Editorial Manager(tm) for Computational Economics
Manuscript Draft

Manuscript Number:

Title: STRUCTURAL CHANGE AND LONG MEMORY IN THE DYNAMIC OF U.S. INFLATION PROCESS

Article Type: Original Research

Section/Category:

Keywords: Long memory, model selection, multiple breaks, penalized least-squares estimate, pure structural change.

Corresponding Author: boutahar mohamed,

Corresponding Author's Institution: University

First Author: belkhouja mustapha

Order of Authors: belkhouja mustapha; boutahar mohamed

Manuscript Region of Origin:

Abstract: In this paper, we consider the problem of estimation of break dates in long memory environment. We focus on the estimation of the number of breaks by using the information criteria, the test of Bai and Perron (1998), and the method of Lavielle (2004). We perform a Monte Carlo experiment to analyse the behaviour of such estimators in finite sample size.
An application to the U.S. inflation process is presented to illustrate the usefulness of these procedures.

STRUCTURAL CHANGE AND LONG MEMORY IN THE DYNAMIC OF U.S. INFLATION PROCESS

Mustapha BELKHOUJA*　　**Mohamed BOUTAHAR**†
GREQAM, Université de la Méditerranée 2 Rue de la charité,
13002 Marseille, France.

December 30, 2007

Abstract

In this paper, we consider the problem of estimation of break dates in long memory environment. We focus on the estimation of the number of breaks by using the information criteria, the test of Bai and Perron (1998), and the method of Lavielle (2004). We perform a Monte Carlo experiment to analyse the behaviour of such estimators in finite sample size. An application to the U.S. inflation process is presented to illustrate the usefulness of these procedures.

Key words: Long memory, model selection, multiple breaks, penalized least-squares estimate, pure structural change.
JEL classification: C22, C82.

1 Introduction

In the econometric literature there is a vast amount of works related to structural change. However, the early efforts have been devoted to the detection of a single change-point in the mean of an i.i.d. process (see Hinkley (1970), Sen Srivastava (1975), Hawkins (1977), Bhattacharya (1987), among many others). Recent contributions considered multiple structural changes in a stationary or non-stationary stochastic processes. Liu et al (1997) and Bai and Perron (1998) used the least-squares method allowing serial correlation in the errors.

The problem of estimating the number of changes was studied by Yao (1988), who proved the consistency of the Schwarz criterion in the case of i.i.d Gaussian errors. Bai and Perron addressed this problem by using a supWald type tests and extented the preceding paper to the case of weekly dependent errors. Lavielle (2004) proposed an algorithm to estimate the dimension K which is the number of regimes existing in the model by assuming a strongly dependent errors.

The rest of this paper is organized as follows. Section 2 deals with the model and the assumptions imposed on the variables and the errors. Section 3 presents different procedures

*E-mail: belkhouja_mustapha@yahoo.fr
†E-mail: boutahar@univmed.fr

for the selection of the number of changes in the model such as the information criteria, the test of Bai and Perron (1998) and the method of Lavielle (2004). Section 4 is devoted to simulation experiments, to compare the three methods in estimating the number and the location of the break dates. The application on the U.S. inflation process is reported in section 5.

2 The Model and Estimators

Consider the following structural change model with m breaks ($K = m + 1$ regimes):

$$y_t = \delta_j + \pi_t, \qquad t = T_{j-1} + 1, ..., T_j \tag{2.1}$$

for $j = 1, ..., m + 1$, $T_0 = 0$ and $T_{m+1} = T$, y_t is the observed dependent variable, $\delta = (\delta_1, ..., \delta_{m+1})'$ is the vector of coefficients with $\delta_i \neq \delta_{i+1} (1 \leq i \leq m)$, and π_t is the disturbance. The break dates $(T_1, ..., T_m)$ are treated as unknown for $i = 1, ..., m$, and we have $\lambda_i = T_i/T$ the break fractions with $0 < \lambda_1 < ... < \lambda_m < 1$. We are in the case of a pure structural change model in the sense that all the coefficients shift in time. Bai and Perron (1998), impose some restrictions on the possible values of break dates and they define the following set for some arbitrary small positive number η:

$\Lambda_\eta = \{(\lambda_1, ..., \lambda_m); |\lambda_{i+1} - \lambda_i| \geq \eta, \lambda_1 \geq \eta, \lambda_m \leq 1 - \eta\}$ to be asymptotically distinct and bounded from the boundaries of the sample. The estimation method considered is that based on the least-squares principle proposed by Bai and Perron (1998). This method is described as follows. For each m-partition $(T_1, ..., T_m)$, denoted $\{T_j\}$, the associated least-squares estimates are obtained by minimizing the sum of squared residuals:

$$\sum_{i=1}^{m+1} \sum_{j=T_{i-1}+1}^{T_i} (y_j - \delta_i)^2 .$$

Let $\widehat{\delta}(\{T_j\})$ the resulting estimates, substituting them in the objective function gives an other sum denoted Q_T :

$$\begin{aligned} Q_T(T_1, ..., T_m) &= \min_{(\delta_1, ..., \delta_{m+1})} \sum_{i=1}^{m+1} \sum_{j=T_{i-1}+1}^{T_i} \left(y_j - \widehat{\delta}_i\right)^2 \\ &= \sum_{i=1}^{m+1} \sum_{j=T_{i-1}+1}^{T_i} \left(y_j - \overline{y}(T_{i-1}, T_i)\right)^2 \end{aligned} \tag{2.2}$$

where

$$\overline{y}(T_{i-1}, T_i) = \frac{1}{T_i - T_{i-1}} \sum_{j=T_{i-1}+1}^{T_i} y_j$$

The estimated break dates are as follows:

$$\left(\widehat{T}_1, ..., \widehat{T}_m\right) = \underset{(T_1, ..., T_m)}{\arg\min} Q_T(T_1, ..., T_m), where T_i - T_{i-1} \geq [\eta T]$$

so the break point estimators are thus global minimizers of the objective function. Finally, the estimated regression parameters are the associated least-squares estimates, i.e.

$\widehat{\delta} = \widehat{\delta}\left(\left\{\widehat{T}_i\right\}\right)$. In our computations, we use the efficient algorithm developed in Bai and Perron (2003) based on the principle of dynamic programming. This algorithm allows the obtention of global minimizers using a number of sums of squared residuals of order $O\left(T^2\right)$ for any $m \geq 2$, unlike a standard gird search procedure which would require least squares operations of order $O\left(T^m\right)$. The basis of the method, for specialized cases, is not new and has been part of the statistics literature (see Guthery (1974), Bellman et Roth (1969) and Fisher (1958)), nevertheless, it seems to have been forgotten at least in the econometrics literature despite its high interest.

3 Estimating The Number of Breaks

3.1 The model selection criteria

To determine the dimension of the model, we can use the information criteria. The basic idea is that the addition of a break point necessarily reduces the sum of squared residuals and one must then penalize this addition. So, a heavy penalty reduces the probability of overestimating the number of breaks. In the other hand the penalty must not be exessive because we can underestimate the model in this case. The penalty term differs from a criterion to an other and consequently has an effect on the results of the information criteria as we will see in the simulation experiments.

Yao (1988) suggested the use of the Bayesian information criteria defined as

$$BIC\left(m\right) = \ln\left(Q_T\left(\widehat{T}_1, ..., \widehat{T}_m\right)/T\right) + p^* \ln\left(T\right)/T \tag{3.1}$$

where $p^* = 2m+1$ is the number of unknown parameters. Yao (1988) showed that the change number estimate, $\widehat{m}$, is consistent (at least for normal sequence of random variables with shifts in mean), provided that $m \leq M$, with M be a known upper bound for m.

An other criteria proposed by Yao and Au (1989) is given by

$$YIC\left(m\right) = \ln\left(Q_T\left(\widehat{T}_1, ..., \widehat{T}_m\right)/\left(T-m\right)\right) + mC_T/T \tag{3.2}$$

where $\{C_T\}$ is any sequence satisfying $C_T T^{-2/n} \to \infty$ and $C_T/T \to 0$ as $T \to \infty$ for some positive integer n since the error term is with finite $2nth$ moment for any $n \geqslant 3$. In our simulation experiments, we use the sequence $C_T = 0.368T^{O.7}$ proposed by Liu et al (1997).

Liu et al (1997) suggested a modified Schwarz criteria that takes the form

$$MIC\left(m\right) = \ln\left(Q_T\left(\widehat{T}_1, ..., \widehat{T}_m\right)/\left(T-p^*\right)\right) + p^* c_0 \left[\ln\left(T\right)\right]^{2+\delta_0}/T \tag{3.3}$$

where $\delta_0 = 0.1$ and $c_0 = 0.299$, a choice based on this criterion for various simulation experiments carried out with several models. The estimated number of break dates $\widehat{m}$ is obtained by minimizing the above-mentioned information criteria given the upper bound M for m. Note that the information criteria cannot directly take into account the presence of serial correlation in the errors.

3.2 The method of Bai and Perron (1998)

3.2.1 Test of l versus l+1 breaks

This section considers a test of the null hypothesis of l breaks against the alternative that an additional break exists. For the model with l breaks, the estimated break points, represented by $\widehat{T}_1, ..., \widehat{T}_l$, are obtained by a global minimization of the sum of squared residuals. The strategy of the test proceeds by testing in each $(l+1)$ regime (obtained using the estimated partition $\widehat{T}_1, ..., \widehat{T}_l$) for the presence of an additional break. So, the idea is to carry out $(l+1)$ tests of the null hypothesis of no structural change versus the alternative hypothesis of a single change. This test is applied to each regime containing the observations $\widehat{T}_{i-1}+1$ to $\widehat{T}_i$ $(i = 1, ..., l+1)$ using the convention that $\widehat{T}_0 = 0$ and $\widehat{T}_{l+1} = T$. We conclude for a rejection in favour of a model with $(l+1)$ breaks if the overall minimal value of the sum of squared residuals (overall regimes where an additional break is included) is sufficiently smaller than the sum of squared residuals obtained by the l breaks model. The break date thus selected is the one associated with this overall minimum. For more details, the test is defined as

$$F_T\left(l+1/l\right) = \left\{Q_T\left(\widehat{T}_1, ..., \widehat{T}_l\right) - \min_{1\leq i\leq l+1} \inf_{\lambda\in\Lambda_{i,\eta}} Q_T\left(\widehat{T}_1, ..., \widehat{T}_{i-1}, \lambda, \widehat{T}_i, ..., \widehat{T}_l\right)\right\}/\widehat{\sigma}^2 \quad (3.4)$$

where $\Lambda_{i,\eta} = \left\{\lambda; \widehat{T}_{i-1} + \left(\widehat{T}_i - \widehat{T}_{i-1}\right)\eta \leq \lambda \leq \widehat{T}_i - \left(\widehat{T}_i - \widehat{T}_{i-1}\right)\eta\right\}$ and $\widehat{\sigma}^2$ is a consistent estimator of σ^2 under the null hypothesis. The critical values of this test for different values of l can be obtained from the distribution function $G_{q,\eta}(x)$ (see Bai and Perron (1998), page 60). Using simulations, Bai and Perron (1998), provide critical values in table 1 for $\eta = 0.10$ and $q = 1$.

	l									
$1-\alpha$	0	1	2	3	4	5	6	7	8	9
0.90	7.42	9.05	9.97	10.49	10.91	11.29	11.86	12.26	12.57	12.84
0.95	9.10	10.55	11.36	12.35	12.97	13.45	13.88	14.12	14.45	14.51
0.975	10.56	12.37	13.46	14.13	14.51	14.88	15.37	15.47	15.62	15.79
0.99	13.00	14.51	15.44	15.73	16.39	16.60	16.78	16.90	16.99	17.04

TABLE 1. Asymptotic critical values of the sequential test $F_T\left(l+1/l\right)$

3.2.2 The sequential selection procedure

We start by estimating a model with a small number of breaks that are thought to be necessary (or we start with no break), then we add a break to a sub-sample associated with a rejection using the test $F_T\left(l+1/l\right)$ and we estimate the break dates. This process is repeated by increasing l sequentially until the test fails to reject the null hypothesis of the test. So, the final number of breaks is equal to the number of rejections using the test $F_T\left(l+1/l\right)$ added to the number of breaks used initially and the estimates of the break dates are not guaranteed to be identical to these obtained by global minimization of the sum of squared residuals. The limiting distribution is the same when the break dates are estimated using the global minimization.

3.3 An automatic procedure for estimating the number of regimes $K = m + 1$: Lavielle (2004)

We consider the sum of squared residual:

$$Q_{T,K}\left(T_1,...,T_{K-1}\right) = \min_{(\delta_1,...,\delta_K)} \sum_{i=1}^{K} \sum_{j=T_{i-1}+1}^{T_i} \left(y_j - \widehat{\delta}_i\right)^2 \tag{3.5}$$
$$= \sum_{i=1}^{K} \sum_{j=T_{i-1}+1}^{T_i} \left(y_j - \overline{y}\left(T_{i-1}, T_i\right)\right)^2$$

where $\overline{y}\left(T_{i-1}, T_i\right) = \frac{1}{T_i - T_{i-1}} \sum_{j=T_{i-1}+1}^{T_i} y_j$ and let K_{MAX} be an upper bound for K.

Lavielle (2004) has proposed the following algorithm for estimating the dimension K.

1. For any $1 \leq K \leq K_{MAX}$, let $\widetilde{Q}_{T,K} = \frac{Q_{T,K_{MAX}} - Q_{T,K}}{Q_{T,K_{MAX}} - Q_{T,1}} \left(K_{MAX} - 1\right) + 1$ with $\widetilde{Q}_{T,1} = K_{MAX}$ and $\widetilde{Q}_{T,K_{MAX}} = 1$. This sequence decreases with an average slope equal to -1.

2. For any $2 \leq K \leq K_{MAX} - 1$, let $D_K = \widetilde{Q}_{T,K-1} - 2\widetilde{Q}_{T,K} + \widetilde{Q}_{T,K+1}$ and $D_1 = \infty$, so the estimate number of regimes is given by:

$$\widehat{K} = \max\left\{1 \leq K \leq K_{MAX} - 1 \text{ such as } D_K > S\right\}$$

   $\widehat{K}$ is defined as the greatest value of K such that the second derivative of $Q_{T,K}$ is greater than a given threshold S. If no second derivative is greater than S, we conclude that there are no changes and consquently $\widehat{K} = 1$. Unfortunately, the probability distribution of the statistic D_K cannot be obtained in a closed form and the threshold S cannot be computed as a quantil of this distribution. Nevertheless, on the basis of different numerical experiments, Lavielle proposes $S = 0.75$ because smaller values of S usually over-estimate the number of regimes, while larger values under-estimate this number.

4 Simulation Experiments

Consider the model (2.1) with $m = 2$, $\lambda_1 = 0.25$, $\lambda_2 = 0.5$, $\delta_1 = 2$, $\delta_2 = 0$ and $\delta_3 = 1$. The disturbance π_t is a long memory process given by $\left(1 - L\right)^d \pi_t = \varepsilon_t$ where L is the lag operator such as $L\pi_t = \pi_{t-1}$ and $\varepsilon_t \rightsquigarrow N\left(0, 1\right)$. We take $d < 0.5$, so π_t is a covariance stationary process with zero-mean whose spectral density verifies

$$f\left(\omega\right) \approx C\omega^{-2d} \; as\omega \rightarrow 0^+ \tag{4.1}$$

where d is the long-range dependence parameter and C is a constant.

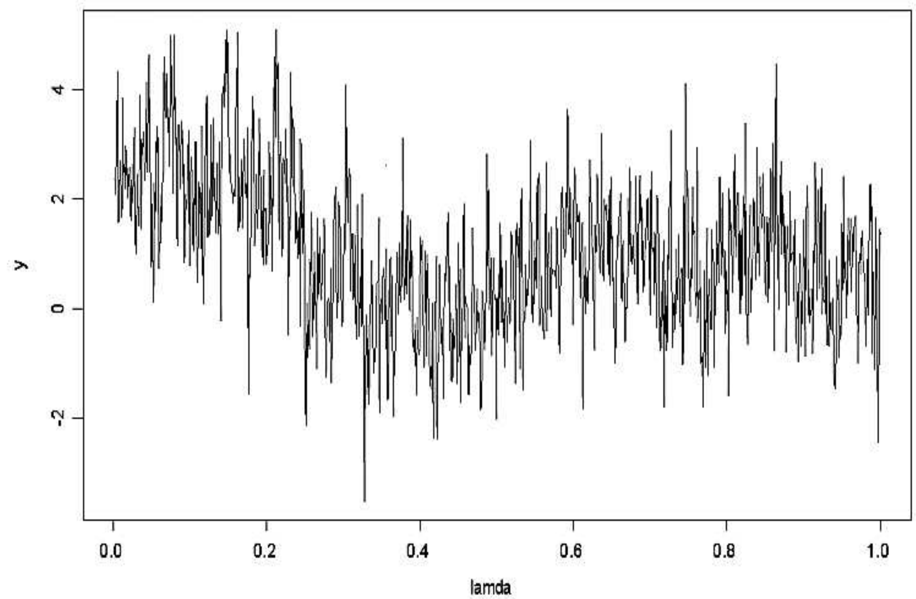


FIGURE 1. A realization of (y_t), $T = 1000$, $d = 0.30$ and with break dates at $\lambda_1 = 0.25$ and $\lambda_2 = 0.5$

The minimal number of observations in each regime is equal to $[\eta T]$ with $\eta = 0.10$.

The maximum permitted number of breaks is $M = 7$. We consider three sample sizes $T = \{500, 1000, 2000\}$ and five values for the long-range dependence parameter $d = \{0.10; 0.20; 0.30; 0.40; 0.49\}$. All the reported simulation results are based on 100 replications and are given in the Appendix.

Experience 1: Analysis of the information criteria

In this experiment, we estimate a model with change in mean using the information criteria BIC, MIC and YIC. The results are given in tables 2, 3 and 4 for different values of d and T. The results are practically similar for the information criteria i.e. they choose a higher number of breaks as d increases because we are close to the non-stationary case. These three tables lead us to conclude that for all the information criteria, the frequency of choosing $\widehat{m} = 2$ is very reasonable for a small values of d, but it decreases as d approches to 0.49. On the other hand, the frequency of choosing $\widehat{m} > 2$ increases as T increases. If we compare the three information criteria, we can say that the results given by MIC and YIC are better than those obtained by BIC, hence, a heavy penalty reduces the tendency of choosing a higher number of breaks. These conclusions are predictable, since as we said, the information criteria cannot directly take into account the presence of serial correlation in the errors.

Experience 2: Analysis of the test $F_T\,(l + 1/l)$: Bai and Perron (1998)

With this type of tests, we have the same problem with respect to the serial correlation in the errors. The results are provided in table 5 for different choices of d and T with the significance level $\alpha = 0.05$. When d and T increase, we notice that the frequency of choosing $\widehat{m} = 2$ decreases and becomes practically equal to zero. When we are close to the non-stationary case, the highest frequency corresponds to $\widehat{m} > 2$ like the information criteria. In conclusion, we can say that these results are not better than the ones obtained with the information criteria.

Experience 3: Analysis of the automatic procedure: Lavielle(2004)

The results are presented in table 6 for different values of d and T. When d increases, the frequency of choosing $\widehat{m} = 2$ decreases but remains the highest in the distribution even for $d = 0.49$. When T increases, we see that the frequency of choosing $\widehat{m} = 2$ increases practically whatever the value of d is, which shows the difference and the advantage of this method over the latters. On the basis of the simulation experiments, we can say that the best method to estimate the number of breaks is the Lavielle's, because its frequency of choosing the true number of changes is the highest particularly when the difference level converges to 0.5.

Experience 4: Estimating the break points

We estimate the break points by using the efficient algorithm developed in Bai and Perron (2003) after constructing the triangular matrix for $m = 2$ and $\eta = 0.10$. The histograms of the estimated break dates $\widehat{\lambda}_1$ and $\widehat{\lambda}_2$ are given in figures 2, 3, 4, 5, and 6, for different values of d and T. We notice that the first break point is estimated with more precision since the frequency of detecting the first true break point is higher than the one detecting the second true break point and the distribution tail of $\widehat{\lambda}_1$ is thicker than that of $\widehat{\lambda}_2$ whatever the value of d is. This can be explained by the fact that the first break point has a larger magnitude of shift than the second. We also see, for a fixed T, that the distribution tails of the two estimated break points become thicker as d varies from 0.10 to 0.49 and the contrary happens as $T \to \infty$ for a fixed d. Thus, we can conclude that the presence of long memory affects the estimation precision of the break dates.

5 Empirical Applications

We use the post war monthly U.S. inflation rate (seasonally adjusted) covering the period 1955:1-2006:10 (see Figure 7). Because of the presence of abrupt structural changes in the mean of the series, we can consider a structural change model with only a constant as regressor. We then determine the existing break dates using the above techniques providing a trimming $\mu = 0.10$ and allowing up to 7 breaks. A serial correlation in the errors and different variances across regimes are permitted. The results reported in table 7 show that BIC selects 4 breaks, MIC and YIC select 3 breaks, the sequential procedure (using a 5% significance level) selects 4 breaks while the Lavielle's method selects only 2 breaks. Given the simulation results above, the Lavielle's method performs better than the others and so we conclude in favour of the presence of two breaks. The break dates are estimated at 1973:1 and 1981:9 which coincide with the period stretching from 1973 Oil-Price shock and the change in the Federal Reserve's operating procedures in 1979. The second break date has a larger confidence interval than the first (see Bai, J (1997), page 555) and the differences in the estimated means over each regime are significant.

A look at the autocorrelation function of the original series suggests the presence of long memory because there is a slow decay in the values and this is confirmed by the periodogram because we observe a peak at the zero frequency. This behaviour can be explained by the real existence of long memory and or the presence of shifts in the series.

Now, the issue to be considered is to model the U.S. inflation rate and then we will execute some misspecification tests on the residuals, to support the fitted model, such as the Ljung and Box (1978) Q statistic for autocorrelation, the ARCH test (Engle, 1982) for autoregressive conditional heteroskedasticity and finally, the Kolmogorov-Smirnov (KS) statistic to test for normality.

5.1 The proposed model

First, we must consider the presence of structural change in the series (see, Figure 7), to that effect we propose the following model which takes into account the presence of two breaks 1973:1 and 1981:9, selected above, using the structural change model:

$$\pi_t = \begin{cases} y_t - \widehat{\delta}_1, \text{ if } t \in [1955:1, 1972:12], \\ y_t - \widehat{\delta}_2, \text{ if } t \in [1973:1, 1981:8], \\ y_t - \widehat{\delta}_3, \text{ if } t \in [1981:9, 2006:10], \end{cases} \tag{5.1}$$

where π_t is the series of the generalized residuals and $\widehat{\delta}_1$, $\widehat{\delta}_2$ and $\widehat{\delta}_3$ are respectively the estimated means over the three regimes. From the graphics of the generalized residuals series π_t, we observe that the long memory is still present but it is less important than the original series (see, Figure 8). We now test for stationarity using the augmented Dickey-Fuller (ADF) (Said and Dickey 1984), the Phillips and Perron (1988) (PP) and the KPSS (Kwiatkowski et al., 1992). We introduce a linear time trend to capture the deterministic nonstationarity and the table 8 shows evidence for stationarity hypothesis with those three tests applied on π_t.

5.2 Fitting a FARIMA model

The long memory behavior can be captured by a fractionally integrated $I(d)$ model which may be presented as

$$(1-L)^d \pi_t = \varepsilon_t \tag{5.2}$$

where ε_t is a covariance stationary process with a bounded spectral density on $[0, \Pi]$.

We use the semiparametric estimator of Geweke-Porter-Hudak (1983) to estimate the degree of integration d. Table 9 gives different estimates of d for different regression ranges $g(T)$ and the one selected is $\widehat{d} = 0.32$ since it corresponds to the smallest standard error. In empirical computations, the $(1-L)^{\widehat{d}}$ filter is approximated as follows

$$(1-L)^{\widehat{d}} \simeq \sum_{j=0}^{n} \widehat{\gamma}_j(\widehat{d}) L^j \;\; , \; \widehat{\gamma}_j(\widehat{d}) = \frac{j-1-\widehat{d}}{j} \widehat{\gamma}_{j-1}(\widehat{d}) \; , \; \widehat{\gamma}_0(\widehat{d}) = 1 \tag{5.3}$$

Following Hassler and Wolters (1995), we may choose n such that the following condition holds:

$$\left| \widehat{\gamma}_{n-1}(\widehat{d}) \right| \geq 0.005, \quad \left| \widehat{\gamma}_n(\widehat{d}) \right| < 0.005 \tag{5.4}$$

For our data we obtain $n = 19$, hence the approximation of the process ε_t is given by:

$$\widehat{\varepsilon}_t \simeq \sum_{j=0}^{19} \widehat{\gamma}_j(0.32)\pi_{t-j}, \qquad t = 20, 21, ..., 622 \tag{5.5}$$

The misspecification tests applied on the filtred series ε_t reveal that the P-values from Ljung-Box, ARCH and KS test are null, hence we must find a model which is more adequate (see Table 10). When analyzing the autocorrelation function of the filtred series (Figure 9), we notice two peaks at the 10th and the 11th lag, which can be removed by using the filter $(1 - \theta_{10}B^{10} - \theta_{11}B^{11})$, where $\theta_{10} = 0.10$ and $\theta_{11} = 0.16$ are respectively the partial correlation coefficients of the 10th and the 11th lagged variables. To eliminate the short memory, we fit an ARMA model. Concerning the heteroskedasticity, we can take as basis the autocorrelation function of the squared filtered series (Figure 9), and since we have two peaks at the first and the second lag, we can conclude in favour of an ARCH(2) model. In conclusion, we have the following FARIMA model with ARCH(2) disturbances:

$$\begin{cases} (1 - 0.78B + 0.34B^2)(1 - 0.10B^{10} - 0.16B^{11})\,(1-d)^{0.32}\,\pi_t = (1 + 0.82B - 0.27B^2)\mu_t \\ \mu_t = \sigma_t z_t \text{ with } z_t \rightsquigarrow iid\ N\,(0,1) \text{ and } \sigma_t^2 = 0.67 + 0.17\mu_{t-1}^2 + 0.15\mu_{t-2}^2 \end{cases} \tag{5.6}$$

To check the validity of the proposed model, the same misspecification tests are applied on the residuals and from the Table 10 we can see that there is no autocorrelation nor heteroskedasticity while the data are still non-normal.

6 Conclusion

We have proposed an hybrid model combining structural change and long memory behaviour. We have also reviewed some techniques to estimate break dates and some model selection procedures as the information criteria, the sequential method and the Lavielle's method to select the number of breaks. Finally, we have illustrated the usefulness of these procedures through an empirical application to the post war monthly U.S. inflation rate. Indeed, we have used the Lavielle's method since it performs better than the others, which are affected by the presence of long memory, according to the simulation experiments. The results obtained are significant since the break date estimates coincide with the period stretching from 1973 Oil-Price shock and the change in the Federal Reserve's operating procedures in 1979. The fitted model for the U.S. inflation rate is a FARIMA model with two shifts in mean and heteroskedastic errors.

References

[28] [1] Andrews DKW. (1991). "Heteroscedasticity and autocorrelation consistent covariance matrix estimation". *Econometrica,* 59:817-858.

[28] [2] Bai, J. (1994a). "Least squares estimation of a shift in linear processes". *Journal of Time Series Analysis,* 15: 453-472.

[28] [3] Bai, J. (1997). "Estimation of a Change Point in Multiple Regression Models". *The Review of Economics and Statistics,* 79: 551-563.

[28] [4] Bai, J. (1997b). "Estimating multiple breaks one at a time". *Econometric Theory,* 13:315-352.

[28] [5] Bai, J. and P. Perron (1998). "Estimating and testing linear models with multiple structural changes". *Econometrica,* 66:47-78.

[28] [6] Bai, J. and P. Perron (2003). "Computation and analysis of multiple structural change models". *Journal of Applied Econometrics,* 18:1-22.

[28] [7] Berran, J. (1994). "A Goodness-of-fit test for time series with long range dependence". *Journal of The Royal Statical Society, Ser. B,* 54:749-760.

[28] [8] Berran, J. (1994). *Statistics for long memory processes.* London : Chapman and Hall.

[28] [9] Boutahar, M. and J. Jouini (2004). "Structural change in the long-memory environment". *Document de travail* $n°2004 - 42$.

[28] [10]Boutahar, M. and J. Jouini (1996). "Spuriousness of information creteria in selecting the number of changes". *Document de travail* $n°03A17$.

[28] [11] Diebold, F. X., and A. Inoue (2001). "Long memory and regime switching". *Journal of Econometrics* , 105, 131-159.

[28] [12] Engle, R. F. And A. D. Smith (1999). "Stochastic permanent breaks". *Review of Economics and Statistics,* 81:553-574.

[28] [13] Fox , R. and M. S. Taqqu (1986). "Large-sample properties of parameter estimates for strongly dependent stationary Gaussian time series". *Annals of Statistics,* 14:517-532.

[28] [14] Geweke, J. and S. Porter-Hudak (1983). "The estimation and application of long memory time series models". *Journal of Time Series Analysis* , 4:15-39.

[28] [15] Giraitis, L. and D. Surgailis (1990). "A central limit theorem for quadratic forms in strongly dependent linear variables and its application to asymptotic normality of Whittle's estimate". *Probability Theory and Related Fields,* 86: 87-104.

[28] [16] Granger , C. W. J. and N. Hyung (1999). "Occasional structural breaks and long memory". *Disscussion paper 99-14, University of California, San Diego.*

[28] [17] Grenander, U. (1954). "On estimation of regression coefficients in the case of autocorrelated disturbance". *Annals of Mathematical Statistics,* 25:252-272.

[28] [18] Hassler, U., and J. "Wolters (1995). Long Memory in Inflations Rates: International Evidence". *Journal of Business and Economic Statistics,* 13:37-45.

[28] [19] Hwang, S. (1996). "Structural breaks in AR(1) process and persistence". *Manuscript, Faculty of Finance, Cass Business School.*

[28] [20] Lavielle, M. and E. "Moulines (1996). Least squares estimations of an unknown number of shifts in a time series". *Journal of Time Series Analysis. 21, No. 1.*

[28] [21] Lavielle, M. (1999). "Detection of multiple changes in a sequence of dependent variables, stoch". *Proc. Appl.* 83:79-102.

[28] [22] Lavielle, M. (2004). "Using penalized contrasts for the change-point problem". *Elsevier.*

[28] [23] Perron, P. (1997). "Estimation de modèles avec changements structurels multiples". *Actualité Economique,* 73:457-505.

[28] [24] Schwarz, G. (1987). "Estimating the dimension of a model". *Annals of Statistics,* 6:461-464.

[28] [25] Yajima , Y. (1988). "On estimation of regression model with long-memory stationary errors". *Annals of Statistics,* 16:791-807.

[28] [26] Yajima , Y. (1988). "Asymptotic properties of the LSE in a regression model with long-memory stationary errors". *Annals of Statistics,* 19:158-177.

[28] [27] Yao, Y.C. (1988). "Estimating the number of break dates via schwarz's criteria". *Statistics and Probability Letters,* 6:181-189.

[28] [28] Yao, Y.C, and S. T. Au (1989). "Least squares estimations of a step function". *Sankhya, Ser. A,* 51:370-381.

APPENDIX: SIMULATION RESULTS

TABLE 2. Percentage of selecting breaks with different values of d, T=500

	BIC					MIC					YIC				
d / $\widehat{m}$	0.10	0.20	0.30	0.40	0.49	0.10	0.20	0.30	0.40	0.49	0.10	0.20	0.30	0.40	0.49
0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
1	0.00	0.00	0.03	0.00	0.00	0.01	0.02	0.14	0.11	0.05	0.01	0.02	0.14	0.10	0.05
2	0.90	0.47	0.16	0.11	0.04	0.99	0.95	0.69	0.42	0.21	0.99	0.93	0.68	0.40	0.19
3	0.07	0.39	0.36	0.18	0.11	0.00	0.03	0.17	0.36	0.35	0.00	0.05	0.17	0.38	0.35
4	0.03	0.09	0.33	0.32	0.21	0.00	0.00	0.00	0.08	0.21	0.00	0.00	0.01	0.09	0.21
5	0.01	0.05	0.08	0.25	0.28	0.00	0.00	0.01	0.03	0.15	0.00	0.00	0.00	0.03	0.16
6	0.00	0.00	0.04	0.13	0.28	0.00	0.00	0.00	0.00	0.03	0.00	0.00	0.00	0.00	0.04
7	0.00	0.00	0.00	0.01	0.08	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

TABLE 3: Percentage of selecting breaks with different values of d, T=1000

	BIC					MIC					YIC				
d / $\widehat{m}$	0.10	0.20	0.30	0.40	0.49	0.10	0.20	0.30	0.40	0.49	0.10	0.20	0.30	0.40	0.49
0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.03	0.04	0.01	0.00	0.01	0.07	0.06	0.03
2	0.80	0.38	0.06	0.04	0.01	1.00	0.94	0.58	0.22	0.14	1.00	0.97	0.65	0.36	0.17
3	0.19	0.37	0.24	0.01	0.02	0.00	0.06	0.32	0.39	0.19	0.00	0.02	0.26	0.39	0.27
4	0.01	0.17	0.32	0.23	0.14	0.00	0.00	0.05	0.18	0.43	0.00	0.00	0.02	0.15	0.39
5	0.00	0.07	0.27	0.37	0.30	0.00	0.00	0.02	0.16	0.18	0.00	0.00	0.00	0.04	0.13
6	0.00	0.01	0.10	0.28	0.34	0.00	0.00	0.00	0.01	0.05	0.00	0.00	0.00	0.00	0.01
7	0.00	0.00	0.01	0.08	0.19	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

TABLE 4: Percentage of selecting breaks with different values of d, T=2000

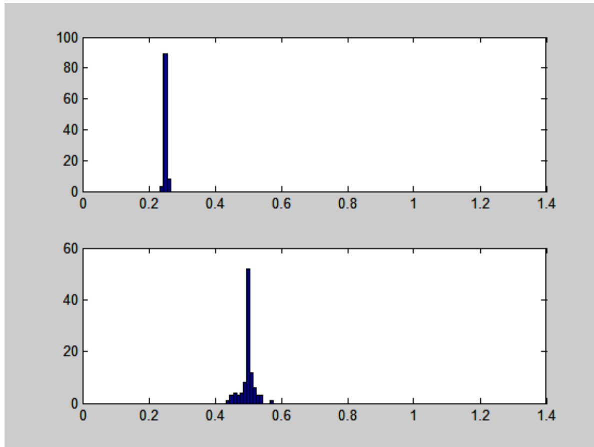
	BIC					MIC					YIC				
d / $\widehat{m}$	0.10	0.20	0.30	0.40	0.49	0.10	0.20	0.30	0.40	0.49	0.10	0.20	0.30	0.40	0.49
0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.02	0.00
1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.02	0.06	0.02
2	0.76	0.26	0.02	0.00	0.00	1.00	0.92	0.46	0.11	0.04	1.00	0.99	0.83	0.30	0.16
3	0.19	0.35	0.15	0.05	0.00	0.00	0.08	0.39	0.23	0.19	0.00	0.01	0.13	0.36	0.37
4	0.04	0.25	0.21	0.07	0.08	0.00	0.00	0.13	0.39	0.27	0.00	0.00	0.01	0.22	0.24
5	0.01	0.14	0.41	0.40	0.21	0.00	0.00	0.01	0.20	0.24	0.00	0.00	0.00	0.04	0.15
6	0.00	0.00	0.18	0.32	0.38	0.00	0.00	0.01	0.04	0.20	0.00	0.00	0.00	0.00	0.06
7	0.00	0.00	0.03	0.16	0.33	0.00	0.00	0.00	0.02	0.05	0.00	0.00	0.00	0.00	0.00

TABLE 5. Percentage of selecting breaks with different values of d and T, $\alpha = 5\%$

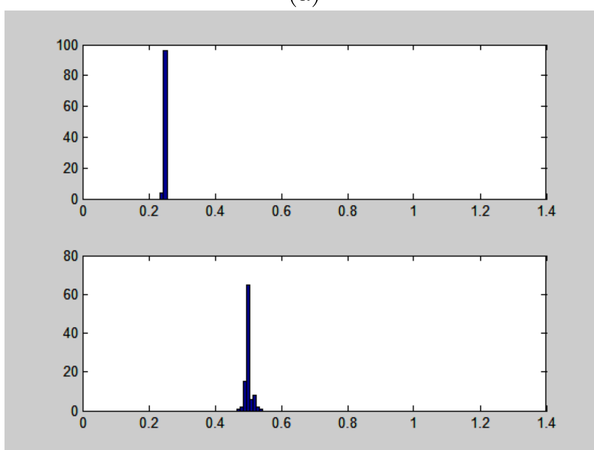
T	500					1000					2000				
d / $\widehat{m}$	0.10	0.20	0.30	0.40	0.49	0.10	0.20	0.30	0.40	0.49	0.10	0.20	0.30	0.40	0.49
0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
1	0.00	0.00	0.02	0.01	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00
2	0.92	0.48	0.22	0.12	0.05	0.74	0.31	0.09	0.04	0.01	0.63	0.18	0.04	0.01	0.00
3	0.08	0.42	0.41	0.30	0.22	0.25	0.45	0.28	0.06	0.05	0.32	0.33	0.10	0.04	0.01
4	0.00	0.09	0.30	0.39	0.24	0.01	0.19	0.38	0.29	0.26	0.04	0.34	0.31	0.15	0.09
5	0.00	0.01	0.04	0.13	0.30	0.00	0.05	0.17	0.39	0.34	0.01	0.14	0.38	0.40	0.32
6	0.00	0.00	0.01	0.05	0.17	0.00	0.00	0.07	0.19	0.27	0.00	0.01	0.15	0.30	0.41
7	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.01	0.03	0.06	0.00	0.00	0.02	0.10	0.17

TABLE 6: Percentage of selecting breaks with different values of d and T

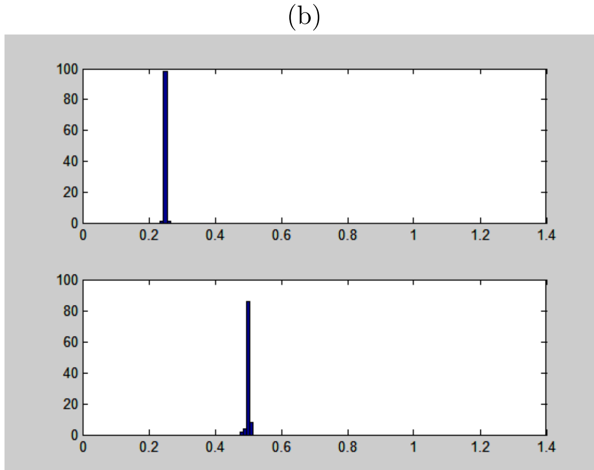
T	500					1000					2000				
d / $\widehat{m}$	0.10	0.20	0.30	0.40	0.49	0.10	0.20	0.30	0.40	0.49	0.10	0.20	0.30	0.40	0.49
0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00
1	0.01	0.09	0.16	0.22	0.22	0.00	0.04	0.15	0.20	0.22	0.00	0.01	0.11	0.13	0.20
2	0.99	0.88	0.69	0.52	0.42	1.00	0.95	0.71	0.52	0.40	1.00	0.99	0.85	0.56	0.42
3	0.00	0.01	0.11	0.21	0.21	0.00	0.01	0.09	0.19	0.17	0.00	0.00	0.03	0.21	0.24
4	0.00	0.00	0.02	0.01	0.07	0.00	0.00	0.01	0.03	0.15	0.00	0.00	0.00	0.01	0.06
5	0.00	0.00	0.00	0.02	0.01	0.00	0.00	0.00	0.02	0.02	0.00	0.00	0.00	0.02	0.01
6	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
7	0.00	0.02	0.02	0.02	0.06	0.00	0.00	0.04	0.04	0.03	0.00	0.00	0.01	0.07	0.07



(a)

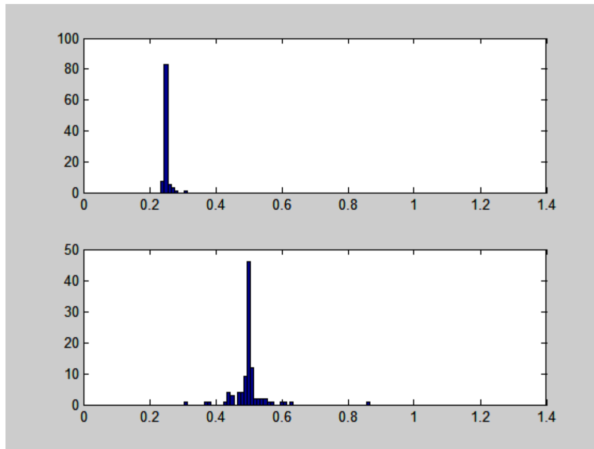


(b)

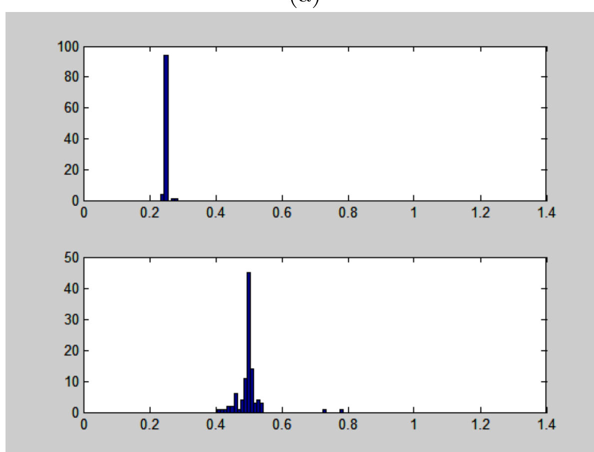


(c)

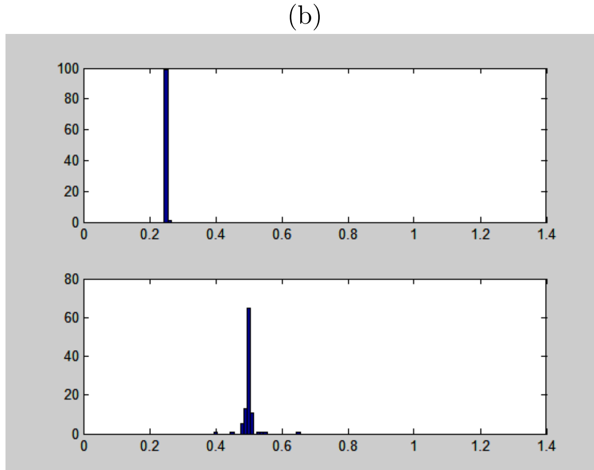
FIGURE 2. Histograms of estimated break dates, d=0.10, (a):T=500, (b):T=1000 and (c):T=2000



(a)

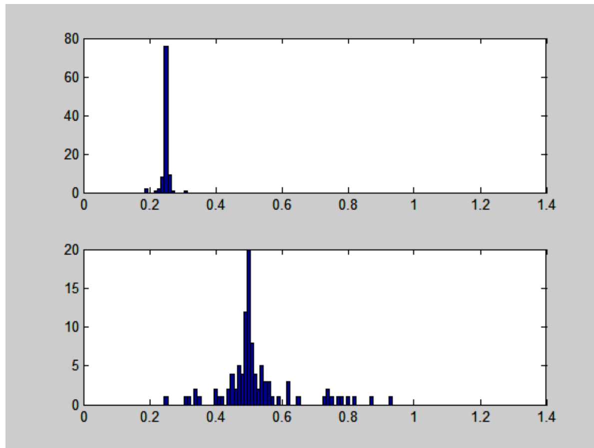


(b)

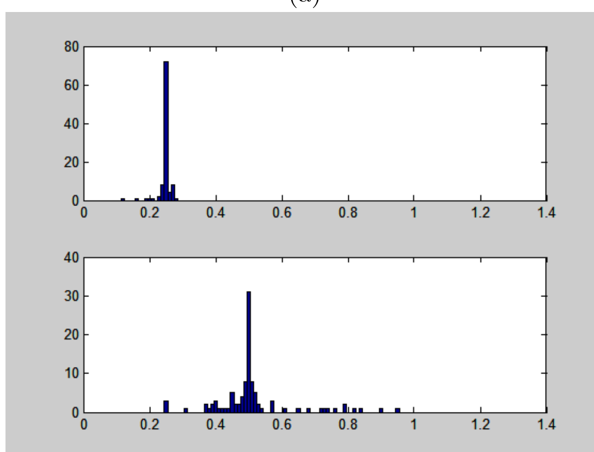


(c)

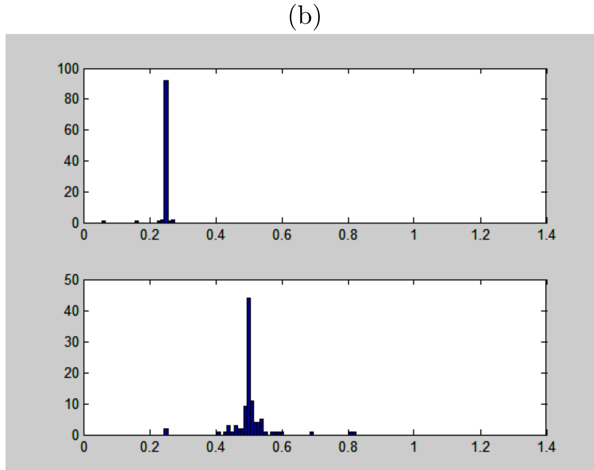
FIGURE 3. Histograms of estimated break dates, d=0.20, (a):T=500, (b):T=1000 and (c):T=2000



(a)

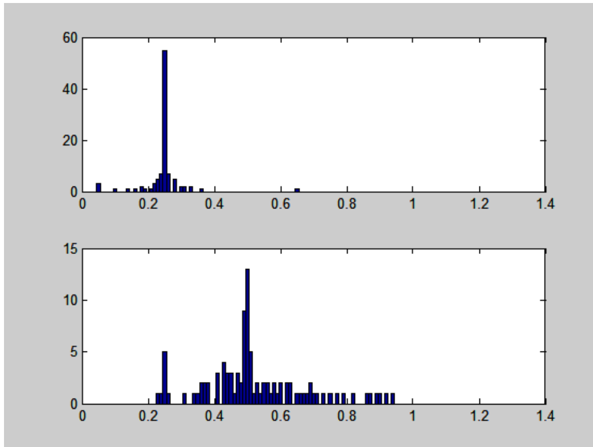


(b)

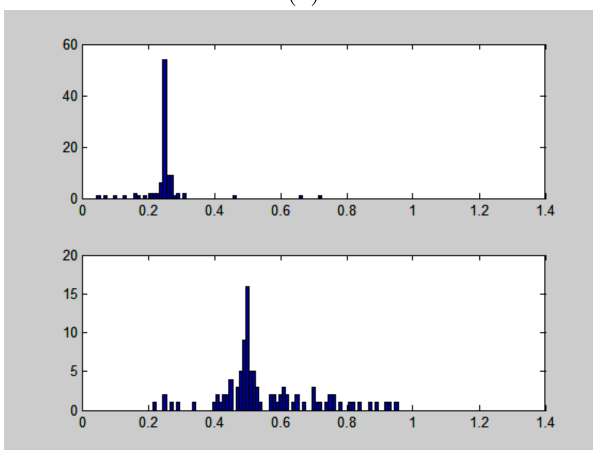


(c)

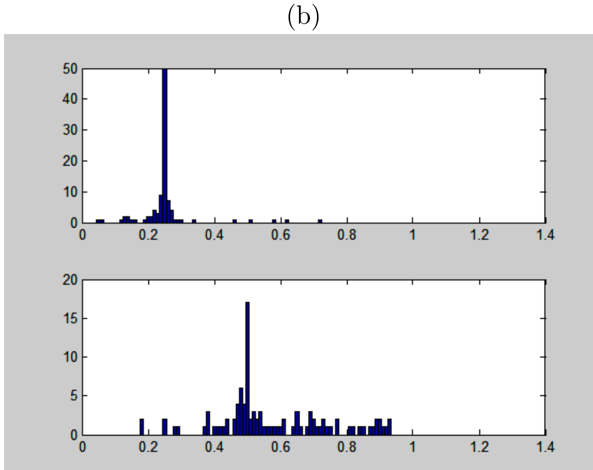
FIGURE 4. Histograms of estimated break dates, d=0.30,
(a):T=500, (b):T=1000 and (c):T=2000



(a)

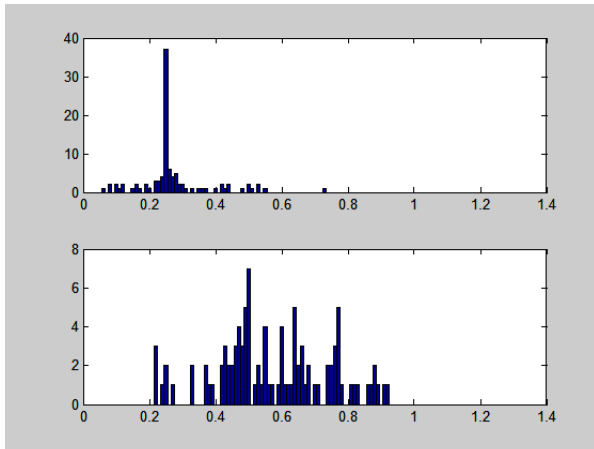


(b)

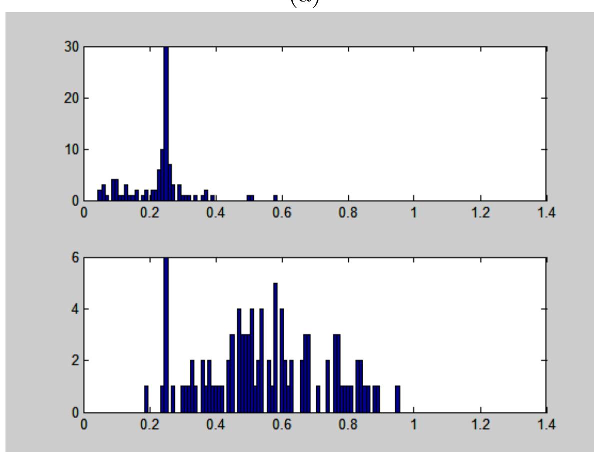


(c)

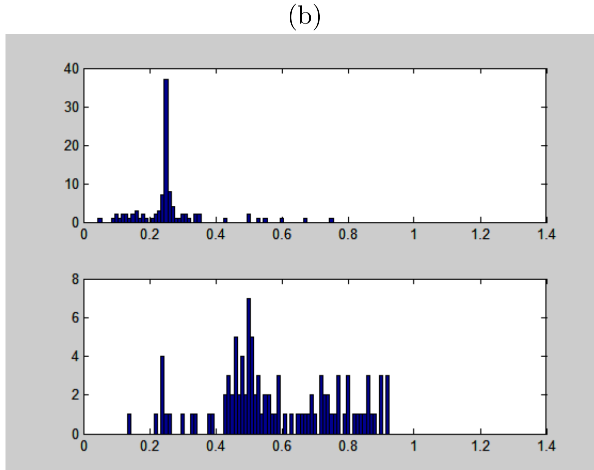
FIGURE 5. Histograms of estimated break dates, d=0.40, (a):T=500, (b):T=1000 and (c):T=2000



(a)



(b)



(c)

FIGURE 6. Histograms of estimated break dates, d=0.49, (a):T=500, (b):T=1000 and (c):T=2000

EMPIRICAL RESULTS

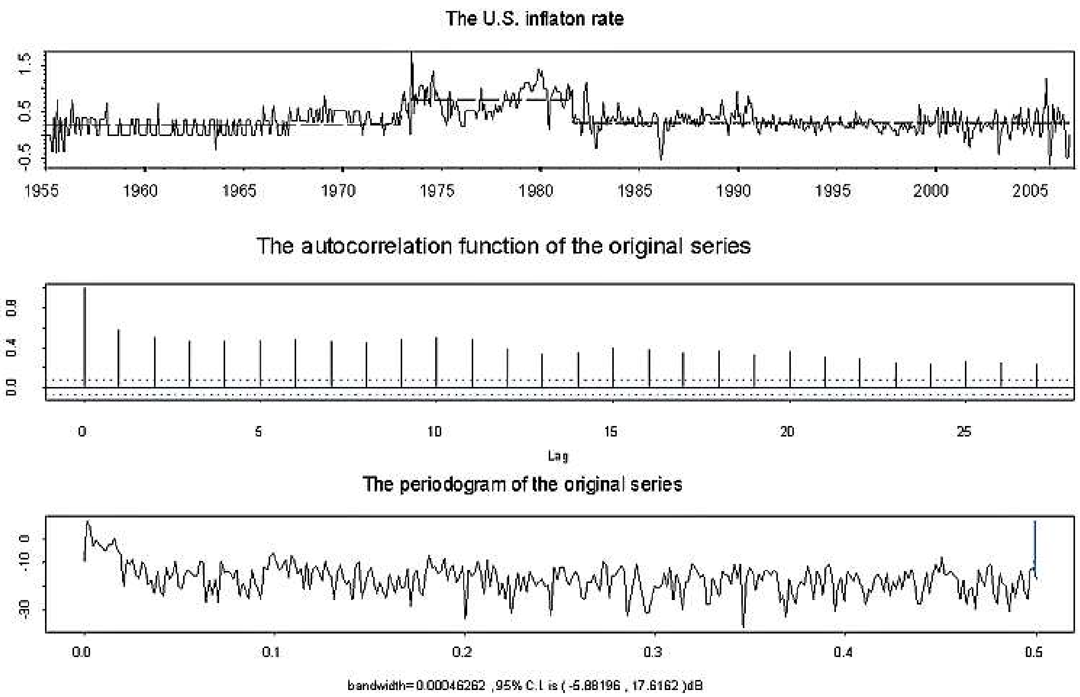


FIGURE 7. Graphics of the original series

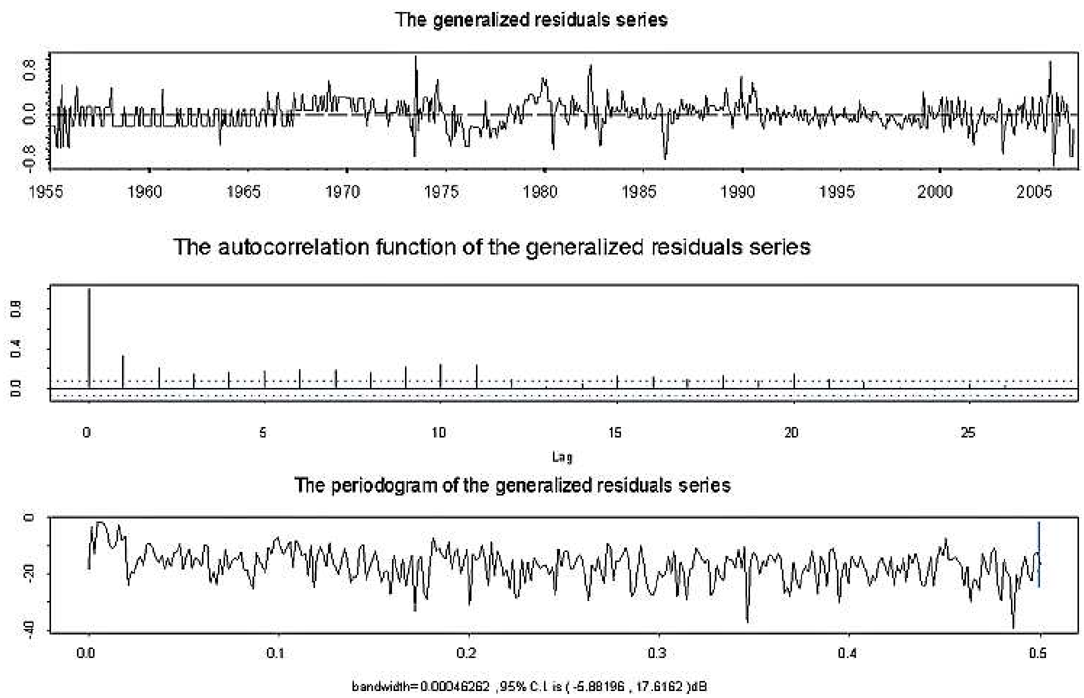


FIGURE 8. Graphics of the generalized residuals series

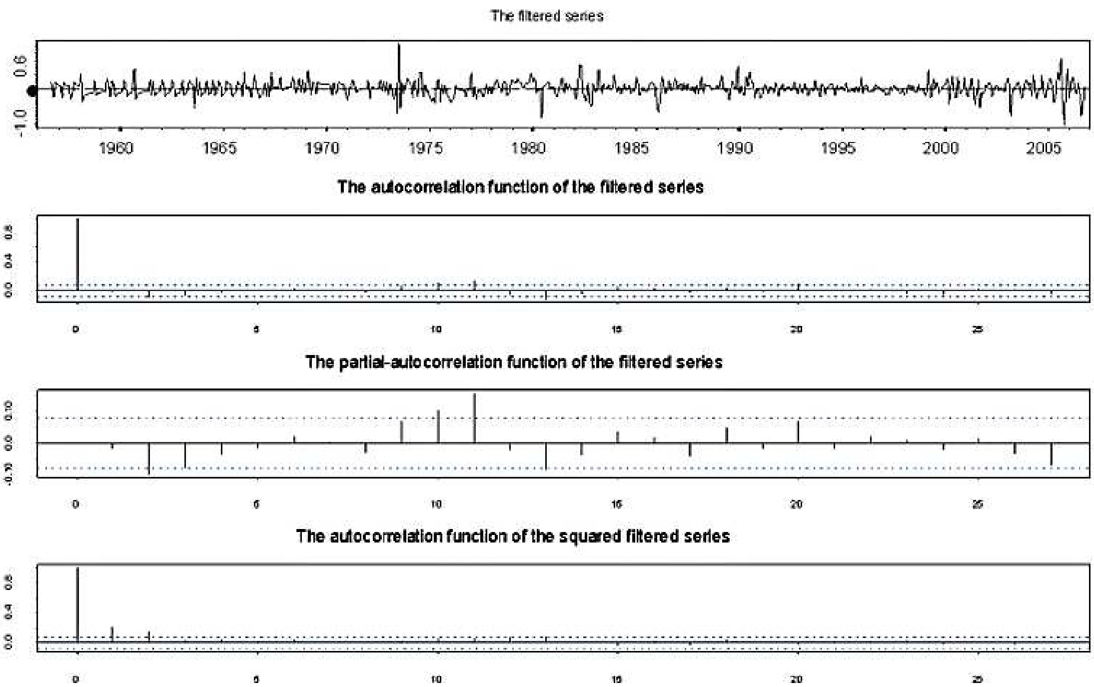


FIGURE 9. Graphics of the filtered series

TABLE 7. U.S. inflation rate(1955:1-2006:10)

			specifications	
	q=1	p=1	M=7	$\mu = 0.10$
			Tests	
$F(1\,\|0)$	$F(2\,\|1)$	$F(3\,\|2)$	$F(4\,\|3)$	$F(5\,\|4)$
67.98*	155.87*	42.84*	17.22*	0.92
			Number of breaks selected[1]	
BIC	4			
YIC	3			
MIC	3			
Sequential	4			
Lavielle	2			
			Estimates with two breaks[2]	
$\widehat{\delta}_1$	$\widehat{\delta}_2$	$\widehat{\delta}_3$	$\widehat{T}_1$	$\widehat{T}_2$
0.22	0.75	0.26	73:1	81:9
(0.05)	(0.09)	(0.05)	(71:8-73:4)	(81:4-83:5)

Notes:

[1] We use a 5% size for the sequential test $F_t(l+1\,|l\,)$.

[2] The values in parenthesis for $\widehat{\delta}_i$ are the standard errors and those for $\widehat{T}_i$ are the 95% confidence intervals.

* Significance at level 5% level.

TABLE 8. Stationarity tests for the generalized residuals series

	k	ADF	PP	KPSS
π_t	3	-10.53 (1.222e-19)	-18.44 (1.979e-37)	0.3192
	6	-6.526 (1.015e-8)		
	12	-4.133 (9.259e-4)		

Notes:

* For the ADF test k represents the number of lagged differences.

* The null hypothesis for ADF and KPSS test is non stationarity, while it is stationarity for the PP test.

*P-values are provided in parenthesis.

*For the KPSS test, the critical values are 0.762(1%), 0.446(5%) and 0.349(10%).

TABLE 9. Estimates of $\widehat{d}$ for the generalized residuals

	$g(T)$	$\widehat{d}$	SE
	$T^{1/2}$	0.488969	0.1611329
	40	0.4482	0.1178005
π_t	80	0.3272	0.0792373
	120	0.3230276	0.06377362
	160	0.3168615	0.05533418
	200	0.322737	0.04967235

Note: SE denotes approximate standard error.

TABLE 10. misspecification

		$I(d)$ process	FARIMA process with Gaussian noise	FARIMA process with ARCH noise
tests	Q(12) [P value]	33.1390 [0.0009]	6.9092 [0.8636]	9.8753 [0.6269]
	ARCH(12) [P value]	40.0507 [0.0001]	44.949 [0.0000]	6.7390 [0.8744]
	KS [P value]	0.0693 [0.0000]	0.0693 [0.0000]	0.0525 [0.0005]