$

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 $(1; 9)$:

Finally, we specify a Markov switching DGP where μ_t now 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 attempt here to experiment different settings for the p_{ij}'s without neglecting their empirical congruence. The variance is the same for the two regimes. For this specific DGP, we consider a sample size of 200 observations, given the persistence in the regimes we are considering.. In all experiments, the number of replications was 5000 and the criterion used for comparisons is the forecast mean-squared error (FMSE).

For comparison purposes, we consider different types and classes of models. In each replication, we fit a simple Markov switching-mean model, ARFIMA $(0, d, 0)$ and ARFIMA $(1, d, 0)$ models, a random walk (RW) and an integrated moving-average model (IMA), and compute the respective forecasts. This last model (IMA) was found to be one of the most robust forecasting devices by Clements and Hendry (1998). We tried different specifications for the ARFIMA models, but in general the ones considered here worked better in terms of forecasting. Prediction for the ARFIMA's from (3) was conducted with $k = 10.$ Regarding the estimation method, we adopted the frequency domain estimator of Fox and Taqqu (1986) throughout the paper[5]. All results were obtained using routines written in GAUSS.

Tables 2 to 3 show the results of the simulations for the three DGP's under study. An overall conclusion, in line with what the literature implies, is that for larger magnitudes of shifts one gets higher estimates for d. That also leads to a decrease in the predictive ability for all models. Our experiments also allow us to conclude that the IMA model is the best predictor for most of the DGP's under study, which reinforces the result in Clements and Hendry (1998)[6]. On the other hand, the ARFIMA specifications are not, in general, robust predictors. Although their ability to forecast for shorter periods is reasonable, it rapidly deteriorates and, on average, it is even worse than the RW. Relatively to the MS approach, it is generally superior to the ARFIMA's, and occasionally better than the IMA, especially for shorter forecasting periods.

The conclusions for each DGP are not very distinct from what was outlined above. Nevertheless, it is worth mentioning that for the DGP in (13), an upward shift in the mean will worsen the predictive ability of all models (except for the IMA, where the difference is negligible), when compared to a "reverting" shift. In this last case, the ARFIMA's are to be preferred to the other models, but

[5] Again, we tried different procedures, such as the GPH estimator (see Geweke and Porter-Hudak, 1983) and the exact maximum likelihood method of Sowell (1992), but the one we adopted seemed to do better.

[6] This may be explained by the fact that the MA component captures the previous error, thus improving the forecast comparatively with the RW with no MA component.

are clearly worse in the …rst situation. As for the MS DGP in (14), we observe that less frequent switching improves the performance of all models. Curiously, the average $\hat{d}$ decreases slightly in this situation (see notes of Table 3), although the estimates are not signi…cantly di¤erent for $p_{ij} = 0:95$ and $p_{ij} = 0:99$ $(i = j)$. Overall, the IMA model is still the best, while the ARFIMA's 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.

4 The In‡ation Rate in Portugal: an Empirical Example

In this section, we analyze, in a simpli…ed manner, the univariate properties of the quarterly CPI in‡ation rate in Portugal for the sample period 1968:1-1999:4. The series is constructed by taking …rst-di¤erences 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 concentrate on the other features of the data[7]. For this period, some major events in Portugal led to changes in economic policy and substantial ‡uctuations in the in‡ation rate: the two oil shocks, the democratic Revolution with the subsequent loss of its colonies (1974, 1975), two agreements with the International Monetary Fund, 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 di¤erent regimes. On the other hand, these events led to an increased persistence in the in‡ation rate in Portugal, when compared to other European countries. In fact, the series shows the typical behavior of a series with long memory, with a very slow return to a low in‡ation regime after a big shock, so one may expect a high estimate for the order of integration.

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

In turn, MS models are particularly suitable to analyze some of the dynamic features of in‡ation rates, namely capturing the apparent changes in mean and variance. Regime shifts in in‡ation rates

[7] We considered di¤erent methods to account for seasonality, but the results of our subsequent analysis did not change qualitatively.

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.

Before proceeding with the forecasting exercise, we present in Table 4 some tests concerning the properties of the data. Different unit root tests (ADF, Phillips-Perron and DF-GLS as in Elliott, Rothenberg and Stock, 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 type 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 for testing whether the series is $I(0)$ or $I(1)$. Looking at the estimates of d and respective standard errors (Table 5), using different estimation methods, it can easily be seen that both the $I(0)$ and $I(1)$ hypothesis are rejected[8]. Therefore, it is difficult to state clearly 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[9]. However, introducing an autoregressive component induces an increase in the estimated d.

Regarding the estimation of MS models, we present in Table 5 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 point to different means and variances within the sample period[10]. Moreover, the estimated transition probabilities indicate that the regimes are quite persistent. Therefore, it is not surprising to find evidence of long memory in the series, considering the results in Diebold and Inoue (1999), inter alia.

Turning to the forecast comparison, in terms of FMSE and forecast mean absolute error (FMAE), we include again the RW and IMA models in the results presented in Table 6. 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

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

[9] The estimates of d range from approximately 0.3 to 0.7 using other estimation methods.

[10] 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 the matter.

of the ARFIMA $(1; d; 0)$ model, which ranks second for the shorter horizon and a tied-second with the IMA model for the longer horizon. The simplest ARFIMA $(0; d; 0)$ also works well, ranking fourth for each prediction period. Using di¤erent 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. 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.

5 Further Monte Carlo Analysis

In this section, we re…ne our previous Monte Carlo simulations by taking an empirical model of the in‡ation rate as the DGP. Although the arti…cial DGP is useful in this context, it is preferable to use more economically meaningful estimated models, even if these only o¤er a poor approximation to the true DGP. This practice also permits to control for sampling variability of a one-shot type of forecast comparison as in the previous section, with the empirical example.

Having considered this, we base our DGP in this experiment on the simple MS model, since it provides a simple, yet rough, description of the data, by estimating changes in mean and variance. However, we restrict the break points to be those obtained from observing the …ltered regime probabilities for the simple MS model. Furthermore, we consider a smaller value for the variance of the last regime, which is in accordance with what is observed in the series.

Thus, the DGP is given by $y_t = \mu_t + \sigma_t \varepsilon_t$; with

$$\mu_t = \begin{cases} \mu_1 = 1.8; & \sigma_1^2 = 2; \qquad t \leq 24 \\ \mu_2 = 5.2; & \sigma_2^2 = 8; \qquad 24 < t \leq 74 \\ \mu_3 = 1.8; & \sigma_3^2 = 1; \qquad t > 74 \end{cases} .$$

We generate 5000 series of 128 observations, retaining 16 observations for forecasts comparisons. While this DGP is not truly a MS process (there is no Markov chain behind it), it may be viewed as one with …xed break points.

As expected , the MS model does relatively well, since it is the closest to the speci…ed DGP (see Table 7). However, the IMA model performs even better, which, again, is not surprising, given the results in Clements and Hendry (1998) and our previous simulations. As for the ARFIMA models, although they provide reasonable forecasts for shorter periods, their performance quickly deteriorates as the forecast horizon increases, which is line with the results we get from the other Monte Carlo study. Of course, for other plausible DGPs, the results and the ranking could be di¤erent.

6 Conclusion

Forecasting is a quite di¢cult 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 high persistence and signs of structural breaks, it is natural to try to compare distinct modelling and forecasting methodologies, which try to address the di¤erent 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 ‡exible 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 with IMA and MS models. Moreover, our …ndings, 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.

Obviously, the results in our paper are speci…c 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 …nancial data, where both long memory and structural change models are commonly used. On the other hand, it would also be interesting to analyze how these results would carry over other forecast settings, namely multivariate forecasting.

References

[1] Andrews, D. W. K. (1993), Tests for Parameter Instability and Structural Changes with Unknown Change Points, Econometrica, 61, 821-856.

[2] Baillie, R. T. (1996), Long Memory Processes and Fractional Integration in Econometrics, Journal of Econometrics, 73, 5-59.

[3] Baillie, R. T., Chung, C.-F, and Tieslau, M. A. (1996), Analysing the In‡ation Rate by the Fractionally integrated ARFIMA-GARCH Model, Journal of Applied Econometrics, 11, 23-40.

[4] Bos, C. S., Franses, P. H. and Ooms, M. (1998), Long Memory and Level-shifts: Re-analyzing In‡ation Rates, Econometrics Institute Report 9811, Erasmus University Rotterdam.

[5] Clements, M. P. and Hendry, D. F. (1998), Forecasting Economic Processes, International Journal of Forecasting, 14, 111-131.

[6] Clements, M. P. and Krolzig, H.-M. (1998), A Comparison of the Forecast Performance of Markov-switching and Threshold Autoregressive models of US GNP, Econometrics Journal, 1, 47-75.

[7] Crato, N. and Ray, B. K. (1996), Model Selection and Forecasting for Long-Range Dependent Processes, Journal of Forecasting, 15, 107-125.

[8] Diebold, F. X. and Inoue, A. (1999), Long Memory and Structural Change, manuscript.

[9] Elliott, G., Rothenberg, T. J. and Stock, J. H. (1996), E¢cient Tests for an Autoregressive Unit Root, Econometrica, 64, 813-836.

[10] Evans, M. and Wachtel, P. (1993), In‡ation Regimes and the Sources of In‡ation Uncertainty, Journal of Money, Credit and Banking, 25, 475-511.

[11] Fox, R. and Taqqu, M. (1986), Large-Sample Properties of Parameter Estimates for Strongly Dependent Stationary Gaussian Time Series, Annals of Statistics, 14, 517-532.

[12] Garcia, R. and Perron, P. (1996), An Analysis of the Real Interest Rate under Regime Shifts, Review of Economics and Statistics, 78, 111-125.

[13] Geweke, J. and Porter-Hudak, S. (1983), The Estimation and Application of Long Memory Time Series Models, Journal of Time Series Analysis, 4, 221-238.

[14] Granger, C. W. J. (1980), Long Memory Relationships and the Aggregation of Dynamic Models, Journal of Econometrics, 14, 227-238.

[15] Granger, C. W. J. and Hyung, N. (1999), Occasional Structural Breaks and Long Memory, UCSD Discussion Paper 99-14.

[16] Granger, C. W. J. and Joyeux, R. (1980), An Introduction to Long-Memory Time Series Models and Fractional Di¤erencing, Journal of Time Series Analysis, 1, 15-39.

[17] Granger, C. W. J. and Teräsvirta, T. (1999), A Simple Nonlinear Time Series Model with Misleading Linear Properties, Economics Letters, 62, 161-165.

[18] Hamilton, J. D. (1989), A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle, Econometrica, 57, 357-384.

[19] Hamilton, J. D. (1994), Time Series Analysis, Princeton University Press: Princeton.

[20] Hassler, U. and Wolters, J. (1995), Long Memory in In‡ation Rates: International Evidence, Journal of Business and Economic Statistics, 13, 37-45.

[21] Hidalgo, J. and Robinson, P.M. (1996), Testing for Structural Change in a Long Memory Environment, Journal of Econometrics, 70, 159-174.

[22] Kim, C.-J. (1994), Unobserved-Component Time Series Models with Markov-Switching Heteroskedasticity: Changes in Regime and the Link between In‡ation and In‡ation Uncertainty, Journal of Business and Economic Statistics, 12, 157-179.

[23] Kim, C.-J. and Nelson, C. R. (1999) State-space Models with Regime Switching, MIT Press: Cambridge.

[24] Nunes, L. C., Newbold, P. e Kuan, C. M.(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, 435-448.

[25] Ooms, M. and Doornik, J. (1999), Inference and Forecasting for Fractional Autoregressive Integrated Moving Average Models with an Application to US and UK In‡ation, Econometrics Institute Report 9947-A, Erasmus University Rotterdam.

[26] Timmermann, A., (2000), Moments of Markov Switching Models, Journal of Econometrics, 96, 75-111.

[27] Sowell, F. (1992), Maximum Likelihood Estimation of Stationary Univariate Fractionally Integrated Time Series Models, Journal of Econometrics, 53, 165-188.

[28] Stock, J. H. and Watson, M. W. (1999), Forecasting In‡ation, Journal of Monetary Economics, 44, 293-335.

7 Appendix

TABLE 1 - Autocorrelation function of an autoregressive Markov switching-mean model

$p_{11} = p_{22} = 0{:}95$									
	Á = 0:2			Á = 0:5			Á = 0:9		
$^1{}_2$	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

$p_{11} = p_{22} = 0{:}98$									
	Á = 0:2			Á = 0:5			Á = 0:9		
$^1{}_2$	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

$p_{11} = p_{22} = 0{:}99$									
	Á = 0:2			Á = 0:5			Á = 0:9		
$^1{}_2$	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

TABLE 2 a)[11]- Monte Carlo FMSE from DGP (12) and (13) with $^{1}{}_{2} = 2$

	(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	
	$^{1}{}_{3} = 1$	$^{1}{}_{3} = 3$	$^{1}{}_{3} = 1$	$^{1}{}_{3} = 3$	$^{1}{}_{3} = 1$	$^{1}{}_{3} = 3$	$^{1}{}_{3} = 1$	$^{1}{}_{3} = 3$	$^{1}{}_{3} = 1$	$^{1}{}_{3} = 3$
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

[11] The values in the row "Average" represents 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 (s.e. = 0:050); mean p_{11} = 0:989 (s.e. = 0:010); mean p_{22} = 0:989 (s.e. = 0:014):

TABLE 2 b)[12]- Monte Carlo FMSE from DGP (12) and (13) with ¹$_2$ = 5

	(0; d; 0)		(1; d; 0)		RW		IMA		M S(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	
	¹$_3$ = 1	¹$_3$ = 9	¹$_3$ = 1	¹$_3$ = 9	¹$_3$ = 1	¹$_3$ = 9	¹$_3$ = 1	¹$_3$ = 9	¹$_3$ = 1	¹$_3$ = 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

[12] See notes of Table 2 a).

From the 5000 replications the following results were obtained for the main parameters: mean d = 0:665 (s.e. = 0:042); mean p_{11} = 0:989 (s.e. = 0:005); mean p_{22} = 0:989 (s.e. = 0:004):

TABLE 2 c)[13]- Monte Carlo FMSE from DGP (12) and (13) with ${}^{1}{}_{2}$ = 10

	(0; d; 0)	(1; d; 0)	RW	IMA	MS(2)
h			NB		
1	1:740	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

[13] See notes of Table 2 a).

From the 5000 replications the following results were obtained for the main parameters: mean d = 0:862 (s.e. = 0:038); mean p_{11} = 0:990 (s.e. = 0:001); mean p_{22} = 0:990 (s.e. = 0:001):

TABLE 3 a) [14]- Monte Carlo FMSE from DGP (14) with $T = 200; \mu_2 = 2$

h	(0; d; 0)		(1; d; 0)		RW		IMA		MS(2)	
	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

[14] The values in the row "Average" represents 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 (s.e. = 0:040);

For 0:99; mean d = 0:319 (s.e. = 0:058);

TABLE 3 b) [15]- Monte Carlo FMSE from DGP (14) with T = 200; $^{1}{}_{2}$ = 5

h	(0; d; 0)		(1; d; 0)		RW		IMA		MS(2)	
	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

[15] See notes of Table 3 a).

From the 5000 replications the following results were obtained:

For 0:95; mean d = 0:636 (s.e. = 0:052);

For 0:99; mean d = 0:570 (s.e. = 0:122):

TABLE 4 [16]- Unit Roots, Stationarity and Structural Changes Tests for the Inflation Rate in Portugal

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^{**}

[16] The lag length for the ADF and DF-GLS tests is selected according to a t-test downward selection procedure, by setting the maximum lag equal to 8 and then testing downward 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 5 [17]- Estimation Results for the Inflation Rate in Portugal (1968:1 - 1998:4)

	(0; d; 0)	(1; d; 0)	MS(2)	MS(2) − AR(4)	MS(3) − AR(2)
d	0.477 (0.055)	0.712 (0.08)			
ϕ		−0.469 (0.094)			
σ^2	2.565 (0.163)	2.396 (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)
σ_1^2			2.025 (0.349)	0.582 (0.121)	0.405 (0.121)
σ_2^2			7.96 (1.599)	7.953 (1.465)	2.104 (0.476)
σ_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)

[17] Standard errors in brackets.

TABLE 6 - 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
$(1, d, 0)$	0.273	0.457	0.308	0.422
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

TABLE 7 [18]- Monte Carlo FMSE for the empirical-based DGP

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

[18] From the 5000 replications the following results were obtained for the main parameters: mean d = 0:388 (s.e. = 0:053); mean p_{11} = 0:983 (s.e. = 0:021); mean p_{22} = 0:968 (s.e. = 0:031):

Figure 1: In‡ation Rate in Portugal