Long memory in the R$/US$ exchange rate: A robust analysis

Márcio Poletti Laurini[1]
Marcelo Savino Portugal[2]

Abstract

This article shows that the evidence of long memory for the daily R$/US$ exchange rate series after the implementation of the Real Plan is not robust when we analyze the existence of structural breaks in this series. We demonstrate that the long memory observed is caused by changes in the structure of variance, captured by a Markov Switching model in all the parameters.

A Monte Carlo study shows that the long memory structure can be induced by changes in the unconditional variance parameters, and that the data generating mechanism is a short memory process.

1 Introduction

The class of models known as Long Memory has recently received great attention in time series models.The most widely used specification of these equations is the one of autoregressive models with fractional order of integration, known as ARFIMA(p,d,q).

The ARFIMA model, introduced by Granger and Joyeaux[21] and Hosking [26], is the generalization of the concept of the order of integration used in ARIMA models, thus allowing the order of integration to be a fractional number. We can represent this process as:

$$(1-L)^d y_t = e_t \tag{1}$$

where e_t is a white noise process. The autocorrelation function has a polynomial decay pattern in fractionally integrated models, differently from other models in which it decays exponentially. This corresponds to the idea of long memory captured in these models. We observed this property when looking at the correlogram of the exchange rate log-return series (figure 1).

2 Properties of the ARFIMA model

The ARFIMA model properties have been widely studied in the literature; the most important of which will be presented next. When the order of integration d of the series is lower than 0.5, the series y_t is covariance-stationary and it has an infinite Moving Average representation:

[1]Ibmec Business School

[2]Professor of Economics at Federal University of Rio Grande do Sul (UFRGS) and associated researcher at CNPq

$$y_t = e_t + \sum_{i=1}^{\infty} \psi_i e_{t-i} \tag{2}$$

Where

$$\psi_i = \frac{d(1+d)...(i-1-d)}{i!} = \frac{(i+d-1)!}{i!(d-1)} \tag{3}$$

When d is greater than -0.5, y_t is invertible and has an infinite autoregressive representation:

$$y_t = \sum_{i=1}^{\infty} \pi y_{t-i} + e_t \tag{4}$$

with

$$\pi_i = \frac{d(1+d)...(i-1-d)}{i!} = \frac{(i+d-1)!}{i!(d-1)} \tag{5}$$

For -0.5 < d < 0.5 the ACF of y_t is given by

$$\rho_i = \frac{d(1+d)...(i-1+d)}{(1-d)(2-d)...(k-d)} \tag{6}$$

in the same case, the PACF is given by

$$\phi_{i,i} = \frac{d}{(i-d)}$$

If the fractionally differenced $(1-L)^d y_t$ series follows an ARMA(p,q) process, we say that this series is an ARFIMA(p,d,q) process. To find the representation of $(1-L)^d$ we can use a binomial expansion for fractional powers:

$$(1-L)^d = \sum_{k=0}^{\infty} (-1)^k \begin{pmatrix} d \\ k \end{pmatrix} L^k \ onde \ \begin{pmatrix} d \\ k \end{pmatrix} = \frac{d(d-1)...(d-k+1)}{d!} \tag{7}$$

The most important characteristic is that, although the ACF of an ARFIMA series is not of a great magnitude, it has a slow decay. This corresponds to the notion of long memory, and the autocorrelations are significant even at very long intervals, as clearly shown in the ACF and PACF graphs of the exchange rate log return series. This characteristic means that shocks to the exchange series have an extremely slow dissipation, that is, high persistence.

3 Sample and Estimation - ARFIMA model

The sample used in our study consists of 1,882 daily observations of the R$\US$ exchange rate log returns from July 1st 1994 to January 4, 2002. The main motivation to work with log returns is that they are usually stationary (covariance-stationary). A second advantage of working with log returns, instead of a series in level form, is that log returns present the behavior of the conditional volatility of the series in a more intuitive manner, as shown in figure 2. The descriptive statistics are shown in table 1 for the exchange rate series in level form, returns $(P_t - P_{t-1})$, and log-returns.

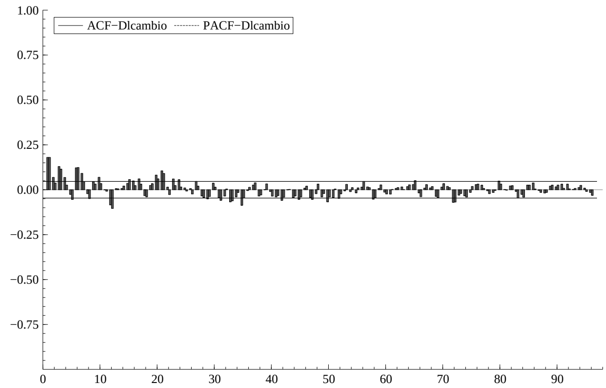

Figure 1: Correlogram - Log-Returns RS$/US$ Exchange Rate

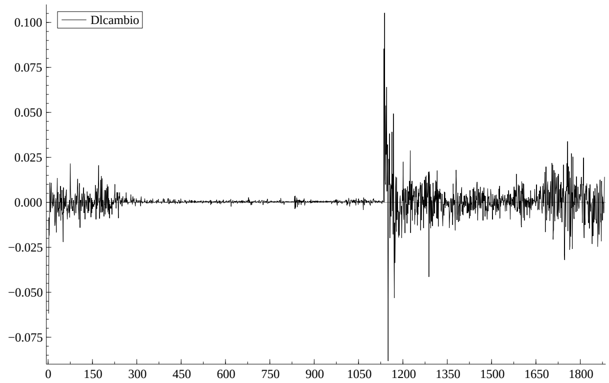

Figure 2: Log-Returns RS$/US$ Exchange Rate

	exchange rate	returns	log-returns
mean	1.417051	0.000753	0.000491
standard error	0.5252	0.0141	0.007634
skewness	0.7734	0.2451	1.8322043
kurtosis	2.3660	30.44971	48.22841
Jarque-Bera	211789	58759.63	160508.6
p-value - JB	0.0000	0.000	0.000

Table 1: Descriptive Statistics

Persistence Measures	sample 1-1181
Lo's RS	1.19579
KPSS	0.173743
Robinson's d	0.112466

Table 2: Persistence Measures

model				sample	12-1881
	d	std. err. (d)	*t-value*(d)	t-*prob*(d)	AIC-T
ARFIMA(2,d,2)	0.149541	0.04600	3.25	0.001	-5.66584723
ARFIMA(1,d,2)	0.159742	0.04760	3.36	0.001	-5.66654016
ARFIMA(0,d,2)	0.18872	0.04231	4.32	0.000	-6.917660
ARFIMA(2,d,1)	0.191031	0.03804	5.02	0.000	-6.9181742
ARFIMA(2,d,0)	0.201205	0.04139	4.86	0.000	-6.92795259
ARFIMA(1,d,1)	0.131943	0.02338	5.64	0.000	-6.91704088
ARFIMA(0,d,1)	0.132439	0.0311	4.26	0.000	-6.91651336
ARFIMA(1,d,0)	0.136198	0.03097	4.40	0.000	-6.91642264
ARFIMA(0,d,0)	0.156432	0.01896	8.25	0.000	-6.91712088

Table 3: ARFIMA - Sample 12-1881

Given the evidence of long memory shown by the ACF and PACF graphs, along with long memory measurements, such as the R/S statistics by Lo, KPSS statistics and Robinson's non-parametric estimation of the order of integration (Table 2), we estimate a series of specifications of an ARFIMA(p,d,q) model by using the Whittle exact maximum likelihood[3] estimation for the exchange rate log return series. Since orders greater than 3 for AR and MA parameters were not significant in this model, we tested the possible combinations with AR and MA parameters lower than 4, as shown in table 3.

By testing all the possible combinations for the ARMA(p,q) part with maximum p and q of 2, we selected the ARFIMA model (2,d,0) by the Akaike information criteria, as shown in table 4.

All the estimated parameters are highly significant, including the parameter that corresponds to the fractional order of integration d of 0.201205. This result would be consistent with the evidence shown by the autocorrelation graph and by the persistence measures (table 2). The estimated value d of 0.201205 defines the process as covariance-stationary, according to the properties seen in section 2. The values forecasted by the ARFIMA (2,d,0) model for the exchange

[3]The model was estimated using the Arfi ma module on PcGive program.

ARFIMA(2,d,0)	sample 12-1881	
parameter	estimated value	standard error
Intercept	0.000722	7.49074e-005
ARFIMA d	0.201205	0.0413921
AR(1)	-0.004658249	0.0007196
AR(2)	-0.054229896	0.0009225

Table 4: Selected Model - ARFIMA(2,d,0)

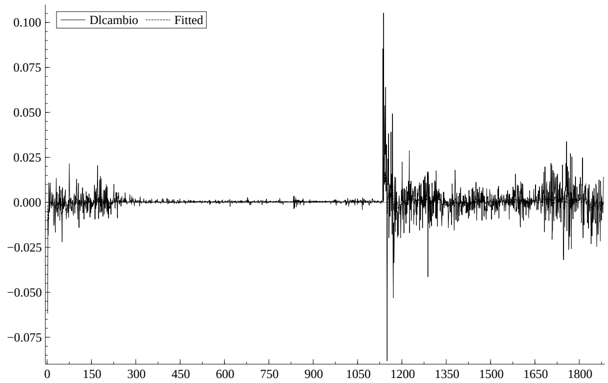

Figure 3: Fitted and Observed Values - ARFIMA(2,d,0) Model

rate log return series can be observed in figure 3.

However, this result should be carefully analyzed. There is a voluminous literature in econometrics that shows that structural breaks and heteroskedasticity can influence the results of econometric procedures. Maybe the most representative example is the work on the influence of structural breaks and heteroskedasticity on the results of unit root tests. Unit root tests have as null hypothesis an I(0) process against an alternative hypothesis of an I(1) process, while in our process we are interested in determining the value of a non-integer order of integration I.

In order to verify whether the estimation methods for long-memory models are robust in the presence of structural breaks, we will carry out a series of procedures to check the behavior of the test in situations of change in heteroskedasticity and in autoregressive parameters.

4 ARFIMA-GARCH Model

A remarkable characteristic of the exchange rate log return series is that the variance of this series is not constant, showing conditional heteroskedasticity. A relevant point is to verify whether conditional heteroskedasticity could induce long-term persistence.

In a first moment, in order to verify this effect, we will estimate a model that combines fractional order of integration and a GARCH (1,1) structure for the conditional variance. The joint estimation allows us to efficiently test, with the control of conditional heteroskedasticity effect, if the order of integration will still be significant.

4.1 ARFIMA-GARCH(1,1)

We estimate a series of ARFIMA models, keeping a GARCH (1,1) structure fixed by using the full sample. The estimation results are shown in table 5. We report the value for the AIC criteria multiplied by the sample size, which is a common procedure in ARFIMA modeling.

model	d	Standard Err. d	AIC-T
ARFIMA(0,d,0)-GARCH(1,1)	0.0176186	(0.0356433)	5632
ARFIMA(1,d,0)-GARCH(1,1)	0.0154518	(0.0591366)	5631.93
ARFIMA(2,d,0)-GARCH(1,1)	-0.00900782	(0.0738184)	5628.18
ARFIMA(3,d,0)-GARCH(1,1)	-0.0160677	(0.0591613)	5643.33
ARFIMA(0,d,1)-GARCH(1,1)	0.883722	(0.0433847)	6077.77
ARFIMA(1,d,1)-GARCH(1,1)	0.999795	(0.000242347)	5678.42
ARFIMA(2,d,1)-GARCH(1,1)	-0.00169456	(0.12095)	5632.05
ARFIMA(3,d,1)-GARCH(1,1)	-0.0156386	(0.0586161)	5642.41
ARFIMA(1,d,2)-GARCH(1,1)	0.999996	(8.88598e-005)	5655.39
ARFIMA(2,d,2)-GARCH(1,1)	not converged	*	*
ARFIMA(1,d,3)-GARCH(1,1)	not converged	*	*
ARFIMA(3,d,2)-GARCH(1,1)	-0.00624241	(0.0518974)	5649.5
ARFIMA(2,d,3)-GARCH(1,1)	-0.008264	(0.0689502)	5637.77
ARFIMA(3,d,3)- GARCH(1,1)	-0.00748883	(0.0553174)	5648.52

Table 5: ARFIMA-GARCH(1,1) Model

ARFIMA(2,d,0)	GARCH(1,1)	sample 1-1881
parameter	estimated value	standard error
Intercept	0.000320622	7.49074e-005
ARFIMA d	-0.00900782	0.0738184
AR(1)	0.078832	0.0779679
AR(2)	0.0512	0.0992569
GARCH Intercept^(1/2)	0.00302543	0.00130698
ARCH(1)	0.1447	0.0505813
GARCH(1)	0.82767	0.0653265

Table 6: Selected ARFIMA(2,d,0)-GARCH(1,1) Model

The model selected by the AIC-T criteria is the ARFIMA (2,d,0)-GARCH (1,1) model. Two phenomena that occur in the estimation of some these models are: the convergence in the maximization of the likelihood function was not achieved; and the value estimated for the fractional order of integration in some models with moving average terms was quite significantly different from the others and very close to one. This difference exists because of the common roots for the autoregressive terms and the moving average in the polynomial that defines the estimated ARMA process. This problem makes the estimation of the order of integration parameter unreliable and, therefore, we disregarded the models in which common roots were present.

In the selected ARFIMA (2,d,0)-GARCH(1,1) model (table 6), we observed that the order of integration parameter is significantly smaller than the one estimated in the ARFIMA(2,d,0) model without the GARCH component. By the value of the standard deviation of this parameter, we do not reject the null hypothesis that it is statistically not different from zero.

When we jointly estimate the coefficient of fractional integration and GARCH parameters, by controling the existence of conditional heteroskedasticity, the existence of long memory is no longer observed; thus, we have some evidence that the phenomenon of long persistence could be generated by changes in the series variance.

However, the non-significance of parameter d could be produced by some structural break in the mean process or by some sampling problem. To verify whether these effects were present, we estimated a Markov Switching model, whereby we captured the structural changes presented by the exchange rate log return series, and we showed that the characteristics of the exchange rate log return series can be generated by alternating short memory processes. This model will also be the basis for the Monte Carlo study, which enhances the conclusions of the present article.

5 Markov Switching Model

The basic idea of Markov Switching model is to decompose a series in a finite sequence of distinct stochastic processes, or regimes, as more widely known in the literature. The current process in each regime is linear, but the combination of processes produces a nonlinear regime.

A simple example is the autoregressive model of first order, which is subject to changes in the autoregressive parameter and is represented by the following system:

$$\begin{array}{c} Y_t = \phi_1 Y_{t-1} + \epsilon_{1t} \,|\, if\, r = 1 \\ Y_t = \phi_2 Y_{t-1} + \epsilon_{2t} | \, if\, r = 2 \end{array} \tag{8}$$

where r represents the current regime. Thus, parameter ϕ_1 describes the behavior of the series when the current regime is 1, but if the current regime is 2, the parameter that will describe the behavior of the series will be ϕ_2.

As we have not observed the stochastic process that determines which the current regime is, we need some way to infer probabilities on which regime is current at period t. The basic idea of Markov Switching model is to describe the stochastic process that determines the switch from one regime to another by means of a Markov Chain. Markov Chain is used to model the behavior of a state variable (or of a combination of variables) that determines which regime is current, as this variable cannot be directly observed.

A Markov chain can be represented as follows: Suppose that the probability of a variable s_t assuming some particular j value, depending only on the previous value s_{t-1}, is given by the following equation

$$P\left\{s_t = j | s_{t-1} = k, ...\right\} = P\left\{s_t = j | s_{t-1} = i\right\} = P_{ij} \tag{9}$$

This process is described as a Markov chain with n-states, whose probability P_{ij} indicates the probability of state i being followed by state j. If we observe that

$$P_{i1} + P_{i2} + P_{in} = 1 \tag{10}$$

we can build the so-called transition matrix, where line i, column j, give the probability of state i being followed by state j.

$$P = \begin{bmatrix} P_{11} & P_{12} & ... & P_{1n} \\ P_{21} & P_{22} & ... & P_{2n} \\ ... & ... & ... & ... \\ ... & ... & ... & ... \\ P_{n1} & P_{n2} & ... & P_{nn} \end{bmatrix} \tag{11}$$

The main characteristic of this Markov transition matrix of first order is that the probability of transition to the next regime relies only on the current regime, which simplifies the modeling and, especially, the estimation methods.

5.1 Estimation and Specification Tests

To find the correct specification for the Markov Switching model that is more appropriate to our data, we used a combination of the general to specific procedures with an analysis of adequacy of the specifications. This mixed procedure was carried out due to the extremely computationally intensive burden for the estimation of these models and also because the test distribution of the number of regimes does not have a standard distribution.

We departed from a general model with 10 regimes and 12 lags and regime switches in the intercept, in autoregressive parameters and in variance which, by using Krolzig [28] notation, is an MSIAH(10)-ARX(12) (Markov Switching Intercept Autoregressive Heteroskedasticity) model. We also included a tendency variable in the specification of the model, which proved to be necessary and relevant.

The models were estimated by means of the EM algorithm proposed by Dempsteir, Laird and Rubin[10] in the form of a BHLK (Baum-Lindgren-Hamilton-Kim) filter, using MSVAR software by Hans Martin Krolzig. In order to discuss this estimation method properly, we would need a presentation that drifts away from the context of application of the method proposed in the present article. The paper written by Krolzig [28] (chapters 5, 6 and 8) discusses this topic in further details.

5.2 Determination of the number of regimes

The major problem with the determination of the appropriate specification for a Markov Switching model is to determine the number of regimes. Tests used to determine the null hypothesis of n-1 regimes against the alternative hypothesis of n regimes do not have a standard distribution, since the null hypothesis is not identified due to the presence of nuisance parameters.

The usual procedure of testing this hypothesis by means of a likelihood ratio test is not valid because the probabilities associated with the additional regime are not identified in the null hypothesis, thus violating the normal conditions of regularity of this test.

Although some procedures used for the derivation of the asymptotic distribution have been proposed by Hansen [25] and Garcia and Perron [18], they are not valid for our general model and require the simulation of the data contained in a grid of values for the nuisance parameters, which would mean a time-consuming simulation for each specification tested. To determine the number of regimes, we will use 3 methods. The first one consists in using information criteria, since Akaike and Schwartz criteria have shown to never underestimate the minimum number of regimes. The second method involves an approximation to the asymptotic distribution of the test, based on Ang, A. and Bekaert, G. [1]. The third method consists in carrying out specification tests to check the necessity for an additional regime preferred procedure.

Table 7 shows the log-likelihood, AIC and BIC information criteria, and the number of parameters, restrictions and nuisance parameters associated with each regime. We estimate MSIAH models with a number of regimes from 10 to 2, and a model with one regime that corresponds to the linear model.

The number of autoregressive lags was selected through AIC and BIC information criteria; a number of 5 lags was considered to be appropriate. According to the table 7, the number of

Regimes	Log-Lik	Parameters	AIC	BIC	Nuisance	Restrictions
10	8944.3934	170	-9.3794	-8.8765	90	72
9	8924.5196	144	-9.3859	-8.9600	72	64
8	8908.4118	120	-9.3943	-9.0394	56	56
7	8883.6830	98	-9.3974*	-9.1016	42	48
6	8858.2377	78	-9.3856	-9.1549	30	40
5	8832.4047	60	-9.3772	-9.1998*	20	32
4	8769.3628	44	-9.3270	-9.1968	12	24
3	8611.4266	30	-9.1731	-9.0844	6	16
2	8309.3226	18	-8.8630	-8.8097	2	8
1	6516.8861	8	-6.9577	-6.9340	0	0

Table 7: Log-Likelihood and Information Criteria

Test	Stat.	Dist	*p-value*
2 against 1	3854.873	$\chi^2(10)$	0.0000 **
3 against 2	604.208	$\chi^2(14)$	0.0000 **
3 against 4	315.8724	$\chi^2(20)$	0.0000 **
5 against 4	126.0838	$\chi^2(28)$	0.0000 **
6 against 5	51.666	$\chi^2(38)$	0.0487*
7 against 6	50.8956	$\chi^2(50)$	0.4381
8 against 7	49.4576	$\chi^2(64)$	0.9095
9 against 8	32.2156	$\chi^2(80)$	1.0000
10 against 9	39.7476	$\chi^2(98)$	1.0000

Table 8: Ang and Bekaert LR Test - Number of Regimes

regimes selected by Akaike information criteria (AIC) corresponds to the model with 7 regimes, whereas Bayes information criteria (BIC) included 5 regimes.

Ang, A. and Bekaert, G. [1] show that the asymptotic distribution of the likelihood ratio test between n-1 and n regimes can be approximated by a Chi-square distribution, where the number of degrees of freedom is given by the number of nuisance parameters of the model with n regimes plus the number of restrictions imposed by regime n on regime n-1. The test statistics is calculated in a usual fashion in likelihood ratio tests, LR=2(log-likelihood(n)-log-likelihood(n-1)), where n and n-1 are the models with n and n-1 regimes.

A test with a significance level of 1% indicates the necessity for a model with 5 regimes, whereas a significance level of 5% shows the necessity for a model with 6 regimes. This test is, however, based on an approximation to the correct critical values and therefore we need further support in order to decide on the optimal number of regimes.

The criterion used to decide on the necessary number of regimes arose from the idea of checking the necessity for an additional regime by means of specification tests. A test through which our viewpoint proves adequate within this context is BDS statistics, a robust test used to determine the presence of remaining structures both in the mean and in the variance of the process. Therefore, BDS statistics is efficient in checking whether the proposed specification can capture the whole structure of n-1 regimes.

Table 9 shows the results of the BDS statistics applied to the residuals of the models with 6 and 7 regimes. In the residuals of the model with 6 regimes, BDS statistics rejects at 5% that these regimes are IID in dimensions 2 and 3, showing some evidence that, with 6 regimes, there is still some uncaptured structure in the mean and/or variance. To check whether 7 regimes are enough to capture all the structure present in the mean and variance, we applied the BDS

BDS Test	Residuals	6	regimes		BDS Test	Residuals	7	regimes	
Dim	BDS Stat.	Std. Err.	z Stat	*p-value*	Dim	BDS Stat.	Std. Err.	z Stat	*p-value*
2	-0.002602	0.001363	-1.909946	0.0461	2	-0.001400	0.001307	-1.071312	0.2840
3	-0.004220	0.002159	-1.954688	0.0406	3	-0.000443	0.002068	-0.214278	0.8303
4	-0.003577	0.002562	-1.395865	0.1628	4	0.001066	0.002451	0.434681	0.6638
5	-0.003316	0.002661	-1.245971	0.2128	5	0.002315	0.002543	0.910344	0.3626
6	-0.002880	0.002557	-1.126457	0.2600	6	0.002775	0.002441	1.137098	0.2555

Table 9: BDS Tests for Specification

Unrestricted Model	Restricted Model	Restrictions	Test Stat	Critical Value	*p-value*
MSIAH(7) 8883.683	MSIA(7) 7667.708	6	2341.94	14.0671	[0.0000] **
MSIAH(7) 8883.683	MSIH(7) 8743.989	35	279.38	49.8118	[0.0000] **
MSIAH(7) 8883.683	MSAH(7) 8866.492	6	34.38	14.0671	[0.0000] **
MSIAH(7) 8883.683	MSI(7) 7134.9259	42	3497.51	58.1240	[0.0000] **

Table 10: LR Tests - Functional Form

statistics to the residuals of the model with 7 regimes. The results of this test show that it was not possible to refute the residuals of the model with MSIAH(7)-ARX(5) as being IID, which indicates that we should work with 7 regimes, thus capturing all the dependence structure present in the exchange rate log-return series.

5.3 Determination of the functional form

Contrariwise to the determination of the number of regimes, the specification test for the most appropriate functional form, in relation to the parameters subject to Markov Switching, has a standard distribution. By using the principle of likelihood ratio, the test statistics is 2(Log-likelihood (unrestricted model)- Log-likelihood (restricted model)), and the distribution is a Chi-square distribution with the number of degrees of freedom correponding to the number of imposed restrictions. We tested the general MSIAH (intercept changes, autoregressive parameters and variance) model against the other possible specifications.

Table10 shows that, according to the LR (Likelihood Ratio) test, the general MSIAH specification is the most appropriate specification, with changes in all parameters, since the other specifications, which determine that some parameters be constant throughout the regimes, are inappropriate.

The analysis of the residuals, presented in table 11and in graph 4 ,shows that the prediction errors of the model have autocorrelation problems, ARCH and heteroskedasticity. However, the standardized residuals, that is, those divided by the variance that corresponds to the regime to which they have great probabilities of belonging, are uncorrelated ,homoskedastic and do not have a conditional ARCH structure. The correction of the switching structure in the unconditional variance by the standardization of residuals shows that, for the exchange rate log-return series, most of the structure present in the series is generated by the changes in unconditional variance.

Standardized Residuals	Dist.	Stat.	*p-value*
Portmanteau(31)	Chi(26)	22.4019	[0.6665]
normality	Chi(2)	10.7503	[0.004] **
asymp. norm.	Chi(2)	9.7392	[0.0077] **
heteroskedasticity	Chi(12)	12.6262	[0.3968]
hetero-χ test	Chi(27)	39.6994	[0.0546]
hetero Squared:	Chi(12)	12.6440	[0.3954]
hetero-χ Squared	Chi(27)	39.7478	[0.0541]
ARCH(5)	Chi(5)	5.44	[0.3536]

Table 11: Specification Tests - Residuals

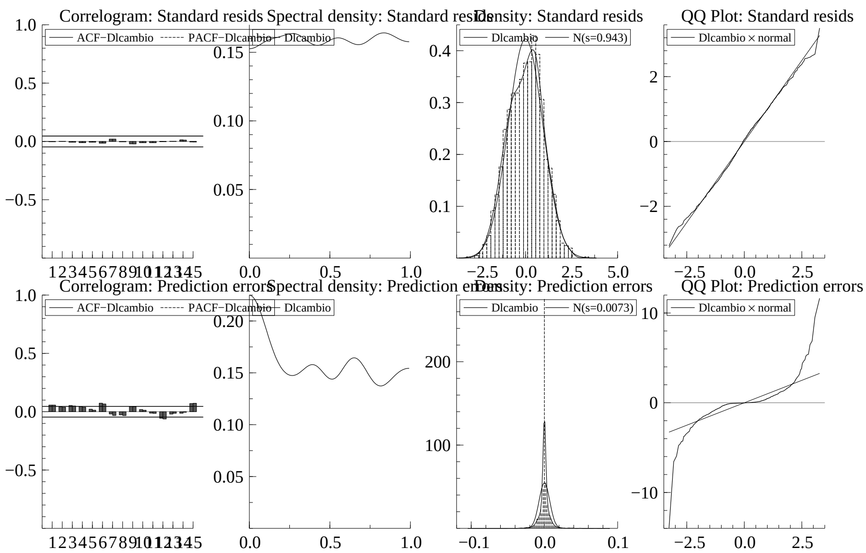

Figure 4: Residual Analysis

5.4 Estimated Model

The estimated MSIAH(7)-ARX(5) model corresponds to the following specification:

$$\left\{\begin{array}{ll} y_{1t}=c_{1t}+\phi_{11}y_{1t-1}+\phi_{12}y_{1t-2}+\phi_{13}y_{1t-2}+\phi_{14}y_{1t-4}+\phi_{15}y_{1t-5}+\beta_1 t+\varepsilon_{1t} & \varepsilon_{1t}\sim N\left(0,\sigma_1^2\right) \\ y_{2t}=c_{2t}+\phi_{21}y_{2t-1}+\phi_{22}y_{2t-2}+\phi_{23}y_{2t-2}+\phi_{24}y_{2t-4}+\phi_{25}y_{2t-5}+\beta_2 t+\varepsilon_{2t} & \varepsilon_{2t}\sim N\left(0,\sigma_2^2\right) \\ y_{3t}=c_{3t}+\phi_{31}y_{3t-1}+\phi_{32}y_{3t-2}+\phi_{33}y_{3t-2}+\phi_{34}y_{3t-4}+\phi_{15}y_{3t-5}+\beta_3 t+\varepsilon_{3t} & \varepsilon_{3t}\sim N\left(0,\sigma_3^2\right) \\ y_{4t}=c_{4t}+\phi_{41}y_{4t-1}+\phi_{42}y_{4t-2}+\phi_{43}y_{4t-2}+\phi_{44}y_{4t-4}+\phi_{45}y_{t-5}+\beta_4 t+\varepsilon_{4t} & \varepsilon_{4t}\sim N\left(0,\sigma_4^2\right) \\ y_{5t}=c_{5t}+\phi_{51}y_{5t-1}+\phi_{52}y_{5t-2}+\phi_{53}y_{5t-2}+\phi_{54}y_{5t-4}+\phi_{55}y_{5t-5}+\beta_5 t+\varepsilon_{5t} & \varepsilon_{5t}\sim N\left(0,\sigma_5^2\right) \\ y_{6t}=c_{6t}+\phi_{61}y_{6t-1}+\phi_{62}y_{6t-2}+\phi_{63}y_{6t-2}+\phi_{64}y_{6t-4}+\phi_{65}y_{6t-5}+\beta_6 t+\varepsilon_{6t} & \varepsilon_{6t}\sim N\left(0,\sigma_6^2\right) \\ y_{7t}=c_{7t}+\phi_{71}y_{7t-1}+\phi_{72}y_{7t-2}+\phi_{73}y_{7t-2}+\phi_{74}y_{7t-4}+\phi_{75}y_{7t-5}+\beta_7 t+\varepsilon_{7t} & \varepsilon_{7t}\sim N\left(0,\sigma_7^2\right) \end{array}\right. \tag{12}$$

Estimated parameters are shown in table 12, along with the standard deviations and t statistics associated with each parameter. While there are regimes in which all the parameters are significant, as in regime 1, no parameter is statistically significant in regime 2. Another interesting fact is that the tendency is only different from zero in regime 7, but the presence of the tendency was confirmed by specification tests. The transition matrix is shown in table14.

Figure 5 shows the estimated probabilities of each regime for each observation in the sample. The graph shows the forecast, filtered (using the information up to period t) and the smoothed probabilities (using the information of the whole sample to infer the probabilities at moment t). This graph shows us that the model associates 3 exclusive regimes (regimes 2, 3 and 4) with the exchange band regime, while the other regimes are present in the remaining regimes of the sample. Regimes 1, 5 and 6 are identified with the periods of free exchange rate variation, whereas regime 7 can be identified, according to the table 12, with the periods of exchange rate crisis, thus combining high variance of the exchange rate with a tendency towards strong devaluation.

Table15 shows the number of observations for each regime, and associates an unconditional probability with each one of the regimes. Regimes 2, 3 and 4, when added, have an associated probability of 0.5933, which corresponds approximately to the percent values of the sample of the band regime (1131/1882=0.6024). Regimes 4, 5 and 6 with a joint probability of 0.4985, correspond to free and controlled periods of fluctuation of the exchange rate. Regimes 1 and 7 correspond the regimes with low probabilities. Regime 1 is associated with periods of continuous exchange rate devaluation, whereas regime 7 is characterized by strong exchange rate devaluation and high volatility.

The association of regimes 3,4 and 5 with the exchange band regime can be seen on the graph of probabilities associated with regimes and also on the transition matrix. The transition matrix shows that the transition probabilities of regimes 3,4 and 5 for the other regimes are almost null, but very significant between them. We can interpret the existence of 3 regimes within the periods of exchange bands as the presence of 2 regimes acting as the upper and lower limits of the band, and a third regime as the exchange rate value within the limits of the band. By looking at the parameters estimated for these regimes, we can identify regime 3 as being the regime with normal values within the limits of the exchange band, while the values of regime 4 are close to the lower limit, and regime 5 contains the returns with values close to the upper limit of the band.

This conclusion is consistent with the mean values of duration of these regimes, since the mean duration of regimes 4 and 5 is 3.5 days on average, while the mean duration of regime 3 is

Regime 1	Coef.	SD	*t-value*
Const	-0.0209	0.0019	-11.1947
Dlcambio_1	-0.1225	0.0940	-1.3037
Dlcambio_2	-0.3699	0.0892	-4.1483
Dlcambio_3	-0.5672	0.0996	-5.6919
Dlcambio_4	-0.2190	0.0906	-2.4176
Dlcambio_5	-0.7564	0.0791	-9.5628
Trend	0.0000	0.0000	0.4548
D.Padrão	0.0095764		

Regime 2	Coef.	SD	*t-value*
Const	-0.0004	0.0002	-1.9478
Dlcambio_1	-0.0007	0.0635	-0.0114
Dlcambio_2	-0.0083	0.0539	-0.1547
Dlcambio_3	-0.0427	0.0527	-0.8097
Dlcambio_4	0.0837	0.0484	1.7289
Dlcambio_5	-0.0021	0.0497	-0.0414
Trend	0.0000	0.0000	0.0422
D.Padrão	0.0042088		

Regime 3	Coef.	SD	*t-value*
Const	0.0002	0.0001	2.0211
Dlcambio_1	-0.0146	0.0663	-0.2201
Dlcambio_2	-0.1437	0.0622	-2.3103
Dlcambio_3	-0.0203	0.0607	-0.3343
Dlcambio_4	-0.1258	0.0526	-2.3922
Dlcambio_5	0.0795	0.0674	1.1790
Trend	0.0000	0.0000	0.0345
SD	0.0004381		

Regime 4	Coef.	SD	*t-value*
Const	-0.0000	0.0000	-1.9315
Dlcambio_1	-0.0150	0.0266	-0.5664
Dlcambio_2	-0.0873	0.0180	-4.8547
Dlcambio_3	0.0309	0.0154	2.0102
Dlcambio_4	-0.0646	0.0136	-4.7414
Dlcambio_5	-0.0241	0.0151	-1.5928
Trend	0.0000	0.0000	0.0333
SD	0.0001678		

Regime 5	Coef.	SD	*t-value*
Const	0.0021	0.0001	15.8368
Dlcambio_1	-0.1979	0.0713	-2.7755
Dlcambio_2	-0.1540	0.0745	-2.0686
Dlcambio_3	-0.1840	0.0735	-2.5049
Dlcambio_4	-0.1806	0.0752	-2.4018
Dlcambio_5	-0.2012	0.0758	-2.6533
Trend	-0.0000	0.0000	-0.2335
SD	0.0013201		

Regime 6	Coef.	SD	*t-value*
Const	-0.0000	0.0004	-0.1106
Dlcambio_1	0.1077	0.0486	2.2137
Dlcambio_2	-0.1328	0.0453	-2.9307
Dlcambio_3	0.1168	0.0458	2.5513
Dlcambio_4	-0.0053	0.0436	-0.1227
Dlcambio_5	0.0421	0.0420	1.0007
Trend	0.0000	0.0000	0.0860
SD	0.0084794		

Regime 7	Coef.	SD	*t-value*
Const	0.0924	0.0062	14.9538
Dlcambio_1	-0.2350	0.1496	-1.5707
Dlcambio_2	0.1601	0.1539	1.0405
Dlcambio_3	0.2282	0.1587	1.4377
Dlcambio_4	-0.1669	0.1628	-1.0256
Dlcambio_5	-0.2677	0.1648	-1.6237
Trend	-0.0001	0.0000	-3.6738
SD	0.0227350		

Table 12: Estimated Parameters

	regime1	regime2	regime3	regime4	regime5	regime6	regime7
regime1	0.7122	0.0009091	2.585e-008	6.118e-009	1.005e-007	0.2869	1.217e-011
regime2	8.171e-007	0.9744	8.937e-008	2.735e-008	0.004562	0.02103	3.011e-011
regime3	2.909e-011	9.485e-008	0.8714	0.06670	0.06194	9.380e-009	1.885e-015
regime4	0.003840	1.068e-007	0.1716	0.7283	0.09189	1.926e-008	0.004342
regime5	0.004077	6.174e-007	0.06717	0.1963	0.7324	7.645e-008	7.992e-015
regime6	0.01168	0.02171	1.034e-008	0.002224	4.566e-008	0.9558	0.008568
regime7	0.1250	8.225e-007	3.491e-011	8.977e-012	1.277e-010	1.128e-007	0.8750
Note: p[i][j]=Pr{s(t+1)=j|s(t)=i}							

Table 14: Transition Matrix

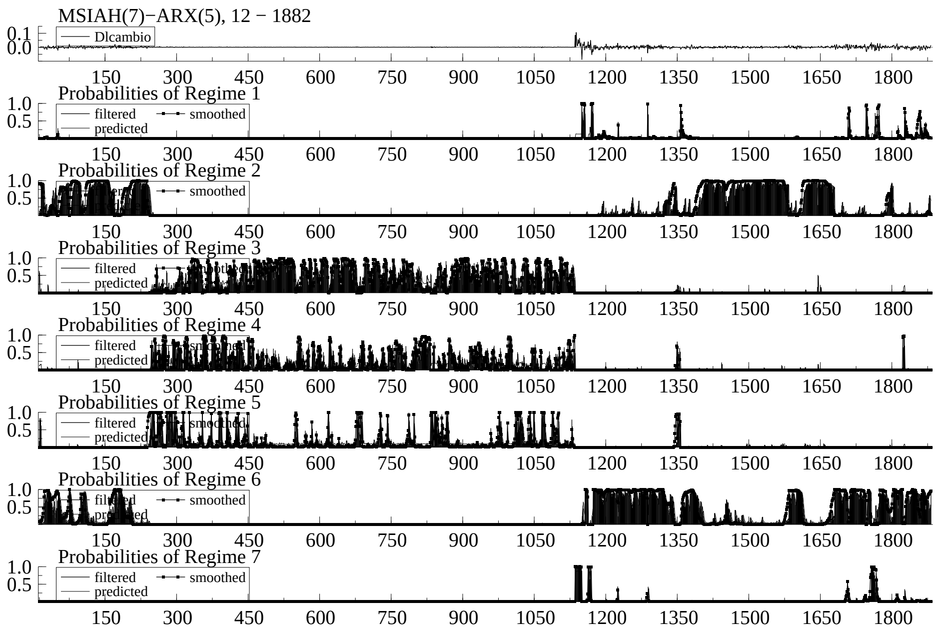

Figure 5: Estimated Probabilities

	Number of Observations	Unconditional Probability	Duration
Regime 1	42.3	0.0233	3.47
Regime 2	420.7	0.2171	39.08
Regime 3	452.9	0.2388	7.77
Regime 4	256.4	0.1374	3.68
Regime 5	201.1	0.1062	3.74
Regime 6	457.3	0.2549	22.63
Regime 7	40.2	0.0222	8.00

Table 15: Regimes and Duration

real	1.0000	0.99353	0.95765	0.85702	0.73215	0.70606	0.60312

Table 16: Eigenvalues of Transition Matrix

7.8 days. The fact that the duration of these regimes in effect during the exchange band regime is short shows that frequent interventions in the exchange rate market were necessary to keep the value within the intervals established by the Central Bank. We can observe that regime 2 is the most persistent, with an average duration of 38.08 days, followed by regime 6 with 22.63 days. These two regimes are characterized by the lowest correlations between the estimated regimes. The other regimes are much less persistent, with frequent switches.

An interesting characteristic is that regime 7, associated with moments of crisis in the exchange market, even with the lowest unconditional probability among all regimes, has the third highest mean persistence, lasting, on average, 8 days. This occurs because the probability of being in regime 7 compared with the probability of remaining in regime 7 is 87%, according to the transition matrix. It is also relevant to observe that regime 1, which is associated with periods of exchange rate devaluation, has a greater probability of preceding regime 7. This characteristic is consistent with the existence of extreme values and groups of high volatility in the series, which is frequently modeled through GARCH models.

The eigenvalues of the transition matrix are shown in table16. Since the first eigenvalue is equal to one and the other eigenvalues are within the unit circle, the transition matrix is ergodic, as the transition matrix is also irreductible. Thus, the eigenvector associated with the unit eigenvalue present in 16 represents the ergodic probabilities of the process. This vector also indicates the unconditional probability of each regime, and therefore we build the table 15. The fact that the transition matrix is ergodic confirms that our regime is stationary since, according to Hamilton [22] (pages 681 and 682), a Markov switching process with an ergodic transition matrix is always covariance-stationary.

6 ARFIMA model applied to the residuals of the MSH model

To check whether the variance structure determined by the Markov Switching model presented in chapter 2 could also be conducive to the presence of long memory, we carried out a two-step procedure to test this hypothesis.

The first step consisted in estimating a Markov switching model for the exchange rate log return series. So as not to affect the autocorrelation structure of the log return series, the only

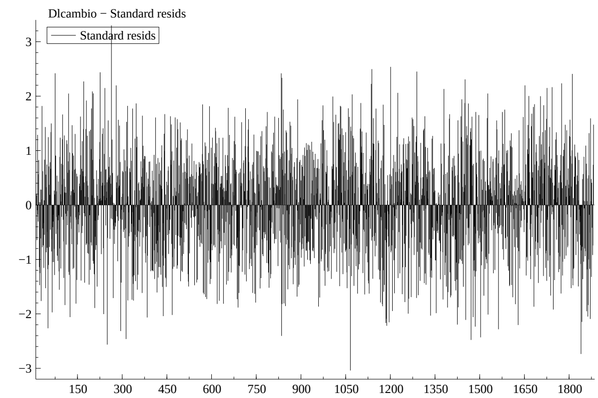

Figure 6: Standardized Residuals

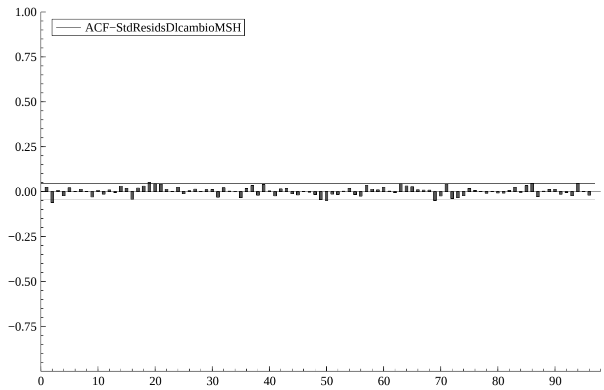

Figure 7: ACF - Standardized Residuals - MSH Model

parameter that was subject to Markov switching was the unconditional variance. By using the same number of regimes established for the model in section 5, we estimated an MSH (Markov Switching Heteroskedasticity) model with 7 regimes.

The second step was to estimate a sequence of ARFIMA(p,d,q) models for the series of standardized residuals of the MSH(7)-AR(0) model estimated by us. Figure 7 shows the graph of the autocorrelation function for this estimated series. One should observe that even though the autocorrelation pattern is the same as that of the original log return series, in the standardization in which the 7 variances estimated by the MSH model were used, all correlations are within the range of significance. Therefore, after the presented autocorrelations were corrected by the variance, they became statistically nonsignificant.

The models estimated for the series of standardized residuals of the MSH model are shown in table 17. Since, by definition, the residuals have mean zero, we restricted the estimation of these ARFIMA models by causing the mean to be zero and by not including a constant in the estimation.

Although, according to the graph of the partial autocorrelation function, none of the cor-

model	d	Standard Error	AIC-T
ARFIMA(0,d,0)	0.00654766	(0.0189905)	847.429
ARFIMA(1,d,0)	-0.0184685	(0.0285412)	847.205
ARFIMA(2,d,0)	0.0328249	(0.0326322)	849.906
ARFIMA(3,d,0)	0.0335371	(0.0374943)	848.907
ARFIMA(0,d,1)	0.053764	(0.0399949)	847.425
ARFIMA(1,d,1)	-0.01643	(0.02187)	849.546
ARFIMA(2,d,1)	0.00869183	(0.0311261)	849.242
ARFIMA(3,d,1)	0.022684	(0.0337418)	848.571
ARFIMA(0,d,2)	0.0469924	(0.0429947)	850.372
ARFIMA(1,d,2)	0.0719771	(0.210842)	849.416
ARFIMA(2,d,2)	0.142631	(0.0913865)	850.248
ARFIMA(0,d,3)	0.0484714	(0.0548457)	849.374
ARFIMA(1,d,3)	0.254321	(0.119458)	850.176

Table 17: ARFIMA - Residuals of MSH Model

ARFIMA(1,d,0)	sample 1-1871	
parameter	estimated value	standard error
ARFIMA-d	-0.0184685	0.0285412
AR(1)	0.043508	0.033922

Table 18: Selected Model - ARFIMA(1,d,0) - Residuals of MSH Model

relations were out of the range, the model selected by the AIC-T information criteria was an ARFIMA(1,d,0) model, which contained an autoregressive term of first order for the specification of the mean. The estimated parameters are shown in table18. None of the estimated parameters was statistically different from zero.

The result of this two-step procedure is also consistent with the hypothesis that the long persistence found by ARFIMA models without adjustment for the variance structure is a spurious phenomenon induced by unconditional switching heteroskedasticity.

Two observations should be made as to the validity of the two-step estimation procedure. The first one is that since we have not jointly estimated parameter d and the variances used to standardize data, the procedure is statistically inefficient. The second one is concerned with the fact that as we are working with a dependent variable constructed with an estimation procedure, the t test of the hypothesis that the parameter is different from zero, using a distribution t (or Normal, by the number of the degrees of freedom), is only valid asymptotically.

7 Monte Carlo Study

The possibility of the long memory observed in our exchange rate log return series being spurious had already been studied in the literature. Bollerslev and Mikkelson [6], in a study of the possibility of long memory in volatility in stock market indices, suggest that the long memory detected in these indices could be a result of aggregation, since individual residuals did not have significant long memory components in volatility.

T	sign.	>0.05	>0.10	>.25	mean	variance
100	99 (4.95%)	145 (7.25%)	96 (4.80%)	25 (1.25%)	0.0185	0.0024
500	203 (10.15%)	298 (14.90%)	136 (6.80%)	5 (0.25%)	0.0238	0.00158
2000	249 (12.45%)	191 (9.55%)	17 (0.35%)	0 (0.0%)	0.0189	0.0004

Table 19: DGP - Unconditional Probabilities, Unit Variance

Granger [20]shows that the sum of autoregressive AR(1) processes of first order randomly taken from some adequate distribution (in this study, Granger used a beta distribution), generates processes with long memory in the process mean when the number of terms in the sum increases.

Based on these results, we used a Monte Carlo study in order to identify whether the long memory in our exchange rate log return series could be generated by the sum of autoregressive components or by structural changes in conditional variance.

7.1 Description of Monte Carlo experiment

The aim of our procedure was to test, by simulation, whether a generating process consistent with the model we estimated in chapter 2 could induce the presence of long memory in a series, even if the memory of this process is short.

The statistical data-generating mechanism (DGP - Data Generating Process) we selected was a Markov switching model for all parameters (intercept, autoregressive parameters and unconditional variance) of an autoregressive process of order 5, according to representation 12.

The parameters chosen to represent the DGP of our data were the parameters estimated by the MSIAH(7)-AR(5)model for the exchange rate log return series. We removed the trend component, which was only significant in regime 7 as a way to make the obtained results easily comparable. By using this DGP, we tested four specifications for the process.

7.2 Simulation using unconditional probabilities and unit variance

The first specification for the generating process was to produce 2,000 replications of a process with the parameters defined in table 12, in which the current regime was determined by the unconditional probability of data. In this first specification for the Monte Carlo experiment, the regimes were taken from a uniform distribution with values determined by table 12. Thus, we did not have the structure of a Markov chain for the determination of the current regime. In this first experiment, the variance of the error component of each replication had a unit variance. In brief, this procedure tests whether the sum of AR(5) processes taken from a uniform distribution can induce long memory. Two thousand replications for sample sizes 100, 500 and 2000 were carried out.

The results of this experiment are shown in table 19. This table shows the values statistically different from zero (sign column) and the percentage in terms of the total number of replications. Columns > 0.05, > 0.10 and > .025 show the total of fractional integration coefficients that were greater than the value indicated in the column, as well as the proportion related to the total. The

T	sign.	0.05	0.10	.25	mean	variance
100	149 (7.45%)	239 (11.95%)	158 (7.90%)	35 (1.75%)	0.0247	0.0030
500	225 (11.25%)	335 (16.75%)	136 (6.80%)	0 (0.0%)	0.0242	0.0014
2000	610 (30.50%)	152 (7.60%)	28 (1.40%)	0 (0.0%)	0.0115	0.0003

Table 20: DGP - Unconditional Probabilities, Markovian Variance

T	sign.	0.05	0.10	.25	mean	variance
100	161 (8.05%)	269 (13.45%)	192 (9.60%)	37 (1.85%)	0.0204	0.0036
500	246 (12.30%)	357 (17.85%)	149 (7.45%)	2 (0.01%)	0.0960	0.0178
2000	279 (13.95%)	234 (11.70%)	19 (0.09%)	0 (0.0%)	0.1695	0.0252

Table 21: DGP - Conditional Probabilities (*Markov Switching*) - Unit Variance

two remaining columns show the mean and variance of the vector of fractional integration values estimated for the experiment's 2,000 replications.

The mean estimated for each sample size is near zero, but the number of estimated significant d is higher than 10% for the samples with sizes 500 and 2000, although it is difficult to establish some behavior pattern in this experiment. However, we prove that the sum of AR(5) components selected by a uniform distribution with the employed probabilities can induce the presence of long memory for the set of parameters used in the experiment's DGP.

7.3 Simulation using unconditional probabilities and variances of the MSIAH model

The second experiment uses the same structure of the previous experiment (transition of regimes determined by a uniform distribution, without using Markov transition matrix), but now the error component of each autoregression has a different variance. For the experiment, we associated the variances estimated in table 12 with each autoregression used.

In this experiment, we could prove that changes in the variance of the process induce long memory on the behavior of the series. The number of significant parameters of fractional integration remarkably increases with the sample size, as observed in the second column of table 20.

7.4 Simulation using the Markov structure for regimes - Fixed variance

In this experiment, the data-generating mechanism is a Markov Switching model in autoregressive parameters and in the intercept, and the transition between regimes is determined by a transition matrix according to 11. To define the DGP used in the experiment, we used the parameters estimated in chapter 2 (table 12) and also the transition matrix (table 14) estimated by the same model. To isolate the effect of the Markov structure on variance, the error component had a unit variance in this experiment. Just as in previous experiments, we generated 2,000 replications for samples with sizes 100, 500 and 2000.

T	sign.	0.05	0.10	0.25	mean	variance
50	52(2.60%)	64(3.20%)	50(2.50%)	19(0.90%)	0.0128	0.0015
100	96(4.80%)	136(6.80%)	105(5.25%)	43(2.15%)	0.0204	0.0036
250	224(11.20%)	554(27.70%)	430(21.50%)	199(9.95%)	0.0635	0.0122
500	264(13.20%)	818(40.90%)	663(33.15%)	287(14.35%)	0.0960	0.0178
750	360(18.00%)	936(46.80%)	808(40.40%)	340(17%)	0.1141	0.0203
1000	378(18.90%)	984(49.2%)	840(42.00%)	400(20%)	0.1251	0.0223
1500	400(20%)	1280(64.00%)	1104(55.20%)	509(25.45%)	0.1618	0.0238
2000	449(22.45%)	1288(64.4%)	1097(54.85%)	626(31.30%)	0.1695	0.0252
5000	736(36.80%)	1776(88.80%)	1616(80.80%)	768(38.40%)	0.2205	0.0117
10000	1264(63.20%)	1952(97.6%)	1824(91.20%)	936(46.80%)	0.2454	0.0112

Table 22: DGP - Conditional Probabilities (*Markov Switching*) - Full Markov Switching in Variance

The results of this experiment (table 21) show a relevant proportion of parameters of fractional order of integration that are statistically significant (8.05% in the sample with size 100 and 13.95% in the sample with size 2000). The number of estimated parameters d greater than 0.05 and 0.10 also increases for samples with size 100 and 500, although part of them is not statistically significant (the proportion of values in column 3 is greater than that in column 2). An interesting phenomenon captured in this experiment is that the increase in the sample size elevates the number of long memory processes. The mean vector for the estimated parameters d has also increased significantly (means of 0.02, 0.09 and 0.16 for samples with size 100,500 and 2000).

7.5 Simulation using total Markov structure in regimes

The DGP of the last experiment allows all parameters in an autoregressive process to change according to a Markov transition chain, since in this experiment, we allow unconditional variance to undergo changes. The data-generating process is an MSIAH(7)-AR(5) model with parameters defined by table 12 and the transition matrix shown in table 14. Again, we generated 2,000 replications for each sample size; however, in this experiment, we used sample sizes 50,100, 250,500,750,1000, 1500,2000, 5000 and 10000.

The results of the Monte Carlo experiment for this DGP, which allows variances to change according to a Markov chain, are quite interesting. Table 22, containing the results of this experiment, shows that the number of spurious long memory processes grows in a hyperbolic fashion with the increase of the sample size. The number of processes with statistically significant values for the order of fractional integration is 2.60% for sample with size 50, amounts to 18.90% in sample size 1000 and reaches 63.20% for sample size 10000.

When we analyze the column that contains the proportion of estimated parameters of the order of fractional integration greater than 0.05, the results are even more striking. The proportion of processes with d greater than 0.05 is 3.20% with sample size 50, reaches 49.2% with a sample size 1000 and for sample size 10000, the number of parameters d amounts to 97.6%, nearly all the generated data. The results for parameters d greater than 0.10 are similar (91.20% of the

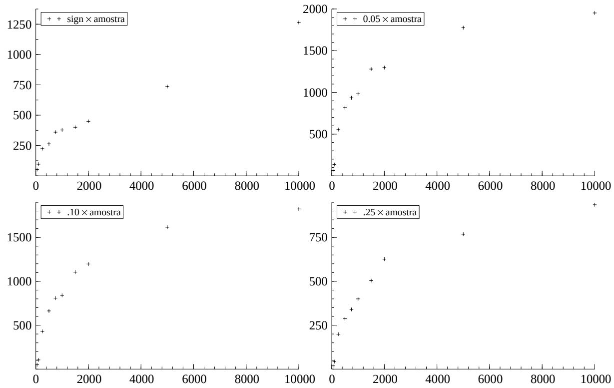

Figure 8: Convergence

processes have d greater than 0.10) and when we analyze parameters d greater than 0.25, the proportion decreases (46.80% in sample size 10000) but even so it is extremely high.

The mean for the estimated values for parameter d in the 2,000 replications of each process also grows with the increase in sample size, which reflects the results shown in the previous paragraph. The mean for sample size 50 is 0.01, reaches 0.12 in sample size 1000 and is 0.2454 in sample size 10000.

Interestingly enough, even though the mean for the estimated d values increases with sample size, the value of d was not greater than 0.5 (stationarity threshold for long memory processes) in any of the experiments carried out with this DGP specification. Thus, we can infer that with the increase of sample size to infinity, the mean for the estimated parameters d must converge towards a value close to but below 0.5. This is due to the fact that the DGP was always a stationary process, as observed by the eigenvalues of each autoregression. A graph showing the tendency of convergence of parameters is presented in figure 8.

The results of Monte Carlo experiment reveal that at least for the DGP used, corresponding to the estimated process for the exchange rate log return series by the MSIAH model, there is a great probability for the long memory estimated in section 1.3 to be spurious.

8 Conclusions

In this article, we analyzed whether the evidence of long persistence (long memory) found in the exchange rate log return series was robust. The estimation of ARFIMA model of Granger and Joyeux, which is the most widely used to model the long memory, obtained a significant parameter for the order of fractional integration, which would be consistent with the long, significant autocorrelations observed in the exchange rate log return series.

Because of conditional heteroskedasticity in the series, we carried out a series of procedures to verify whether the procedures for estimating the long persistence were robust in the presence of conditional heteroskedasticity.

The first step was to estimate an ARFIMA(p,d,q)-GARCH(1,1) model in which we jointly modeled the order of fractional integration with a GARCH model. The result of this procedure showed that the evidence of long memory was weak when we controlled heteroskedasticity in the sample. The existence of a long memory structure in this series, captured by the parameter of the order of fractional integration in the ARFIMA model, did not prove robust. Although the long memory models estimated without control over conditional heteroskedasticity were significant when we estimated ARFIMA models in which we had the autoregressive structure of the conditional variance of this series by means of a GARCH component, the parameter of fractional integration was no longer statistically significant.

Next, a two-step procedure was used in order to attempt to control heteroskedasticity by means of an MSH (Markov Switching Heteroskedasticity) model. First, we estimated an MSH model for the exchange rate log return series. Afterwards, we estimated an ARFIMA model for the standardized residuals of this model, which corrected the changes in the unconditional variance of the series. The results of this procedure showed that the evidence of long memory could not be sustained when we controlled heteroskedasticity by means of a Markov Switching model.

Finally, we carried out a series of Monte Carlo studies in order to confirm that there was a great probability of spurious long memory in the processes with changes in the unconditional variance controlled by a Markov chain. The probability of a Markov switching model inducing long memory on the series grew in a hyperbolic fashion with sample size. Therefore, we confirmed by a controlled experiment that, in addition to Granger's conclusion that the sum of AR(1) processes generates fractional integration processes, the sum of autoregressive processes with different unconditional variances controlled by a Markov Chan can generates spurius long memory. The Monte Carlo experiment showed that the processes with long memory were present in almost all the generated series, due to the increase in sample size, whereas the short memory process was the actual data-generating mechanism.

The existence of heteroskedasticity does not allow the analysis of the dependence structure in autocorrelations to be statistically reliable. This occurs due to the fact that the definition of significant autocorrelations by means of procedures such as Box-Pierce test are explicitly based on the hypothesis that variance is constant. The fact that variance is not constant in our series can make statistically weak dependence structures significant.

This is an even more serious problem when we observe that the joint tests for n autocorrelations are based on the fact that the squared sum of the square root of tested autocorrelations, multiplied by the sample size, have a chi-square distribution. Since our series consists of the sum of components from different distributions, the hypothesis of normality is violated and so is the hypothesis of a chi-square distribution (squared sum of normal standard distributions) used as distribution of autocorrelation tests.

By using the same reasoning, we observe that the estimation of the order of fractional integration in the ARFIMA model is not robust in the presence of heteroskedasticity. The Monte Carlo procedure showed that if we generate a short memory series but with Markov switching in unconditional heteroskedasticity, we induce a spurious long memory process. As the long memory process is defined as the existence of significant autocorrelations even at very distant intervals, the problem of heteroskedasticity and non-normality in returns is present in the significance test for the order of fractional integration parameter, which corresponds to the long memory process.

Such evidence shows that the long dependence structure between autocorrelations could be generated by a short memory process with structural breaks in the unconditional variance parameters; therefore, the existence of long memory in this series should be carefully considered.

Even though the literature on observational equivalence between Markov switching and long memory models is vast, our study shows that structural breaks originated by changes in the unconditional variance parameters of this series are the major component that triggers this behavior.

The final conclusions are that the long memory models of the ARFIMA class are not robust in the presence of heteroskedasticity and that the results found by these models in this situation should be carefully analyzed. It is still necessary to develop corrections for heteroskedasticity in the estimators of ARFIMA models.

This study allows us to infer that the persistence of shocks to the exchange rate log return series is quite low. The hypothesis that shocks to the exchange rate log return series have an effect after relatively long periods is not supported by the results of the procedures carried out in the present study.

Additional Notes

The ARFIMA models were estimated by means of the ARFIMA package by Marium Ooms and Jurgen Doornik for PcGive. We used the LMMod class for Ox by James Davidson to estimate the ARFIMA-GARCH models. We estimated the MSH model (Markov Switching Heteroskedasticity) by using MSVAR class 1.3.0 by Hans-Martin Krolzig for Ox language. The Monte Carlo study was programmed through the R language, with the use of the fracdiff package to estimate the ARFIMA model parameters.

References

[1] Ang, A. and Bekaert, G. (1998). "Regime Switches in Interest Rates". **Research Paper** 1486, Stanford University.

[2] Andersen, T. G., Bollerslev, T. Diebold, F. X. and Labys, P. (1999). "The Distribution of Exchange Rate Volatility". **NBER Working Paper** 6961.

[3] Baillie, R.T, Bollerslev, T. and Mikkelsen, H.-O. (1996). "Fractionally integrated generalized autoregressive conditional heteroskedasticity". **Journal of Econometrics** 74:3-30.

[4] Bollerslev, T. (1986). "Generalised Autoregressive Conditional Heteroskedasticity". **Journal of Econometrics** 31:307-327

[5] Bollerslev. T, Chou, R. and Kroner, K. (1992). "ARCH Modelling in Finance:A Review of Theory and Empirical Evidence". **Journal of Econometrics** 53:5-60

[6] Bollerslev. T. and Mikkelson, H.O. (1996). "Modelling and pricing long memory in stock market volatility". **Journal of Econometrics** 73:151-184

[7] Brock, W., Dechert. W. and Scheinkman, J. (1987). "A Test for Independence Based on the Correlation Dimension". Unpublished Working Paper , University of Wincosin at Madison, University of Houston and University of Chicago.

[8] Campbell, John Y., Lo, Andrew W. and MacKinlay, Archie G. (1997). **The Econometrics of Financial Markets**. Princeton University Press

[9] Davidson, J. (2002). "Moment and memory properties of linear conditional heteroskedasticity models". manuscript - http://www.cf.ac.uk/carbs/econ/davidsonje/hygarch4.pdf

[10] Dempster, A.P. , Laird, N.M. and Rubin, D.B. (1977). "Maximum Likelihood from Incomplete Data via the EM algorithm". **Journal of The Royal Statistical Society**, Series B, 39:1-38

[11] Ding, Z. and Granger, C.W.J (1996). "Modeling Volatility Persistence of Speculative Returns: a New Approach". **Journal of Econometrics** 73:185-215.

[12] Doornik, J.A. (2001). **Ox: An Object-Oriented Matrix Language** (4th edition), London: Timberlake Consultants Press

[13] Doornik, J.A. and Ooms,Marium (2000). "Multimodality and the GARCH Likelihood". manuscript

[14] Engel, Charles and Hamilton, James D. (1990). "Long Swings in the Dollar: Are They in the Data and Do Markets Know It?" . **American Economic Review** 80:689-713

[15] Engle, Robert. F. (1982). "Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of U.K. Inflation". **Econometrica** 46:987-1007

[16] Engle. R. and Bollerslev, T. (1986). "Modelling the persistence of conditional variances". **Econometric Reviews** 5:1-50.

[17] Franses, Philip Hans and Dijk, Dick van. (2000) . **Non-linear time series models in empirical finance**. Cambridge University Press

[18] Garcia, R. and Perron, P. (1990). "An Analysis of the Real Interest Rate under Regime Shifts". **Université of Montreal working paper**.

[19] Gouriéroux, Christian. (1997). **ARCH Models and Financial Applications**. Springer Series in Statistics. Springer

[20] Granger, C.(1980) "Long-memory relationships and the aggregation of dynamic models" **Journal of Econometrics**, 14, , 227-238.

[21] Granger. C. and Joyeaux (1980). "An Introduction to long memory time series models and fractional differencing". **Journal of Time Series Analysis** 1:15-19

[22] Hamilton, James D. (1994). **Time Series Analysis**. Princeton University Press

[23] Hamilton, James D. (1996). "Specification Testing in Markov Switching Time Series Models". **Journal of Econometric**s 70:127-157

[24] Hamilton, James D. and Susmel, Raul (1994). "Autoregressive heteroskedasticity and changes in regime". **Journal of Econometrics** 64:307-333

[25] Hansen, Bruce E. (1992). "The Likelihood Ratio Test under Non-Standard Condition: Testing the Markov Switching Model of GNP". **Journal of Applied Econometrics** 7:S61-82

[26] Hosking, J.R.M (1981). "Fractional differencing". **Biometrika** 68:165-178

[27] Kim, Chang-Jin, Nelson, Charles R. and Startz. R. (1998). "Testing for Mean Reversion in Heteroskedastic Data Based on Gibbs-Sampling-Augmented-Randomization". **Journal of Empirical Finance** 5:131-154

[28] Krolzig, Hans-Martin (1997). **Markov-Switching Vector Autoregressions: Modeling, Statistical Inference and Application to Business Cycle Analysis**. Lecture Notes in Economics and Mathematical Systems . Springer