A Generalized ARFIMA Process with Markov-Switching Fractional Differencing Parameter

Wen-Jen Tsay*
Wolfgang Härdle**

* Academia Sinica, Taiwan
** Humboldt-Universität zu Berlin, Germany

ECONOMIC RISK   BERLIN

SFB 649

A Generalized ARFIMA Process with Markov-Switching Fractional Differencing Parameter

Wen-Jen Tsay [1] and Wolfgang Karl Härdle [2] *

April 24, 2007

[1] The Institute of Economics, Academia Sinica, Taiwan

[2] CASE – Center for Applied Statistics and Economics

Humboldt-Universität zu Berlin,

Spandauer Straße 1, 10178 Berlin, Germany

Abstract

We propose a general class of Markov-switching-ARFIMA processes in order to combine strands of long memory and Markov-switching literature. Although the coverage of this class of models is broad, we show that these models can be easily estimated with the DLV algorithm proposed. This algorithm combines the Durbin-Levinson and Viterbi procedures. A Monte Carlo experiment reveals that the finite sample performance of the proposed algorithm for a simple mixture model of Markov-switching mean and ARFIMA$(1, d, 1)$ process is satisfactory. We apply the Markov-switching-ARFIMA models to the U.S. real interest rates, the Nile river level, and the U.S. unemployment rates, respectively. The results are all highly consistent with the conjectures made or empirical results found in the literature. Particularly, we confirm the conjecture in Beran and Terrin (1996) that the observations 1 to about 100 of the Nile river data seem to be more independent than the subsequent observations, and the value of differencing parameter is lower for the first 100 observations than for the subsequent data.

Key words: Markov chain; ARFIMA process; Viterbi algorithm; Long memory

JEL classification: C14, C22, C32, C52, C53, G12

1   Introduction

It is well known that many time series data exhibit long memory, or long-range dependence, including the Nile river level, *ex post* real interest rate, forward premium, and the dynamics of aggregate partisanship and macroideology. Among the many other examples that Beran (1994) gives the Nile river data has been known for its long memory behavior since ancient times, and this is one of the time series that led to the discovery of the Hurst effect (Hurst, 1951) and motivated Mandelbrot and his co-workers (Mandelbrot and van Ness, 1968; Mandelbrot and Wallis, 1969) to introduce fractional Gaussian noise to model long memory phenomenon.

Long range dependence also has been observed in financial data. As demonstrated by Ding et al. (1993), de Lima and Crato (1993) and Bollerslev and Mikkelsen (1996) that the volatility of most financial time series exhibits strong persistency and can be well described as a long memory process. Evidence of financial market volatility's strong persistency inspired Breidt et al. (1998) to propose a class of long memory stochastic volatility (LMSV) models. Deo et al. (2006) also show that the LMSV model is useful for forecasting realized volatility (RV) which is an important quantity in finance.

Figure 1 displays the yearly Nile river minima based on measurements at the Roda gauge near Cairo during the years 622-1284. Beran (1994, p.33) documents that "When one only looks at short time periods, then there seem to be cycles or local trend. However, looking at the whole series, there is no apparent persisting cycle." The changing pattern of the Nile river data leads Bhattacharya et al. (1983) to argue that the so-called Hurst effect can also be explained as if the observations are composed as the sum of a weakly dependent stationary process and a deterministic function. As a consequence it is important to distinguish between a long memory time series and a weakly dependent time series with change-points in the mean. This question has been intensively considered in the literature, including Künsch (1986) and Heyde and Dai (1996). Berkes et al. (2006) presents an overview about this strand of literature. Similarly, Diebold and Inoue (2001) shows that long memory also may be easily confused with a Markov-switching mean. Thus, most of the existing literature considers long memory as a competing modeling framework against the structural change and Markov-switching models.

The Nile river level time series is far more complicated than a pure long memory or
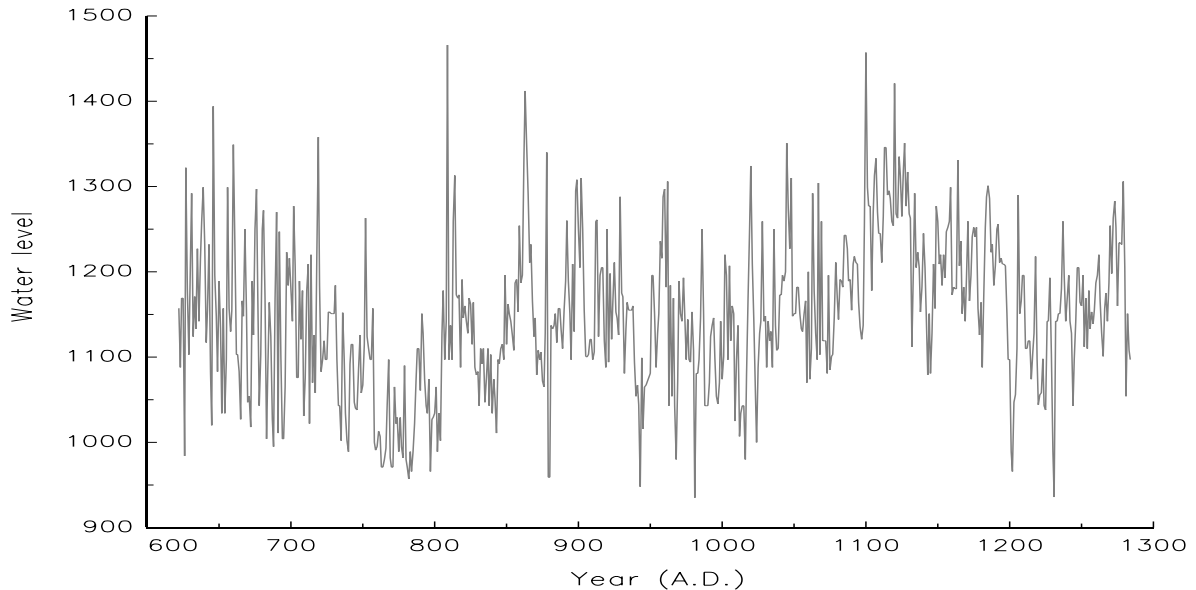
Figure 1: Yearly Nile river minima based on measurements at the Roda gauge near Cairo.

a weakly dependent time series with change-points in the mean to describe. Beran and Terrin (BT) (1996) suggest therefore that the Hurst parameter characterizing the yearly Nile river might change over time. When estimating the Nile river data with the autoregressive fractionally-integrated moving-average (ARFIMA) or $I(d)$ process introduced by Granger (1980), Granger and Joyeux (1980) and Hosking (1981), Beran and Terrin (1996, p.629) show that the data can be well fitted with an ARFIMA$(0, d, 0)$ model with $d = 0.4$, where the fractional differencing parameter d of ARFIMA process acts like the Hurst parameter H of fractional Gaussian noise in characterizing the hyperbolic decay of the autocovariance function of a long memory process. BT further claim that the observations 1 to about 100 seem to be more independent than the subsequent observations, and the value of the fractional differencing parameter might be *lower* for the first 100 observations than for the subsequent data. If this claim is right, then there should be a structural change in the long range persistence of the Nile river data around the year 720, and the Nile river data neither can be described with a pure long memory nor a weakly dependent time series with change-points in the mean.

The possible change of the differencing parameter stimulate BT to propose a statistic for testing the stability of the fractional differencing parameter. This testing statistic has been further discussed and extended in Horváth and Shao (1999) and Horváth (2001). However,

3

their methods can not identify the change points of the fractional differencing parameter. A Bayesian random persistent-shift (RPS) method for detecting structural change in the differencing parameter and the process level has been considered in Ray and Tsay (2002). Nevertheless, the RPS method is not built on the Markov-switching framework, thus may not fully characterize the cycling behavior of the data series, i.e., "seven years of great abundance" and "seven years of famine" — the Joseph effect named by Mandelbrot and van Ness (1968) and Mandelbrot and Wallis (1969).

The above considerations lead us to combine the long memory and Markov-switching literature into a unified framework. We introduce a Markov-switching-ARFIMA (MS-ARFIMA) process by extending the hidden Markov model. Given that the hidden Markov model has become extremely popular in speech recognition as shown in Juang and Rabiner (1991) and Qian and Titterington (1991), and in econometrics, finance, genetics, and neurophysiology as outlined in Robert et al. (2000), the MS-ARFIMA model provides a flexible modeling framework for many applications to these fields. Moreover, the research conducted in this paper also solve the puzzle raised by Diebold and Inoue (2001) by estimating the differencing parameter allowing for the parameters of interest are Markov-switching.

The remaining parts of this paper are arranged as follows: Section 2 presents the MS-ARFIMA process and the algorithms for estimating the parameters of interest. In Section 3 we consider the finite sample performance of the proposed algorithm under the simple mixture of a Markov-switching mean and an ARFIMA$(1, d, 1)$ process. We then apply the proposed methodology to the U.S. real interest rates, the Nile river data, and the U.S. unemployment rates in Section 4. Section 5 provides a conclusion.

2 Models and Main Results

The objective of this paper is to propose a general class of Markov-switching-ARFIMA processes in order to combine strands of long memory and Markov-switching literature. This class of models offers a rich dynamic mixture of a Markov chain and an $I(d)$ process.

Let $\{s_t\}_{t=1}^{T}$ be the latent sample path of an N-state Markov chain. At each time s_t can

assume only an integer value of $1, 2, \cdots, N$, and its transition probability matrix is

$$
\mathcal{P} \equiv \begin{bmatrix}
p_{11} & p_{21} & \cdots & p_{N1} \\
p_{12} & p_{22} & \cdots & p_{N2} \\
\vdots & \vdots & \ddots & \vdots \\
p_{1N} & p_{2N} & \cdots & p_{NN}
\end{bmatrix},
$$

where $p_{ij} = P(s_t = j \mid s_{t-1} = i)$ and $\sum_{j=1}^{N} p_{ij} = 1$ for all i.

An $I(d)$ process, x_t, is defined as:

$$
(1 - L)^d x_t = h_t,
$$

where L is the lag operator ($Lk_t = k_{t-1}$) and h_t is a short memory process. When $d > 0$, the $I(d)$ process is often called the long memory process, because its autocovariance function is not summable so as to capture the long range dependence of a time series. In addition, the $I(d)$ process is nonstationary when $d \geq \frac{1}{2}$, otherwise, it is covariance stationary.

Combining the defining feature of a Markov chain and that of an $I(d)$ process, we propose the following MS-ARFIMA(p, d, q) process:

$$
w_t = \mu_{s_t} I\{t \geq 1\} + (1 - L)^{-d_{s_t}} \sigma_{s_t} z_t I\{t \geq 1\} = \mu_{s_t} I\{t \geq 1\} + y_{s_t}, \tag{1}
$$

where $I\{.\}$ is the indicator function and z_t is stationary process with mean zero and bounded positive spectral density $f_u(\lambda) \sim G_0$ as $\lambda \to 0$ at each possible regime, thus including stationary and invertible ARMA process as its special case. The most distinguished feature of the process is that the fractional differencing parameter d_{s_t} well known in the long memory literature is allowed to be a Markov chain satisfying the following Assumption A:

Assumption A. *s_t is independent of z_τ for all t and τ.*

The model in (1) subsumes many interesting models in the literature. When $N = 1$, w_t reduces to the specification in (7) of Shimotsu and Phillips (2005):

$$
w_t = \mu_0 + (1 - L)^{-d_0} \sigma_0 z_t I\{t \geq 1\} \tag{2}
$$

which also can be represented as:

$$
w_t = \mu_0 + \sum_{k=0}^{t-1} \frac{(d_0)_k}{k} \sigma_0 z_{t-k}, \tag{3}
$$

where

$$(d_0)_k = \frac{\Gamma(d_0 + k)}{\Gamma(d_0)} = (d_0)(d_0 + 1) \dots (d_0 + k - 1) \tag{4}$$

is Pochhammer's symbol for the forward factorial and $\Gamma(.)$ is the gamma function. Moreover, under the model in (1) and $d_{s_t} = 0$, w_t still includes the Markov-switching AR model considered in Hamilton (1989) as one of its special cases. We will show that the estimation of the model in (1) can be easily implemented with the algorithm proposed in this paper, even though the parameter estimation from a noisy version of realizations of Markov models is extremely difficult in all but very simple examples as well documented in Qian and Titterington (1991).

Let the total sample size be T, and denote $\mathcal{W}_t \equiv (w_1, w_2, \cdots, w_t)^\top$ the column vector containing the observations from time 1 to time t, while $\mathcal{S}_t = (s_1, s_2, \cdots, s_t)^\top$ represents the corresponding states, and $\mathcal{Y}_t = (y_1, y_2, \cdots, y_t)^\top$ in (1) is similarly defined. The column vector $\alpha = (\mu_1, \dots, \mu_N, \sigma_1, \dots, \sigma_N, \phi_{11}, \dots, \phi_{1p}, \phi_{21}, \dots, \phi_{Np}, d_1, \dots, d_N, \theta_{11}, \cdots, \theta_{Nq})^\top$ and p_{ij} (transition probabilities) consist of the parameters characterizing the conditional density function (cdf) of w_t. After stacking the parameter vector α and the transition probabilities p_{ij} into one column vector ξ, we can represent the cdf of w_t as $f(w_t \mid \mathcal{S}_t, \mathcal{W}_{t-1}; \xi)$, clearly showing that the cdf of w_t depends on the entire past routes of states (in general). Indeed, there are N^T possible paths of states running throughout the observations $\mathcal{W}_T$.

To illustrate the proposed algorithm for the model in (1), we first consider the simplest case where w_t in (1) is generated as:

$$w_t = \mu_{s_t} I\{t \geq 1\} + (1 - L)^{-d_0} \sigma_0 \varepsilon_t I\{t \geq 1\} = \mu_{s_t} I\{t \geq 1\} + y_t, \tag{5}$$

where $d < \frac{1}{2}$ and ε_t is a zero mean normally, independently and identically distributed white noise (i.i.d.) with $E(\varepsilon_t^2) = 1$. That is, w_t in (5) is a special type of MS-ARFIMA$(0, d, 0)$ process whose differencing parameter is fixed across different regimes. Under Assumption A and $\varepsilon_t \sim N(0, 1)$ i.i.d. process, the likelihood function of $\mathcal{W}_T$, $L(\mathcal{S}_T, \mathcal{W}_T; \xi)$ hereafter, for the hidden Markov model in (5) equals

$$L(\mathcal{S}_T, \mathcal{W}_T; \xi) = (2\pi)^{-T/2} |\Lambda|^{-1/2} \exp\left(-\frac{1}{2} \mathcal{Y}_T^\top \Lambda^{-1} \mathcal{Y}_T\right) \prod_{t=1}^{T} \Pr(s_t \mid s_{t-1}), \tag{6}$$

where $\Lambda = E(\mathcal{Y}_T \mathcal{Y}_T^\top)$, and $\Pr(s_1 \mid s_0)$ is evaluated with the unconditional probability that the process will be in regime s_1. Given that y_t in (5) is a simple ARFIMA$(0, d, 0)$ process,

we can use the Durbin-Levinson algorithm to derive

$$(2\pi)^{-T/2}|\Lambda|^{-1/2}\exp\left(-\frac{1}{2}\mathcal{Y}_T^\top\Lambda^{-1}\mathcal{Y}_T\right) = \prod_{t=1}^T (2\pi)^{-1/2}v_{t-1}^{-1/2}\exp\left\{-\frac{(y_t-\widehat{y}_t)^2}{2v_{t-1}}\right\}, \qquad (7)$$

where $\widehat{y}_t$ denotes the one-step ahead predictor of y_t with the observation $\mathcal{Y}_{t-1}$ as $j \geq 2$, and v_{t-1} is the corresponding one-step ahead prediction variance. Deriche and Tewfik (1993) also have employed the Durbin-Levinson algorithm to estimate a univariate ARFIMA$(0, d, 0)$ processes without Markov-switching characteristic. Note that as $t = 1$, $\widehat{y}_1 = 0$, and $v_0 = \gamma_0$ corresponds to the variance of y_t. As a result, the likelihood function in (6) can be rewritten as:

$$L(\mathcal{S}_T, \mathcal{W}_T; \xi) = \prod_{t=1}^T (2\pi)^{-1/2}v_{t-1}^{-1/2}\exp\left\{-\frac{(y_t-\widehat{y}_t)^2}{2v_{t-1}}\right\}\Pr(s_t \mid s_{t-1}), \qquad (8)$$

indicating that the *unconditional* likelihood function of the mixture model in (5) can be exactly and recursively evaluated provided that we can identify the true path of s_t, $\mathcal{S}_T^*$.

We do not know in reality the value of $\mathcal{S}_T^*$. However, the recursive structure shown in (8) is especially suitable for implementing the Viterbi (1967) algorithm in the digital communication literature to identify the most likely path of states among the N^T possible routes within $\mathcal{W}_T$. We thus combine the Durbin-Levinson algorithm and the Viterbi algorithm to suggest a *Durbin-Levinson-Viterbi* (DLV) algorithm for the model in (5). When compared to the original Viterbi algorithm designed for solving the problem of maximum a posteriori probability estimate of the state sequence of a finite-state discrete-time Markov process observed in white noise, the DLV algorithm proposed in this paper is concerned with the hidden Markov process observed in a much more general ARFIMA noise. Since the DLV algorithm can estimate the differencing parameter of a time series allowing for the presence of a Markov-switching mean, the puzzle raised by Diebold and Inoue (2001) that long memory can be easily confused with a Markov-switching mean is thus resolved by using this DLV algorithm.

To locate the most likely path running through the data $\mathcal{W}_T$ with the idea of Viterbi (1967), we note first that, for each time t, there are N possible states ending at time t, i.e., $(s_t = i)$, $i = 1, \ldots, N$. For a particular node of these N end points at time t, say $(s_t = j)$, there exists a corresponding most likely path:

$$(\mathcal{S}_{t-1}(s_t = j), s_t = j) = (s_1(s_t = j), s_2(s_t = j), \cdots, s_{t-1}(s_t = j), s_t = j), \qquad (9)$$

which ends at this particular node ($s_t = j$). We refer to the path $(\mathcal{S}_{t-1}(s_t = j), s_t = j)$ in (9) as the *survivor* associated with the node ($s_t = j$). Note that, with little loss of clarity, we do not explicitly specify that the path depends on the parameter ξ and the observations $\mathcal{W}_t$ in order to simplify the notation. The likelihood function generated from this survivor $(\mathcal{S}_{t-1}(s_t = j), s_t = j)$ and the formula in (8) is recorded as $L(\mathcal{S}_{t-1}(s_t = j), s_t = j, \mathcal{W}_t; \xi)$ and is crucial for locating the most likely path running from time 1 to time T. In short, for each node ($s_t = j$) at time t, there exists a most likely path, survivor $(\mathcal{S}_{t-1}(s_t = j), s_t = j)$, and its associated likelihood function $L(\mathcal{S}_{t-1}(s_t = j), s_t = j, \mathcal{W}_t; \xi)$. Most importantly, the number of survivors at each time t is always equal to N.

Given the N survivors at time t and in order to locate the survivor $(\mathcal{S}_t(s_{t+1} = i), s_{t+1} = i)$ for a particular node ($s_{t+1} = i$) at time $t + 1$, among the N segments connecting the node ($s_{t+i} = i$) and the N time-t survivors $(\mathcal{S}_{t-1}(s_t = j), s_t = j)$ recorded at time t, we select the one producing the largest likelihood function $L(\mathcal{S}_t(s_{t+1} = i), s_{t+1} = i, \mathcal{W}_{t+1}; \xi)$ among these N possible candidates, and name it as the survivor $(\mathcal{S}_t(s_{t+1} = i), s_{t+1} = i)$ for this particular node ($s_{t+1} = i$). The computation of the aforementioned likelihoods is simple, because we record the likelihood functions of the N time-t survivors at each time t.

This recursive updating process proceeds from time 1 to time T and results in N time-T survivors $(\mathcal{S}_{T-1}(s_T = i), s_T = i)$ and their associated likelihood function $L(\mathcal{S}_{T-1}(s_T = i), s_T = i, \mathcal{W}_T; \xi)$, for each $i = 1, \ldots, N$. From these N time-T survivors we select the one producing the largest likelihood function, say $L(\mathcal{S}_{T-1}(s_T = g), s_T = g, \mathcal{W}_T; \xi)$, as the most likely path running from time 1 to time T. Combining a numerical optimization procedure and this chosen likelihood function $L(\mathcal{S}_{T-1}(s_T = g), s_T = g, \mathcal{W}_T; \xi)$ generated from the Viterbi algorithm and the Durbin-Levinson algorithm displayed in (7), we can estimate the parameters ξ and identify the states $\mathcal{S}_T$ hidden in the observations $\mathcal{W}_T$.

We now consider another special type of MS-ARFIMA(p, d, q) process:

$$w_t = \mu_{s_t} I\{t \geq 1\} + y_t = \mu_{s_t} I\{t \geq 1\} + (1 - L)^{-d_0} \sigma_0 z_t I\{t \geq 1\}, \quad \phi(L)z_t = \theta(L)\varepsilon_t, \quad (10)$$

where

$$\phi(L) = 1 - \phi_1 L - \ldots - \phi_p L^p, \quad \theta(L) = 1 + \theta_1 L + \ldots + \theta_q L^q, \quad (11)$$

and the roots of the polynomial $\phi(L)$ and those of $\theta(L)$ in (11) are all outside the unit circle and share no common roots. The model in (10) is much more general than that in (5), but

still can be estimated with the preceding Viterbi algorithm after some modifications. Please note that the value of fractional differencing parameter is unchanged across different regimes as that imposed in (5).

Note that the term y_t in (10) can be rearranged as

$$y_t = (1 - L)^{-d_0} \sigma_0 \phi(L)^{-1} \theta(L) \varepsilon_t, \qquad t = 1, 2, \ldots. \tag{12}$$

We then have

$$\phi(L) y_t = (1 - L)^{-d_0} \sigma_0 \theta(L) \varepsilon_t = \sigma_0 \theta(L)(1 - L)^{-d_0} \varepsilon_t = \sigma_0 \theta(L) \widetilde{y}_t, \qquad t = 1, 2, \ldots, \tag{13}$$

where $\widetilde{y}_t = (1 - L)^{-d_0} \varepsilon_t$ is an ARFIMA$(0, d, 0)$ process. Dueker and Serletis (2000) use the same transformation method for estimating an ARFIMA(p, d, q) process. Conditional on a set of $\phi(L)$ and $\theta(L)$ and a suitable starting value, the *conditional* likelihood function of y_t in (12) can still be evaluated exactly with the transformed ARFIMA$(0, d, 0)$ $\widetilde{y}_t$ in (13) and the Durbin-Levinson algorithm defined in (7). For example, conditional on y_0 being equal to 0, we can extract an ARFIMA$(0, d, 0)$ process from an ARFIMA$(1, d, 1)$ process as follows:

$$\sigma_0 \widetilde{y}_t = y_t - \phi_1 y_{t-1} - \sigma_0 \theta_1 \widetilde{y}_{t-1}, \qquad t = 1, \ldots, T. \tag{14}$$

Conditional on a set of $\phi(L)$ and $\theta(L)$ and a suitable starting value for the parameter ξ, we can recursively and exactly evaluate the conditional likelihood function of the hidden Markov model using the DLV algorithm proposed previously.

The same idea also applies to the class of MS-ARFIMA(p, d, q) processes in (1) where d can be Markov-switching. However, we cannot use the Durbin-Levinson algorithm when the fractional differencing parameter is allowed to be Markov-switching. Nevertheless, the Viterbi algorithm is still powerful enough to locate the most likely path under this circumstance. That is, conditional on a suitable starting value for the parameter ξ, we employ the recursive structure inherent in Viterbi algorithms to identify the most likely path running through the data set.

3   Monte Carlo Experiment

In this section we consider a Monte Carlo experiment to demonstrate the finite sample performance of the proposed DLV algorithm on a special version of the model in (1):

$$w_t = \mu_{s_t} I\{t \geq 1\} + (1 - L)^{-d_0} \sigma_0 (1 - \phi_1 L)^{-1} (1 + \theta_1 L) \varepsilon_t I\{t \geq 1\}. \tag{15}$$

We employ three different values of the fractional differencing parameter:

$$d_0 = \{0.2, 0.3, 0.4\}, \tag{16}$$

along with the following parameters:

$$\mu_1 = 4, \ \mu_2 = 1, \ \phi_1 = 0.5, \ \theta_1 = 0.5, \ p_{11} = p_{22} = 0.95, \tag{17}$$

and σ_0 is chosen to ensure that the variance of the ARFIMA$(1, d, 1)$ noise in (15) is equal to 1 across different configurations. Note that the positive values of d_0 in (16) are chosen to reflect the variations used in the long memory literature.

All the computations are performed with GAUSS. Two hundred replications are conducted for each specification at 3 different sample sizes ($T = 100, 200, 400$) usually encountered in the empirical applications. For each sample size T, 200 additional values are generated in order to obtain random starting values. The optimization algorithm used to implement the DLV algorithm is the quasi-Newton algorithm of Broyden, Fletcher, Goldfarb, and Shanno (BFGS) contained in the GAUSS MAXLIK library. The maximum number of iterations for each replication is 100.

Table 1 contains the simulation results when the true value of parameters are used as the initial values for estimation procedure. The results reveal that the bias performance from the DLV algorithm is satisfactory (especially when the sample size is larger) for all configurations considered. Moreover, the associated root-mean-squared error (RMSE) almost always decreases with the increasing sample size. We find only two cases where the pattern of RMSE change is not what we expect, i.e., when $d_0 = 0.4$, the RMSE of estimating the parameters μ_1 and μ_2 as $T = 400$ is found to be a little higher than that of estimating the parameters μ_1 and μ_2 as $T = 200$. These two observations demonstrate the ability of the DLV algorithm to deal with the mixture model considered in this section. The performance of DLV algorithm for estimating the fractional differencing parameter is particularly displayed with the box-plots in Figure 2. The above-mentioned observations are clearly borne out in this figure.

We also check the robustness of the preceding simulation results by changing the choice of initial values for estimation. The simulations in Table 1 are replicated by setting the initial values for parameters at the true values except that of d_0 is set at zero. The results

Table 1. Finite sample performance of the DLV algorithm:
Initial values of parameters are set at the true values of parameters

Parameter		μ_1	μ_2	p_{11}	p_{22}	σ_0	d_0	ϕ_1	θ_1
					$d_0 = \mathbf{0.4}$				
$T = 100$	Bias	-0.010	-0.106	0.010	0.016	0.008	0.183	-0.128	-0.040
	RMSE	1.008	0.991	0.039	0.060	0.022	0.294	0.233	0.138
$T = 200$	Bias	-0.094	-0.098	0.006	0.006	0.003	0.135	-0.101	-0.026
	RMSE	0.978	0.978	0.028	0.025	0.015	0.233	0.191	0.086
$T = 400$	Bias	-0.074	-0.076	0.004	0.004	0.001	0.096	-0.073	-0.013
	RMSE	0.990	0.990	0.019	0.017	0.010	0.192	0.163	0.060
					$d_0 = \mathbf{0.3}$				
$T = 100$	Bias	-0.057	-0.070	0.009	0.017	0.012	0.175	-0.109	-0.041
	RMSE	1.042	1.024	0.037	0.060	0.030	0.319	0.245	0.131
$T = 200$	Bias	-0.058	-0.055	0.006	0.006	0.005	0.122	-0.079	-0.030
	RMSE	0.947	0.944	0.027	0.025	0.020	0.260	0.212	0.086
$T = 400$	Bias	-0.038	-0.043	0.004	0.004	0.001	0.090	-0.061	-0.016
	RMSE	0.885	0.883	0.019	0.017	0.015	0.217	0.185	0.061
					$d_0 = \mathbf{0.2}$				
$T = 100$	Bias	-0.017	-0.041	0.009	0.017	0.014	0.201	-0.115	-0.044
	RMSE	0.874	0.853	0.037	0.060	0.037	0.341	0.258	0.128
$T = 200$	Bias	-0.042	-0.047	0.006	0.006	0.006	0.167	-0.106	-0.037
	RMSE	0.795	0.792	0.028	0.025	0.024	0.297	0.239	0.088
$T = 400$	Bias	-0.038	-0.046	0.004	0.004	0.002	0.122	-0.085	-0.019
	RMSE	0.670	0.669	0.019	0.017	0.018	0.239	0.203	0.061

Notes: Simulations are based on 200 replications. The data is generated from the mixture model defined in (15), (16) and (17). DLV algorithm is the Durbin-Levinson-Viterbi algorithm proposed in this paper. Bias is computed as the true parameter minus the corresponding average estimated values.
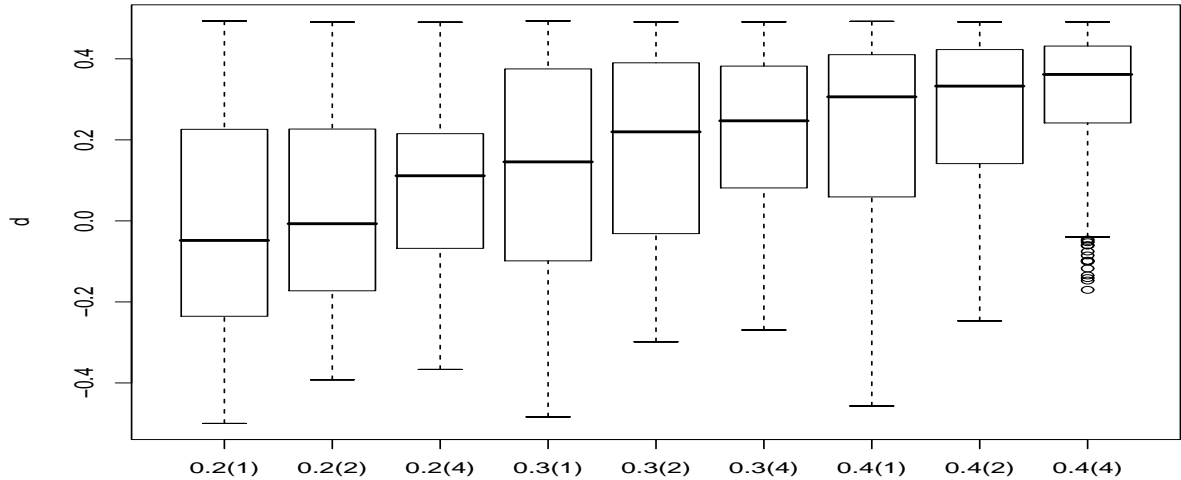
Figure 2: Box-plots of the estimated d from the model defined in (15), (16) and (17) with 200 realizations. The initial values of parameters are set at the true values of parameters. The value $f(g)$ denotes the model specification where $d = f$ and $T = 100 \times g$.

contained Table 2 and Figure 3 indicate that the finite sample performance of our procedure is not sensitive to the initial values used for estimation.

4   Empirical Applications

The methodology developed in this paper is motivated by the dynamic pattern of long memory behavior. Evidence has been given by many methods for such a changing covariance behavior of the Nile river. The applications of the proposed MS-ARFIMA model to actual data are far reaching. For that reason, we consider three data set. The first one is the U.S. real interest rates, the second one is the Nile river data, and the third one is the U.S. unemployment rates.

4.1 Example with real interest rates

In this subsection we first consider the U.S. ex post monthly real interest rate constructed from monthly inflation and Treasury bill rates from January 1953 to December 1990 in Mishkin (1990). The reason we use the original dataset of Mishkin (1990) is to employ it as a benchmark for a clear comparison between the results from the MS-ARFIMA model and

Table 2. Finite sample performance of the DLV algorithm:
Initial values of parameters are set at the true values of parameters
except that of d_0 is set at zero

Parameter		μ_1	μ_2	p_{11}	p_{22}	σ_0	d_0	ϕ_1	θ_1
					$d_0 = \mathbf{0.4}$				
$T = 100$	Bias	-0.116	-0.122	0.010	0.017	0.009	0.188	-0.130	-0.041
	RMSE	1.030	1.017	0.039	0.060	0.021	0.298	0.235	0.137
$T = 200$	Bias	-0.093	-0.096	0.006	0.006	0.003	0.138	-0.103	-0.027
	RMSE	0.979	0.979	0.028	0.025	0.015	0.238	0.193	0.087
$T = 400$	Bias	-0.074	-0.076	0.004	0.004	0.001	0.096	-0.073	-0.013
	RMSE	0.990	0.990	0.019	0.017	0.010	0.192	0.163	0.060
					$d_0 = \mathbf{0.3}$				
$T = 100$	Bias	-0.021	-0.034	0.010	0.017	0.012	0.186	-0.115	-0.040
	RMSE	0.972	0.949	0.039	0.060	0.030	0.325	0.241	0.127
$T = 200$	Bias	-0.046	-0.049	0.006	0.006	0.005	0.126	-0.081	-0.030
	RMSE	0.936	0.937	0.028	0.025	0.021	0.261	0.212	0.086
$T = 400$	Bias	-0.040	-0.044	0.004	0.004	0.002	0.088	-0.059	-0.016
	RMSE	0.912	0.912	0.019	0.017	0.015	0.217	0.184	0.060
					$d_0 = \mathbf{0.2}$				
$T = 100$	Bias	-0.018	-0.038	0.009	0.016	0.014	0.195	-0.110	-0.045
	RMSE	0.892	0.864	0.037	0.060	0.037	0.340	0.260	0.130
$T = 200$	Bias	-0.036	-0.040	0.006	0.006	0.006	0.160	-0.100	-0.037
	RMSE	0.804	0.801	0.028	0.025	0.024	0.294	0.238	0.087
$T = 400$	Bias	-0.044	-0.051	0.004	0.004	0.002	0.117	-0.082	-0.018
	RMSE	0.674	0.673	0.019	0.017	0.018	0.235	0.200	0.060

Notes: Simulations are based on 200 replications. The data is generated from the mixture model defined in (15), (16) and (17). DLV algorithm is the Durbin-Levinson-Viterbi algorithm proposed in this paper. Bias is computed as the true parameter minus the corresponding average estimated values.
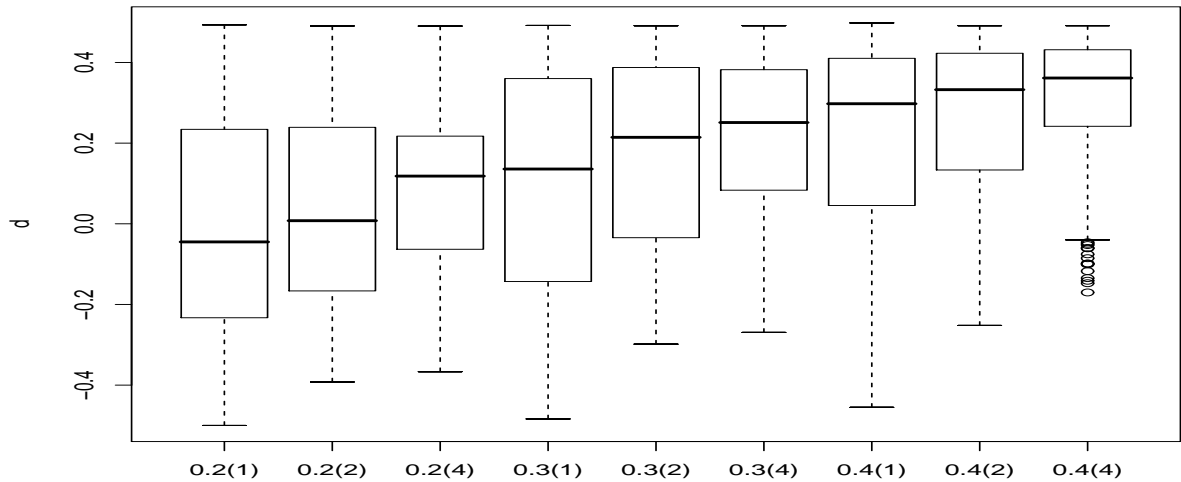
Figure 3: Box-plots of the estimated d from the model defined in (15), (16) and (17) with 200 realizations. The initial values of parameters are set at the true values except that of d_0 is set at zero. The value $f(g)$ denotes the model specification where $d = f$ and $T = 100 \times g$.

those generated from the methodology employed in earlier papers.

The main feature of the real interest rate is that the whole dataset can be split into three subperiods, January 1953-October 1979, November 1979-October 1982, and November 1982-December 1990, because the operating procedure of the monetary authority changed in October 1979 and October 1982 as argued in Mishkin (1990). Another interesting feature of the real interest rate is that the data of these three subperiods can be well described with the ARFIMA models as shown in Tsay (2000). The simultaneous presence of structural break and long memory within the real interest rate allows itself to be an ideal subject to be investigated with the MS-ARFIMA model.

Allowing the break points to be endogenously determined, Table 3 contains the parameter estimates from the following mixture model with a 2-state Markov chain and an ARFIMA$(1, d, 1)$ noise:

$$w_t = \mu_{s_t} I\{t \geq 1\} + (1 - L)^{-d_0} \sigma_0 z_t I\{t \geq 1\}, \quad (1 - \phi_1 L) z_t = (1 + \theta_1 L) \varepsilon_t, \qquad (18)$$

where ϕ_1 or θ_1 is assumed to be zero depending on the noise specification. Following Hamilton (1989), asymptotic standard errors are calculated numerically.

Table 3 shows that the estimates of μ_1, μ_2, p_{11}, p_{22}, σ_0, and d_0 from the DLV algorithm

14

Table 3. Estimates of Parameters Based on Data for U.S. Monthly Real Interest Rate and the DLV Algorithm

	ARFIMA$(0, d, 0)$		ARFIMA$(0, d, 1)$		ARFIMA$(1, d, 0)$		ARFIMA$(1, d, 1)$	
	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
μ_1	5.3455	0.7494	5.3168	0.7162	5.3116	0.7124	5.3626	0.7706
μ_2	0.7226	0.4814	0.7194	0.4383	0.7184	0.4322	0.7352	0.4958
p_{11}	0.9833	0.0150	0.9833	0.0150	0.9833	0.0150	0.9833	0.0150
p_{22}	0.9977	0.0023	0.9977	0.0023	0.9977	0.0023	0.9977	0.0023
σ_0	2.5094	0.0831	2.5091	0.0831	2.5091	0.0831	2.4979	0.0827
d_0	0.2225	0.0367	0.2062	0.0520	0.2034	0.0653	0.2337	0.0376
ϕ_1	-	-	-	-	0.0324	0.0946	-0.9847	0.0155
θ_1	-	-	0.0279	0.0663	-	-	0.9675	0.0200
L^*	1079.0875		1079.0009		1078.9918		1077.0173	

Notes: The results are based on the MS-ARFIMA model defined in (18). S.E. stands for the standard error of the estimate. L^* represents the negative of the log-likelihood function of the switching model. DLV algorithm is the Durbin-Levinson-Viterbi algorithm proposed in this paper.
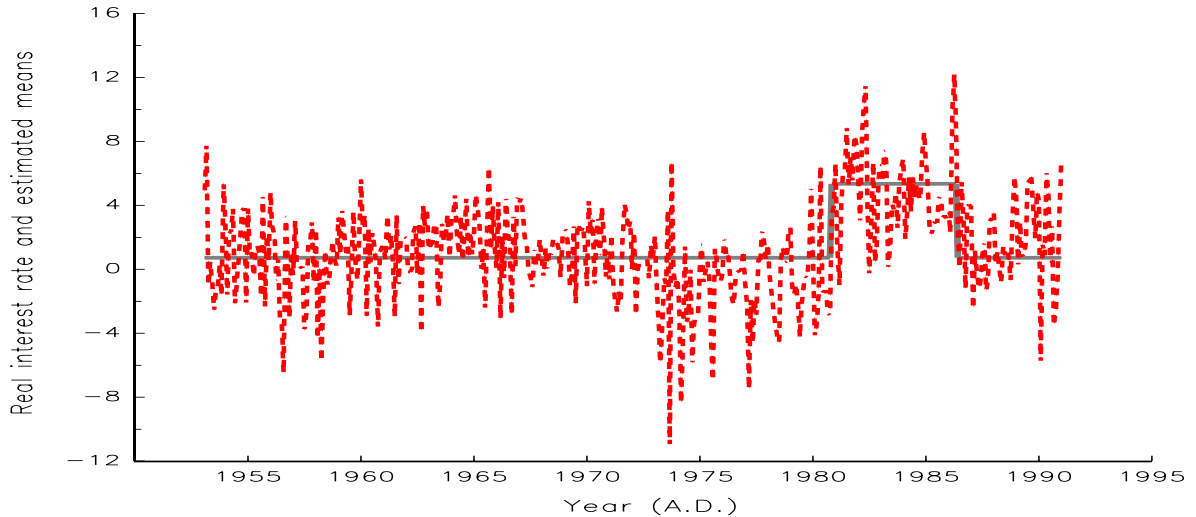
Figure 4: US monthly ex post real interest rates, January 1953-December 1990. Solid line denotes the path of estimated switching means from the specification ARFIMA$(0, d, 0)$ in Table 3, while dotted line denotes the observed monthly ex post real interest rates.

are quite robust across all 4 different configurations. More importantly, two identical break points are identified with these four models, thus divide the whole data into three subperiods as suggested in Mishkin (1990). The endogenous break points identified are November 1980 and May 1986, respectively.

Figure 4 displays the U.S. monthly ex post real interest rates and the path of estimated switching means generated from the DLV algorithm. Without loss of generality, only the path of the estimated switching means from the specification ARFIMA$(0, d, 0)$ in Table 3 is reported. Figure 4 shows that the model in (18) provides a satisfactory fitting of the U.S. monthly real interest rates. Although the endogenously identified break points are later than the well-known monetary operating procedure change points (October 1979 and October 1982), this finding is quite reasonable, because it takes some time for the ex post real interest rate to adjust its path after new information arrives. This argument is buttressed with the findings in Figure 4 that the endogenously identified break points are more closely connected to the observed path of the U.S. monthly ex post real interest rates than the monetary operating procedure change points are.

Table 3 also shows that a long memory phenomenon is found in the real interest rate as has been documented in Tsay (2000). Nevertheless, the estimate of the fractional differencing parameter in Table 3 is much lower than that of 0.666 in Table 3 of Tsay (2000) where the

16

change points are exogenenously determined, and it is more in line with the estimates of 0.204, 0.275, and 0.193 from the individual subperiod data presented in Table 3 of Tsay (2000). This implies that the persistence of long memory in the real interest rate is much more mitigated, once we take the potentially switching mean of the data into account, thus confirming the arguments of Diebold and Inoue (2001) that the presence of Markov-switching level might increase the persistence of the data under investigation.

4.2 Example with Nile river data

In this subsection we apply the Viterbi algorithm to the Nile river data with the following model:

$$w_t = \mu_{s_t} I\{t \geq 1\} + (1 - L)^{-d_{s_t}} \sigma_{s_t} \varepsilon_t I\{t \geq 1\}, \tag{19}$$

where N is assumed to be 2. For the purpose of comparison, we estimate a fixed regime ARFIMA$(0, d, 0)$ model for the Nile river data, i.e., $N = 1$ is imposed on this model. The estimated value of d from such a fixed regime ARFIMA$(0, d, 0)$ model is 0.3986 and is almost identical to the finding in Beran and Terrin (1996).

When estimating the model in (19) with the Viterbi algorithm, we find that the value of the differencing parameter in Table 4 is 0.5770 (nonstationary) for one state, and is 0.2143 (stationary) for the other one. In addition, we identify 5 transitions within the Nile river data in the year 720, 805, 815, 878, and 1070. The estimated path of d_{s_t} from the MS-ARFIMA$(0, d, 0)$ model in Table 4 is graphed in Figure 5.

Most impressively, the first transition data occurs in the year of 720, and the associated estimated value of d_{s_t} within the period 622 to 719 is 0.2143 which is lower than the 0.5770 observed in the other regime. These two findings correspond closely to the conjectures in Beran and Terrin (1996) that the observations 1 to about 100 seem to be more independent than the subsequent observations and the value of differencing parameter might be lower for the first 100 observations than for the subsequent data.

In Figures 6 and 7 we present the observations and the fitted values generated from the estimated parameters displayed in Table 4. It is clear that the fitted value from the MS-ARFIMA$(0, d, 0)$ model is much closer to the real data than that generated from the model whose differencing parameter is not Markov switching. Combining the findings of the likelihood values in Table 4, we find that the MS-ARFIMA$(0, d, 0)$ model is a promising

Table 4. Estimates of MS-ARFIMA$(0, d, 0)$ Model based on the Nile River Data

	MS-ARFIMA$(0, d, 0)$		ARFIMA$(0, d, 0)$	
	Estimate	S.E.	Estimate	S.E.
μ_1	10.8593	0.6903	11.4847	0.2607
μ_2	11.4939	0.0917	-	-
p_{11}	0.9930	0.0042	-	-
p_{22}	0.9918	0.0050	-	-
σ_1	0.5430	0.0202	0.6995	0.0192
σ_2	0.8143	0.0332	-	-
d_1	0.5770	0.0430	0.3986	0.0309
d_2	0.2143	0.0510	-	-
L^*	687.5642		703.8541	

Notes: The MS-ARFIMA$(0, d, 0)$ model is defined in (19). S.E. stands for the standard error of the estimate based on numerical derivative. L^* represents the negative of the log-likelihood function of the estimated model.
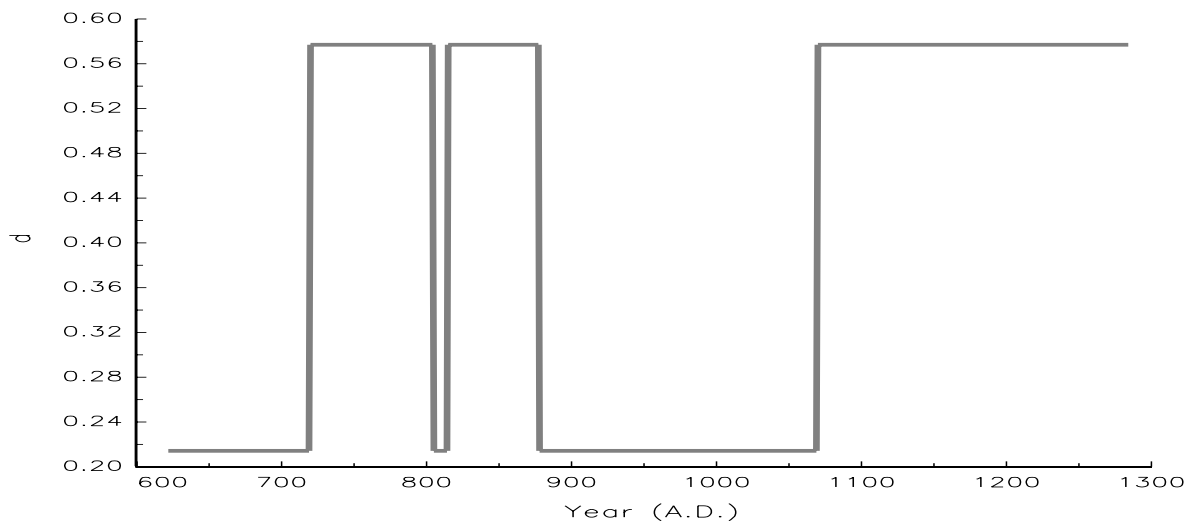
Figure 5: Estimated d_{s_t} from the MS-ARFIMA$(0, d, 0)$ model in Table 4.
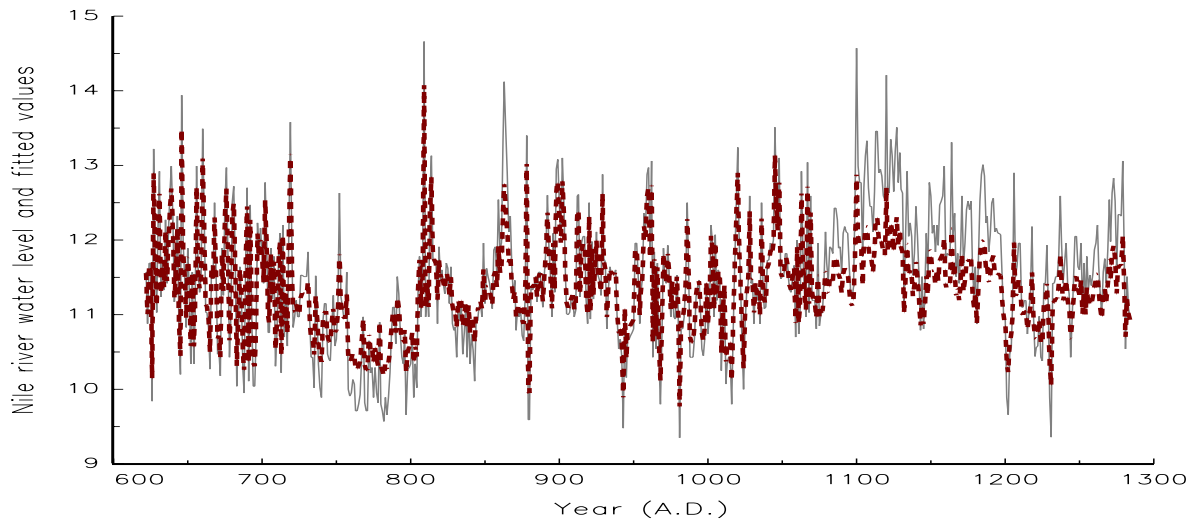
Figure 6: Solid line denotes the Nile river water level divided by 100, while dotted line denotes the corresponding fitted values from the MS-ARFIMA$(0, d, 0)$ model in Table 4.
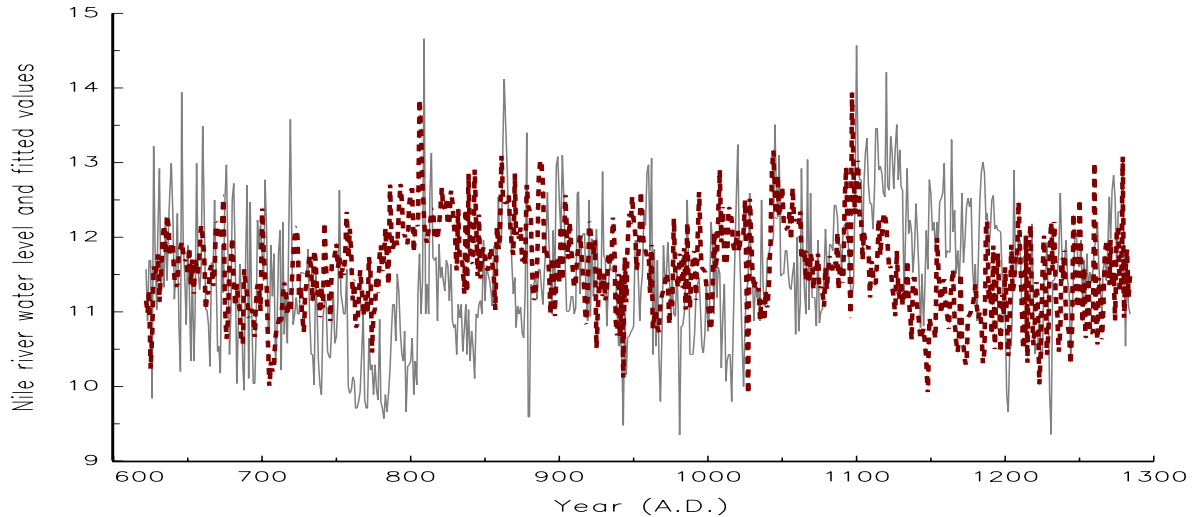
Figure 7: Solid line denotes the Nile river water level divided by 100, while dotted line denotes the corresponding fitted values from the ARFIMA$(0, d, 0)$ model in Table 4.

alternative to describe the Nile river data.

4.3 Example with unemployment rates

In this subsection we apply the Viterbi algorithm to the U.S. quarterly unemployment rates rates from 1948 to 2006. This data is based on the monthly unemployment rates contained in *Bureau of Labour Statistics* as those employed in van Dijk et al. (2002) for estimating a fractionally integrated smooth transition autoregressive (FI-STAR) model. However, van Dijk et al (2002) employ the original monthly unemployment rates ranging from July 1986 to December 1999, while we use all the data contained in *Bureau of Labour Statistics*, but focusing on the quarterly frequency usually considered in the business cycle related studies.

As clearly argued in van Dijk et al. (2002) and shown in Figure 8, there are two important empirical features of U.S. unemployment rates, i.e., the shocks to the series is quite persistent and the series seem to rise faster during recessions than it falls during expansions. van Dijk et al. (2002) find that the estimated d is 0.43 from a FI-STAR model presented in their Table 1. This implies that a time series model describing long memory and nonlinearity simultaneously may be useful for modeling U.S. unemployment rates and many other applications.

The aforementioned two features contained in U.S. unemployment also provide another good opportunity to test the applicability of the MS-ARFIMA model. As a consequence we estimate the U.S. quarterly unemployment rates with the following MS-ARFIMA$(p, d, 0)$
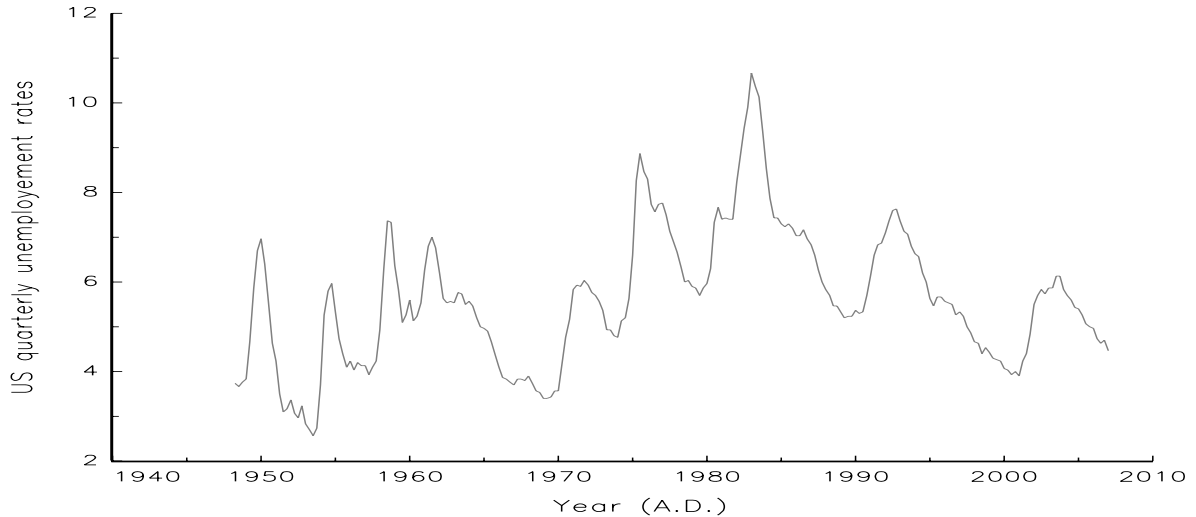
20

Figure 8: U.S. quarterly seasonally adjusted unemployment rates, 1948-2006.

model:

$$w_t = \mu_{s_t} I\{t \geq 1\} + (1 - L)^{-d_{s_t}} \sigma_{s_t} \phi(B)^{-1} \varepsilon_t I\{t \geq 1\}, \tag{20}$$

where N is assumed to be 2, and $p = \{3, 4\}$. The choice of $p = 4$ is adopted by following the model specification in (30) of van Dijk et al (2002), while $p = 3$ is chosen to check the robustness of the estimation results from the specification $p = 4$. The major objective of this subsection is to investigate whether the long memory observed in van Dijk et al. (2002) can also be retained from the MS-ARFIMA methodology.

When estimating the model in (20) with the Viterbi algorithm, we find that the values of the estimated fractional differencing parameter from both MS-ARFIMA$(3, d, 0)$ and MS-ARFIMA$(4, d, 0)$ models in Table 5 are very close to that found in van Dijk et al. (2002), thus confirming that long memory phenomenon seems to be present in the U.S. unemployment rates. For clarity of exposition, the estimated path of d_{s_t} from the MS-ARFIMA$(3, d, 0)$ model and that of d_{s_t} from the MS-ARFIMA$(4, d, 0)$ one are graphed in Figure 9 and Figure 10, respectively. These figures clearly show that d_{s_t} are around 0.4-0.5 for both regimes estimated in each MS-ARFIMA$(p, d, 0)$ model in Table 5.

We also check to what extent the fitted values generated from the models in Table 5 can capture the feature of U.S. unemployment rates. This task is not taken in van Dijk et al. (2002) when estimating their FI-STAR model for the U.S. monthly unemployment rates. It is interesting to find in Figure 11 and Figure 12 that the MS-ARFIMA$(p, d, 0)$ model in (20)

Table 5. Estimates of MS-ARFIMA$(p, d, 0)$ Model based on the U.S. quarterly unemployment rates

	MS-ARFIMA$(3, d, 0)$		MS-ARFIMA$(4, d, 0)$	
	Estimate	S.E.	Estimate	S.E.
μ_1	3.8080	0.1552	3.4572	0.3711
μ_2	5.1358	0.4403	3.8254	0.3334
p_{11}	0.9939	0.0067	0.9877	0.0093
p_{22}	0.9896	0.0083	0.9867	0.0101
σ_1	0.1973	0.0135	0.1535	0.0101
σ_2	0.3380	0.0206	0.3921	0.0274
d_1	0.4919	0.1215	0.4429	0.0987
d_2	0.4143	0.1337	0.4342	0.1058
ϕ_1	1.2570	0.1415	1.1325	0.1215
ϕ_2	-0.3822	0.1510	-0.2301	0.1239
ϕ_3	-0.0666	0.0788	-0.0141	0.1053
ϕ_4	-	-	-0.0495	0.0712
L^*	36.1377		14.7662	

Notes: The results are based on the MS-ARFIMA$(p, d, 0)$ model defined in (20). S.E. stands for the standard error of the estimate based on numerical derivative. L^* represents the negative of the log-likelihood function of the estimated model.
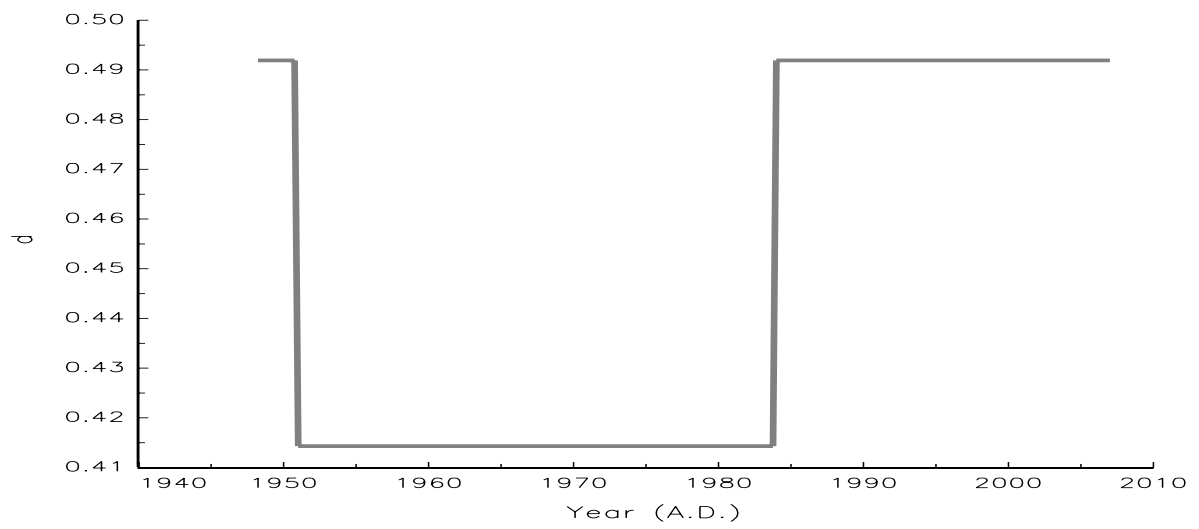
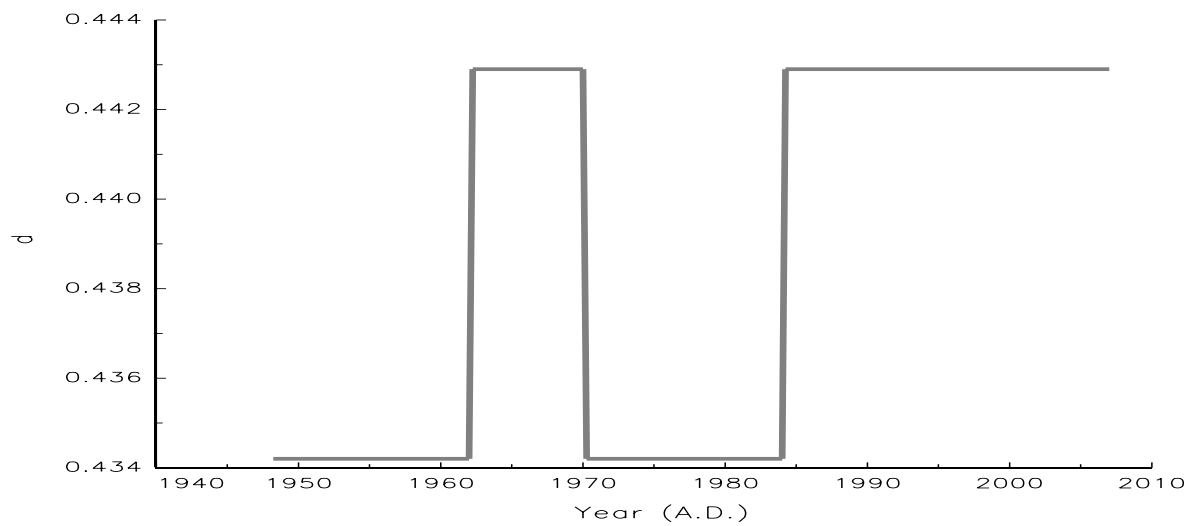Figure 9: Estimated d_{s_t} from the MS-ARFIMA$(3, d, 0)$ model in Table 5.

Figure 10: Estimated d_{s_t} from the MS-ARFIMA$(4, d, 0)$ model in Table 5.
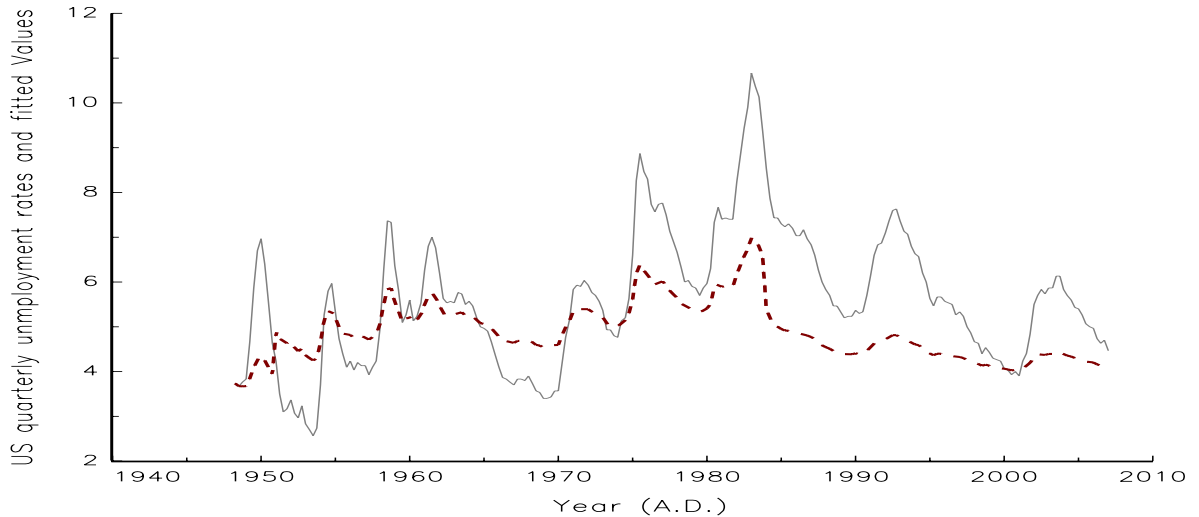
Figure 11: Solid line denotes the U.S. quarterly seasonally adjusted unemployment rates (1948-2006), while dotted line denotes the corresponding fitted values from the MS-ARFIMA$(3, d, 0)$ model in Table 5.

provides a reasonable fit to the data, even though we do not include some seasonal control variables, like seasonal difference operator, as van Dijk et al. (2002) have done for their empirical studies.

5   Conclusions

A general class of MS-ARFIMA processes is suggested to combine long memory and Markov-switching models into one unified framework. The coverage of this class of MS-ARFIMA models is far-reaching, but we show that they still can be easily estimated with the original Viterbi algorithm or the DLV algorithm proposed in this paper. In addition, the simulation reveals that the finite sample performance of the DLV algorithm for a simple mixture model of Markov-switching mean and ARFIMA$(1, d, 1)$ process is satisfactory. When applying the MS-ARFIMA models to the U.S. real interest rates, the Nile river level, and the U.S. unemployment rates, the estimation results are both highly compatible with the conjectures made in the literature. Accordingly, the MS-ARFIMA model considered in this paper not only can be used for solving the puzzle raised by Diebold and Inoue (2001), but can also find many potential applications in several scientific research fields.
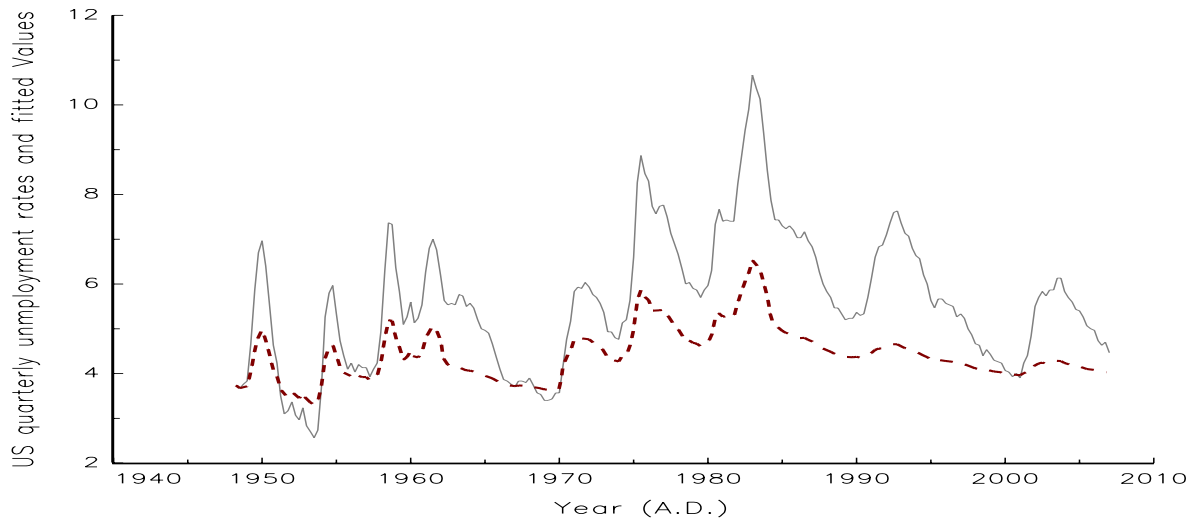
24

Figure 12: Solid line denotes the U.S. quarterly seasonally adjusted unemployment rates (1948-2006), while dotted line denotes the corresponding fitted values from the MS-ARFIMA$(4, d, 0)$ model in Table 5.

References

Beran, J. (1994), *Statistics for Long-Memory Processes*. New York: Chapman and Hall.

Beran, J. and Terrin, N. (1996), "Testing for a Change of the Long-Memory Parameter", *Biometrika*, 83, 627-638.

Berkes, I., Horváth, L., Kokoszka, P. and Shao, Q.M. (2006), "On Discriminating between Long-Range Dependence and Changes in Mean", *The Annals of Statistics*, 34, 1140-1165.

Bhattacharya, R.N., Gupta, V.K. and Waymire, E. (1983), "The Hurst Effect under Trends", *Journal of Applied Probability*, 20, 649-662.

Bollerslev, T. and Mikkelsen, H.O.A. (1996), "Modeling and Pricing Long-Memory in Stock Market Volatility", *Journal of Econometrics*, 73, 151-184.

Breidt, F.J., Crato, N. and de Lima, P. (1998), "The Detection and Estimation of Long Memory in Stochastic Volatility", *Journal of Econometrics*, 83, 325-348.

de Lima, P. and Crato, N. (1993), " Long-Range Dependence in the Conditional Variance of Stock Returns", *Proceedings of the Business and Economic Statistics Section*, August 1993 Joint Statistical Meetings, San Francisco.

Deo, R., Hurvich, C. and Lu, Y. (2006), "Forecasting Realized Volatility Using a Long-Memory Stochastic Volatility Model: Estimation, Prediction and Seasonal Adjustment", *Journal of Econometrics* 131, 29-58.

Deriche, J.A. and Tewfik, A.H. (1993), "Maximum Likelihood Estimation of the Parameters of Discrete Fractionally Differenced Gaussian Noise Process," *IEEE Transactions on Signal Processing*, 41, 2977-2989.

Diebold, F.X. and Inoue, A. (2001), "Long Memory and Regime Switching", *Journal of Econometrics*, 105, 131-159.

Ding, Z., Granger, C.W.J. and Engle, R.F. (1993), "A Long Memory Property of Stock Market Returns and a New Model", *Journal of Empirical Finance*, 1, 83-106.

Dueker, M. and Serletis, A. (2000), "Do Real Exchange Rates Have Autoregressive Unit Roots? A Test under the Alternative of Long Memory and Breaks", Working Paper 2000-016A, Federal Reserve Bank of St. Louis.

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

Granger, C.W.J. and Joyeux, R. (1980), "An Introduction to Long-Memory Time Series Models and Fractional Differencing", *Journal of Time Series Analysis*, 1, 15-29.

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

Heyde, C.C. and Dai, W. (1996), "On the Robustness to Small Trends of Estimation based on the Smoothed Periodogram", *Journal of Time Series Analysis*, 17, 141-150.

Horváth, L. (2001), "Change-Point Detection in Long-Memory Processes. *Journal of Multivariate Analysis*, 78, 218-234.

Horváth, L. and Shao, Q.M. (1999), "Limit Theorems for the Union-Intersection Test", *Journal of Statistical Planning and Inference*, 44, 133-148.

Hosking, J.R.M. (1981), "Fractional Differencing", *Biometrika*, 68, 165-176.

Hurst, H.E. (1951), "Long-Term Storage Capacity of Reservoirs", *Transactions of the American Society of Civil Engineers*, 116, 770-799.

Juang, B.H. and Rabiner, L.R. (1991), "Hidden Marker Models for Speech Recognition", *Technometrics*, 33, 251-272.

Künsch, H. (1986), "Discrimination between Monotonic Trends and Long-Range Dependence", *Journal of Applied Probability*, 23, 1025-1030.

Mandelbrot, B.B. and van Ness, J.W. (1968), "Fractional Brownian Motions, Fractional Noises and Applications", *SIAM Review*, 10, 422-437.

Mandelbrot, B.B. and Wallis, J.R. (1969), "Some Long-Run Properties of Geophysical Records", *Water Resources Research*, 5, 321-340.

Mishkin, F.S. (1990), "What does the Term Structure of Interest Rate Tell Us about Future Inflation? *Journal of Monetary Economics*, 25, 77-95.

Qian, W. and Titterington, D.M. (1991), "Estimation of Parameters in Hidden Markov Models", *Philosophical Transactions: Physical Sciences and Engineering*, 337, 407-428.

Ray, B.K. and Tsay, R.S. (2002), "Bayesian Methods for Change-Point Detection in Long-Range Dependent Process", *Journal of Time Series Analysis*, 23, 687-705.

Robert, C.P., Rydén, R. and Titterington, D.M. (2000), "Bayesian Inference in Hidden Markov Models through the Reversible Jump Markov Chain Monte Carlo Method", *Journal of the Royal Statistical Society B*, 62, 57-75.

Shimotsu, K. and Phillips, P.C.B. (2005), "Exact Local Whittle Estimation of Fractional Integration", *The Annals of Statistics*, 33, 1890-1933.

Tsay, W.J. (2000), "Long Memory Story of the Real Interest Rate", *Economic Letters*, 67, 325-330.

van Dijk, D., Franses, P.H. and Raap, R. (2002), "A Nonlinear Long Memory Model, with an Application to US Unemployment", *Journal of Econometrics*, 110, 135-165.

Viterbi, A.J. (1967), "Error Bounds for Convolutional Codes and an Asymptotic Optimum Decoding Algorithm", *IEEE Transactions on Signal Processing*, IT-13, 260-269.

SFB 649 Discussion Paper Series 2007

For a complete list of Discussion Papers published by the SFB 649, please visit http://sfb649.wiwi.hu-berlin.de.

001 "Trade Liberalisation, Process and Product Innovation, and Relative Skill Demand" by Sebastian Braun, January 2007.

002 "Robust Risk Management. Accounting for Nonstationarity and Heavy Tails" by Ying Chen and Vladimir Spokoiny, January 2007.

003 "Explaining Asset Prices with External Habits and Wage Rigidities in a DSGE Model." by Harald Uhlig, January 2007.

004 "Volatility and Causality in Asia Pacific Financial Markets" by Enzo Weber, January 2007.

005 "Quantile Sieve Estimates For Time Series" by Jürgen Franke, Jean-Pierre Stockis and Joseph Tadjuidje, February 2007.

006 "Real Origins of the Great Depression: Monopolistic Competition, Union Power, and the American Business Cycle in the 1920s" by Monique Ebell and Albrecht Ritschl, February 2007.

007 "Rules, Discretion or Reputation? Monetary Policies and the Efficiency of Financial Markets in Germany, 14th to 16th Centuries" by Oliver Volckart, February 2007.

008 "Sectoral Transformation, Turbulence, and Labour Market Dynamics in Germany" by Ronald Bachmann and Michael C. Burda, February 2007.

009 "Union Wage Compression in a Right-to-Manage Model" by Thorsten Vogel, February 2007.

010 "On σ−additive robust representation of convex risk measures for unbounded financial positions in the presence of uncertainty about the market model" by Volker Krätschmer, March 2007.

011 "Media Coverage and Macroeconomic Information Processing" by Alexandra Niessen, March 2007.

012 "Are Correlations Constant Over Time? Application of the CC-TRIG$_t$-test to Return Series from Different Asset Classes." by Matthias Fischer, March 2007.

013 "Uncertain Paternity, Mating Market Failure, and the Institution of Marriage" by Dirk Bethmann and Michael Kvasnicka, March 2007.

014 "What Happened to the Transatlantic Capital Market Relations?" by Enzo Weber, March 2007.

015 "Who Leads Financial Markets?" by Enzo Weber, April 2007.

016 "Fiscal Policy Rules in Practice" by Andreas Thams, April 2007.

017 "Empirical Pricing Kernels and Investor Preferences" by Kai Detlefsen, Wolfgang Härdle and Rouslan Moro, April 2007.

018 "Simultaneous Causality in International Trade" by Enzo Weber, April 2007.

019 "Regional and Outward Economic Integration in South-East Asia" by Enzo Weber, April 2007.

020 "Computational Statistics and Data Visualization" by Antony Unwin, Chun-houh Chen and Wolfgang Härdle, April 2007.

021 "Ideology Without Ideologists" by Lydia Mechtenberg, April 2007.

022 "A Generalized ARFIMA Process with Markov-Switching Fractional Differencing Parameter" by Wen-Jen Tsay and Wolfgang Härdle, April 2007.