

Simulation Evidence on Long Memory and Regime  
Switching in the Second Moment for Modelling of  
Financial Returns

Hans Isaksson

First Year Master Thesis

Department of Economics

Supervisor: Joakim Westerlund

To be defended on April 17, 2015.



**LUNDS UNIVERSITET**  
Ekonomihögskolan

## Abstract

It is well-known that long memory and regime switching in the first moment of a stochastic process are easily confused. But the relation between long memory in the second moment and regime switching in the second moment is less well understood. We perform a simulation study in which we assess the possibility to distinguish the two properties in the case that data with long memory in the second moment is generated as a FIGARCH(0,d,0)-process, and when regime switching is modelled by HMM with switching variance. Those model specifications are common in the modelling of financial returns, and conform to several well-known stylised facts of financial data. The simulation study lends evidence to the risk of confusing long memory and regime switching in the case studied.

**Keywords:** Long memory, Regime switching, HMM, FIGARCH, Simulation study, Returns

# Contents

- 1 Introduction** **1**
  
- 2 Statistical Models** **2**
  - 2.1 The Confusion of Long Memory and Regime Switching . . . . . 2
  - 2.2 The FIGARCH Model . . . . . 4
  - 2.3 HMM . . . . . 6
  
- 3 Simulation Study** **9**
  - 3.1 Simulation of FIGARCH data . . . . . 9
  - 3.2 Results from HMM estimation of FIGARCH Data . . . . . 11
  - 3.3 Parameter Estimation in FIGARCH models . . . . . 14
  
- 4 Conclusion** **18**

# 1 Introduction

In an influential paper by Diebold and Inoue (2001), it is argued that long memory of and regime switching in a stochastic process are properties that are easily confused in empirical economic and financial work. They even argue that they are "intimately related concepts" (p. 131) and that they are "effectively different labels for the same phenomenon" (p.157), at least if only a small amount of regime switching occurs. That is perhaps surprising as the models they consider appear to be very different, but the risk of confusing the two phenomena that they warned of was taken seriously by many authors. They support that claim by theoretical results and simulation evidence from a number of different models with the property of regime switching. Among others they consider models of fractional integration as a prototype for long memory, and the Hidden Markov Model (HMM) as a model of regime switching. However, to substantiate the claim that long memory and regime switching are effectively the same phenomenon, we should also expect series generated with long memory to be hard to distinguish from regime switching ones. They do not simulate models of long memory, however, to show that those are in a similar fashion easily misrepresented as models of regime switching. A simulation study in which a special case of this is done is presented in the working paper (Shi and Ho, 2014a), and another special case is presented in (Shi and Ho, 2014b). We therefore perform a simulation study in which we simulate series with long memory and examine the possibility of correctly rejecting regime switching models in favour of long memory ones in a case that we find particularly interesting for applications in finance.

To keep relevant to applications in financial modelling, we will examine models with long memory in the second moment, rather than in the first moment. This is because commonly accepted stylised facts regarding financial returns data state that there is no long memory in the mean of the returns series, but that there is long memory in squared returns (Cont, 2001). Since the empirical mean of daily returns series is close to zero, and

$$V(x) = E(x^2) - E(x)^2 \tag{1}$$

the squared returns series can be seen as an approximation of the conditional variance of the series. Having a good model for the conditional variance of financial returns is crucial in applications as it affects risk management (such as calculation of Value at Risk, or such as the Artzner et al. (1999) Expected Shortfall risk metric), as well as for the valuation of derivative assets (Hull and White, 1987). The model for long memory in the second moment that we will study is the Fractionally Integrated Generalised Auto Regressive Conditional Heteroscedasticity FIGARCH(0,d,0) model. Our model for regime switching is the two-state HMM in which the conditional variance is state-dependent. We chose those models based on their ability to reproduce stylised facts, while also being well-known in the literature.

We simulate data generated by a FIGARCH model, and fit FIGARCH as well as HMM models to the data. The aim is to assess whether there is a clear risk of confusing the two models in this case. The model selection procedure relies on residual analysis that is standard in small sample inference, as suggested by Diebold and Inoue (2001). This means that the distribution and dependence structure of the residual series are compared with what is expected under the model estimated. If the residual analysis can reject the estimated HMM model, but cannot reject the estimated FIGARCH model when fitted to simulated data, we argue that there is no clear risk of confusing the two phenomena in our case.

We find that our simulation evidence, restricted to our case, supports the claim by Diebold and Inoue (2001) that there is a risk of confusing long memory with regime switching in the second moment when applying standard small-sample inference. However, the claim that long memory and regime switching are so intimately related that they should be seen as the same phenomenon can be challenged. In Shi (2015) and Perron and Qu (2010), estimation procedures have been suggested which are claimed to be able to distinguish the two phenomena when generated by restricted families of models. In the simulation study at hand, we find preliminary results indicating there are cases in which the small sample inference can reject the HMM model in favour of the FIGARCH one. We argue that these results might be possible to formalize to a statistical test, similar to the one in Shi (2015).

The rest of the text will continue by presenting previous research on the risk of confusing long memory and regime switching, together with a presentation of the statistical models investigated as well as their relevance for modelling financial data in Section 2. In Section 3, the results from the simulation study are presented. We conclude in Section 4.

## 2 Statistical Models

In this section, we discuss previous results on the relation between long memory and regime switching. We will study long memory in the second moment, and our model for that will be the FIGARCH model. The model for regime switching in the second moment will be the HMM model with switching variance. Those are defined and the models' respective ability to reproduce stylised facts of financial returns data is discussed as well. We also present the parameter estimation procedure that will be employed when fitting those models to simulated data.

### 2.1 The Confusion of Long Memory and Regime Switching

Diebold and Inoue (2001) argue that long memory of and regime switching in a stochastic process are properties that are easily confused in empirical economic and financial work.

They even argue that they are "intimately related concepts" (p. 131) and that they are "effectively different labels for the same phenomenon" (p.157), at least if only a small amount of regime switching occurs.

They support that claim by theoretical results and simulation evidence from a number of different models with the property of regime switching. Their theoretical results are valid in some restricted cases, but show that there is a possibility of confusion between long memory and regime switching in those cases. In the simulation study, they apply statistical tests for long memory to data simulated from models with regime switching. The aim is to "characterize the finite-sample inference to which a researcher armed with a standard estimator of the long-memory parameter would be led" (p.141). The result is that the rejection frequency is well above the confidence level of the tests, indicating significant long memory. One of the models with regime switching from which they simulate data is a HMM in which the mean is a switching parameter, but the variance is not. They find that the empirical size of the tests depend on the length of the series estimated, as well as on the Markov transition probabilities. The dependence appears not to be a linear one.

However, they do not simulate models of long memory, to show that those are in a similar fashion easily misrepresented as models of regime switching. A simulation study in which a special case of this is done is presented in the working paper of Shi and Ho (2014a). That simulation study focuses on models with long memory and a heavy tail distribution, and fit HMM with t-distributed mixture components. The long-memory model from which they simulate data is a t-ARFIMA(0,d,0) model, that is a ARFIMA(0,d,0) model with innovations having a Students' t distribution. It is a well-known property of financial data that the distribution has fat tails (Cont, 2001), which is the motivation for their choice of model. A feature of empirical financial data that is not captured by the t-ARFIMA(0,d,0) model is long memory in squared returns. To reproduce that property, Shi and Ho (2014b) simulate data from a variant of a FIGARCH(1,d,1)-model to assess whether long memory in the second moment is easily confused with regime switching in the second moment. The data they simulate is not from the FIGARCH(1,d,1)-model defined by Baillie et al. (1996), but from a modification of it in which the errors are t-distributed.

The t-FIGARCH(1,d,1)-model is one which of relatively little use. We see no justification in letting  $p=q=1$  in the FIGARCH(p,d,q)-model, as this does not make the generated data conform better to the stylised facts of (Cont, 2001). We will therefore perform a simulation experiment similar to the one in Shi and Ho (2014b), but we will simulate data from a FIGARCH(0,d,0)-model with normal distributed errors. We furthermore want to examine if the confusion of long memory with regime switching in the second moment is a property that can be reproduced also with normal distributed errors.

While accepting the notion of a risk of confusing long memory and regime switching using standard modelling procedures and tests, Perron and Qu (2010) challenge the

interpretation of (Diebold and Inoue, 2001) that this means that the long memory and regime switching are "different labels for the same phenomenon" (Diebold and Inoue, 2001, p.157). Instead they propose a test that they claim can disentangle the two phenomena in the first moment, for a special case of model for long memory. In the same fashion, Shi (2015) propose a two-stage two-state ARFIMA model, and perform a simulation study which they claim shows that their proposed model can distinguish a pure-ARFIMA process from a pure HMM one.

## 2.2 The FIGARCH Model

A generalisation of the GARCH model (Bollerslev, 1986) is the FIGARCH( $p,d,q$ )-model (Baillie et al., 1996), in which  $p$  is the number of autoregressive (AR) lags,  $d$  is the  $\mathbb{R} \cap [0, 1)$ -valued<sup>1</sup> order of integration and  $q$  is the number of moving average (MA)-terms. It is defined for the process  $\{r_t\}_{t=-\infty}^T$  as

$$r_t = z_t \sigma_t \tag{2}$$

$$(1 - \alpha(B) - \beta(B))(1 - B)^d r_t^2 = \omega + (1 - \beta(B))\nu_t \tag{3}$$

where  $\{z_t\}_{t=-\infty}^T$  is an iid  $N(0, 1)$  stochastic process and  $\nu_t$  is defined as  $\nu_t = r_t^2 - \sigma_t^2$ .  $B$  is the back-shift operator and  $\alpha$  and  $\beta$  are polynomials in  $B$ , of order  $p$  and  $q$  respectively. The fractional difference operator is defined as

$$(1 - B)^d = \sum_{k=0}^{\infty} \frac{\Gamma(k - d)}{\Gamma(k + 1)\Gamma(-d)} B^k \tag{4}$$

and  $\Gamma$  denotes the Gamma function

$$\Gamma(t) = \int_0^{\infty} x^{t-1} e^{-x} dx \tag{5}$$

While introduced and defined by Baillie et al. (1996), the existence of a stochastic process satisfying the definition was not proven until Douc et al. (2008).

It is possible to interpret the polynomials  $\alpha$  and  $\beta$  as a model for the short term, so that it is the integration order  $d$  that controls the long memory of the process (Baillie et al., 1996). Our interest lie in the long run, and so we will set  $\alpha = \beta = 0$  when simulating data. That is, we will simulate FIGARCH(0,d,0) models.

The reason for our interest in long memory in the second moment is its application in modelling of financial returns. Cont (2001) presents an empirical overview of stylised facts of financial data, which we list below:

---

<sup>1</sup>The term "fractional" is a misnomer in this context. But it is standard in the econometrics literature, as well as in operator theory.

1. *Absence of autocorrelations.*
2. *Heavy tails.*
3. *Gain/loss asymmetry.*
4. *Aggregational Gaussianity.*
5. *Intermittency.*
6. *Volatility clustering.*
7. *Conditional heavy tails.*
8. *Slow decay of autocorrelation function in absolute returns.*
9. *Leverage effect.*
10. *Volume/volatility correlation.*
11. *Assymetry in time scales.*

The FIGARCH model presented above satisfies several of those stylised facts. The process  $\{r_t\}$  is unconditionally uncorrelated (Baillie et al., 1996), and the squared process  $\{r_t^2\}$  as well as the process for the absolute value,  $\{|r_t|\}$ , has a slowly decaying<sup>2</sup> autocorrelation function (Tayefi and Ramanathan, 2012). The unconditional distribution has heavy tails (Baillie et al., 1996). The model is defined such that the conditional variance is itself a random variable, and volatility clustering (the property that "large changes tend to be followed by large changes, of either sign, and small changes tend to be followed by small changes" (Mandelbrot, 1963, p.26)) is also a defining property of the model. This means that the stylised fact number 1,2,5,6,7 and 8 of Cont (2001) are satisfied. Stylised fact number 4, 10 and 11 deal with properties that are not defined by the stochastic process, or that concern continuous time and are thus not relevant for our purpose. Two of the stylised facts are, however, not obviously satisfied by the FIGARCH model. Stylised fact number 3 is that financial returns typically exhibit gain/loss-asymmetry, meaning that "one observes large drawdowns in stock prices and stock index values but not equally large upward movements" (Cont, 2001, p.224). To reproduce this stylised fact, a model with stochastic jumps could have been examined instead of the FIGARCH(0,d,0)-models. While interesting for applications in finance, doing so means that the model considered are quite different from the ones considered in Diebold and Inoue (2001), Shi and Ho (2014a) and Shi (2015), however. We therefore avoid that altogether. Stylised fact number 9, the "leverage effect", is that "most measures of volatility of an asset are negatively correlated with the returns of that asset" (Cont, 2001, p.224). A generalization of the FIGARCH

---

<sup>2</sup>The rate of decay is hyperbolic.



model that can capture this effect is proposed by Hwang (2001), but Ruiz and Perez (2003) show that model to be ill specified, and proposes another specification. This was not explored further.

An alternative to the model specified above for example is to let  $z_t$  in equation (2) to be t-distributed instead of  $N(0, 1)$  distributed (Bollerslev, 1987) or normal-inverse Gaussian (NIG)-distributed (Kiliç, 2007). By such distributional assumptions, it is argued that the model can fit empirical data better by allowing for fatter tails of the distribution. However, the number of parameters to estimate also increases. In the case of t-distributed error terms for example, a degrees of freedom-parameter is to be estimated from data. We will study models with normal distributed errors only.

Different approaches to estimating the parameters in a FIGARCH model have been proposed (Tayefi and Ramanathan, 2012). We follow the approach by Baillie et al. (1996), which is to maximise the likelihood function in (6). The log likelihood function can be derived from an  $ARCH(\infty)$ -representation of (2).

$$\log L(\theta, r_1, r_2, \dots, r_T) = -0.5T \log(2\pi) - 0.5 \sum_{t=1}^T (\log(\sigma_t) + r_t^2 \sigma_t^{-1}) \quad (6)$$

Each observation is dependent on the entire history of a realisation, and so there are out-of-sample conditional variances that are unavailable in finite-sample inference. Those are set to the unconditional sample variance in the estimation procedure. This is standard procedure when estimating the parameters of GARCH-family models and is suggested in Baillie et al. (1996) and Tayefi and Ramanathan (2012) also for FIGARCH models.

The infinite sum (4) has to be truncated to perform simulation of FIGARCH-series or when estimating parameters. This is commonly done at the 1000th term. The expression (4) for the fractional difference operator was introduced by Hosking (1981) and is the one that is given in Baillie et al. (1996) and that is standard in time series analysis. But there are also alternatives that could have been used, such as defining it through the Laplace transform. Doing so leads to an equivalent operator. But it is not obvious that the truncation of operator (4) is equivalent to a truncation of the Laplace-fractional difference operator. We stick to the procedure of Baillie et al. (1996) since it is standard in the literature. But it might be an interesting question for future work to examine the properties of the alternative operator.

## 2.3 HMM

In a HMM (Hamilton, 1988, 1989)<sup>3</sup>, the probability distribution of the stochastic process at a time  $t$  depends on an underlying and unobserved Markov process. The HMMs that

---

<sup>3</sup>Hidden Markov Models are also named Markov Regime Switching models.

will be estimated are defined as

$$r_t = \beta_{S_t} + \varepsilon_{t,S_t} \quad (7)$$

$$\varepsilon_{t,S_t} \sim N(0, \sigma_{S_t}^2) \quad (8)$$

where  $S_t \in \{1, \dots, k\}$  is the (unobserved) states of the Markov chain  $\{S_t\}_{t=1}^{\infty}$ , and where  $k$  is the number of states. The transition probability  $p_{ij}$  is defined by

$$p_{ij} = P(S_{t+1} = j | S_t = i) = P(S_{t+1} = j | S_t = i, S_{t-1} = k, \dots) \quad (9)$$

and the second equality is the definition of the Markov (chain) property. The parameters  $\beta$  and  $\sigma^2$  are said to be non-switching if they are restricted to be equal in any state, that is they do not depend on  $S_t$ . Otherwise the parameters are said to be switching. A model with switching  $\beta$  and switching  $\sigma^2$  will be estimated.

One feature that makes HMM models popular to practitioners is that the states are seen as easily interpretable, at least as long as the number of states is small. For example, in a two state model for financial applications, they are sometimes interpreted as one state representing "normal times" and the other state representing some kind of "financial stress situation". As an example of the interpretation of what the states represent, Hamilton (1988) applies a two-state HMM model to the US Treasury bill interest rate. He finds that "[t]he period 1979:4-1982:3 is thus identified as a time of dramatically higher and more volatile short-term interest rates than that seen before or since" (p.408), and writes that "[t]his dating of an apparent shift in the process for the short-term interest rate corresponds precisely with a profound change in Federal Reserve operating procedures" (p.408). This means that his interpretation of the different states is that the low volatility state represents the normal operating procedure of the Federal Reserve, and that the high volatility state represents a different operating procedure that was conducted during a period of three years. Another property that makes HMM useful for applications is that the Markov property for the state process greatly simplifies forecasting from them. The reason for that is that the Kolmogorov-Chapman equation can be used for simplifying the computation of the n-step ahead transition probabilities, which are in turn needed for the computation of the n-step ahead forecast of  $r_t$ . Let the transition probabilities be collected in a matrix  $P$ , and the n-step ahead transition probabilities in a matrix  $P^{(n)}$ . Then  $P^{(n)} = P^n$ , which is a special case of the Kolmogorov-Chapman equation (Shiryayev, 1995, p.116).

Rydén et al. (1998) examine if HMM models can generate series that have the properties that are described by the stylised facts of financial data. They compare the models with a different set of stylised facts than those in (Cont, 2001), but their conclusion is that the HMM models can reproduce all of the stylised facts examined apart from one:

the form of the decay of the autocorrelation function. While they are able to reproduce a slow decay, the form of it is not the same as in the stylised facts. Nystrup et al. (2015) suggest that continuous-time HMM can give a better fit to daily financial returns data while restricting the number of parameters to estimate. The number of parameters in a continuous-time HMM grows linearly with the number of states, rather than quadratically which is the case for the HMM (Nystrup et al., 2015). In particular, they show that the slow decay of the empirical autocorrelation function from a set of real-world financial returns data is better captured by a continuous-time HMM than with a HMM with a comparable number of parameters. Of the stylised facts presented in Cont (2001), it is easy to see that a HMM with switching variance can reproduce fact number 2 that the unconditional distribution has heavy tails. That is because the distribution of the errors of a HMM is a mixture distribution of normal distributions, which is a convex combination of distributions. An so, depending on model parameters, an error distribution with fat tails can be reproduced. See figure 1 for an example of a mixture probability density function. Fact number 6, Volatility clustering, is also a feature of HMM with switching variance. That is an immediate consequence of the definition, as the variance at  $t$  depends on what state is realised at  $t$ .

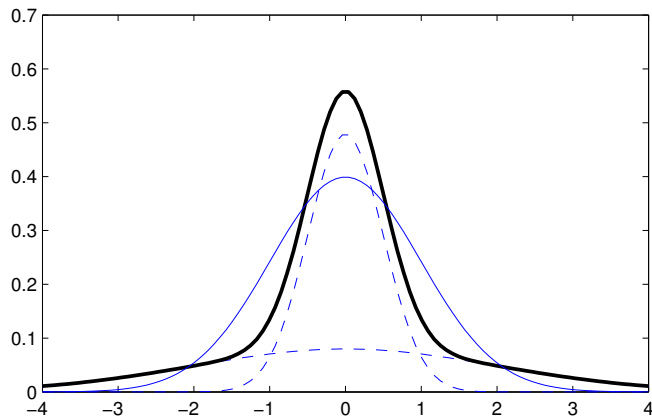


Figure 1: Probability density function of a mixture distribution (thick, black) with heavy tails. A  $N(0,1)$ -pdf (thin, blue) for comparison. The scaled mixture components are depicted as dashed blue lines.

Estimation of HMM can be done either by optimisation of the likelihood function, or through Gibbs-sampling. Gibbs-sampling is a Markov Chain Monte Carlo (MCMC) method. Optimisation of the likelihood function is implemented here, and the log likelihood function is given by

$$\log L(\theta, r_1, r_2, \dots, r_T) = \sum_{t=1}^T \log \sum_{j=1}^k (f(r_t | S_t = j, \theta) \Pr(S_t = j)) \quad (10)$$

where  $f(r_t|S_t = j, \theta)$  is the likelihood for state  $j$  conditional on the parameter vector  $\theta$ . The probabilities  $\Pr(S_t = j)$  are latent and estimated with Hamilton's filter (Hamilton, 1988, 1989). This is done before the optimisation of (10). The HMM-estimation of the simulated data is performed using the MS Regress Package of Perlin (2014), implemented in Matlab. Both HMM-models where the mean and the variance are switching variables and models where only the variance is switching between regimes are estimated.

The number of parameters to estimate increases quadratically with respect to the number of states  $k$ , which means that computation time increases dramatically with  $k$ . Because of that, it is not deemed feasible to estimate models with more than two states. The likelihood function for the two state model is then

$$\log L(\theta, r_1, r_2, \dots, r_T) = \sum_{t=1}^T \log \sum_{j=1}^2 (f(r_t|S_t = j, \theta) \Pr(S_t = j)) \quad (11)$$

$$= \sum_{t=1}^T \log \sum_{j=1}^2 \left( \frac{1}{\sqrt{2\pi}\sigma_j} \exp\left(-\frac{(r_t - \mu_j)^2}{2\sigma_j}\right) \Pr(S_t = j) \right) \quad (12)$$

### 3 Simulation Study

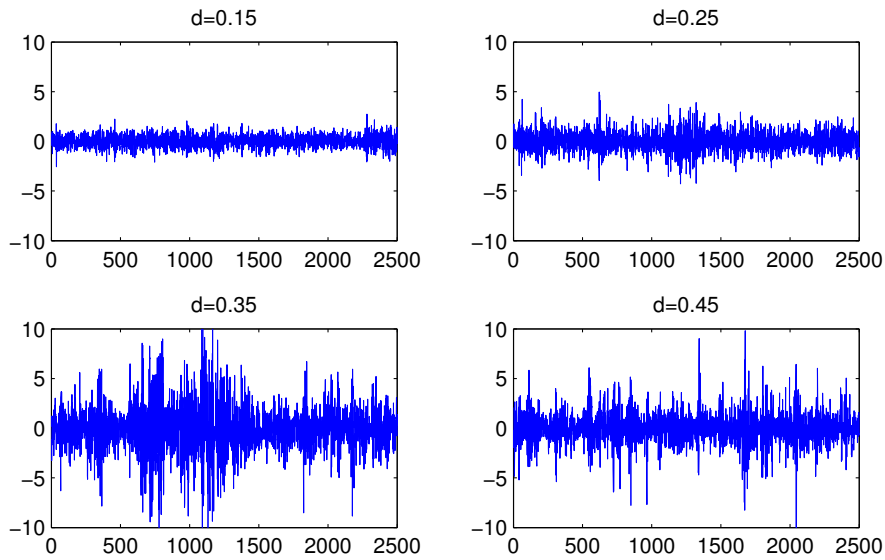
We simulate data from a FIGARCH(0,d,0) process. Then we fit HMM with switching variance as well as FIGARCH(0,d,0)-models to the data. The distribution and dependence structure of the residual series from the estimation procedures are analysed to assess the risk of not being able to reject the incorrect HMM specification in favour of the correct FIGARCH(0,d,0) specification.

#### 3.1 Simulation of FIGARCH data

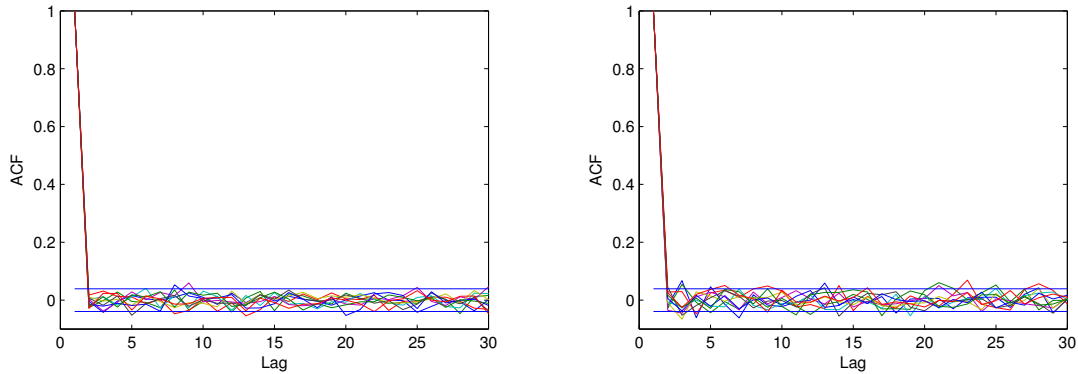
Data from FIGARCH(0,d,0) models are simulated. The choice of setting the AR and the MA parameter to zero is to conform to the stylised facts of Cont (2001) regarding empirical financial returns. We simulate series with fractional integration order  $d$  equal to 0.15, 0.25, 0.35 and 0.45, following Shi and Ho (2014a). The process has well defined variance only if  $d < 0.5$ . The length of the simulated series is chosen to 7000, and we then truncate all but the last 2500 points away in each series. The aim of the truncation is to remove entries that are affected by initial value bias as a result of the truncation of the operator in equation (4). The chosen series length corresponds to roughly ten years of daily trading data, calculated at 252 trading days per year. In Diebold and Inoue (2001), the length of the simulated series range between 100 and 5000, and so 2500 is in the middle of that interval. Investigating the effect of the length of the series is outside the scope of the present simulation study. The number of simulated series is 5000, which is chosen so that computations can be done in a feasible length of time. We set  $\omega = 0.1$

in all simulations, following Shi and Ho (2014b). The simulation is done with the MFE Toolbox of Sheppard (2014), implemented in Matlab.

To build some intuition for the data, figure 2 shows one series from each FIGARCH(0, $d$ ,0) model. Volatility clustering appears to be more apparent with rising fractional integration order  $d$ . No autocorrelation is expected in the simulated series. This is supported by the autocorrelation plots in figure 3 where the autocorrelation functions for ten simulations with  $d = 0.15$  and  $d = 0.45$  are shown. In figure 4 the same plot is presented for the square of the simulated series. We can see that there are a large number of significant lags in the autocorrelation function of the squared series. The number of significant lags increases when  $d$  is closer to 0.5. This means that the simulated data appears to have an autocorrelation structure that is consistent with the stylised facts in Cont (2001).



*Figure 2: Simulated FIGARCH series*



*Figure 3: Autocorrelation function of simulated series,  $d = 0.15$  (left) and  $d = 0.45$  (right)*

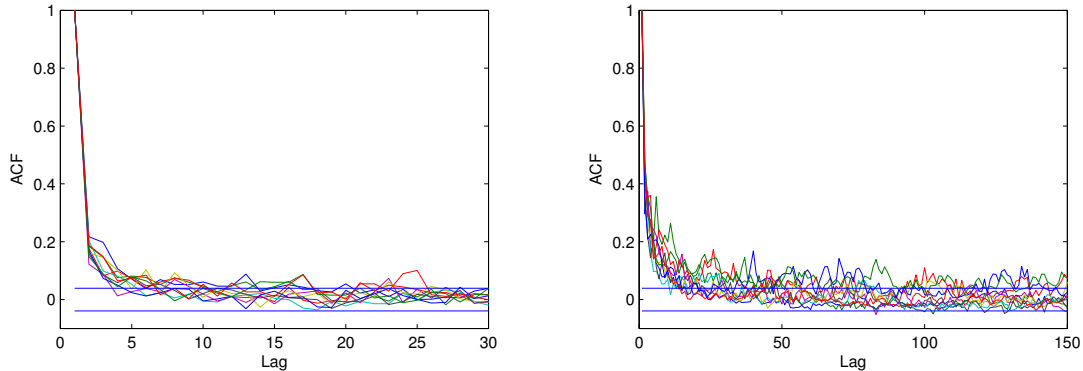


Figure 4: Autocorrelation function of squared simulated series,  $d = 0.15$  (left) and  $d = 0.45$  (right). Note the difference in scale of the x-axis.

### 3.2 Results from HMM estimation of FIGARCH Data

Two state HMM in which the variance and the mean are allowed to switch are fitted to the simulated data, and residual diagnostics are performed on the estimated residuals. Since the estimation of HMM is more computationally intensive than the estimation of FIGARCH models, only the first 1000 of the simulated series are estimated, rather than 5000 which is the case in the estimation of FIGARCH models. This means that the models are not fitted to exactly the same data series, which could be an error source in the Monte Carlo experiment. The alternative would have been to restrict also the number of series to which the FIGARCH model is fitted to 1000.

Estimation of a HMM is computationally expensive. The computations were made in parallel on computers with Intel Core i7 processors, and the total run time when performing the parameter estimations was about three days (70 hours). The HMM-estimation of the simulated data was performed using the MS Regress Package of Perlin (2014), implemented in Matlab. The parallelisation of the code was made by splitting the simulated data in four subsets, one with each value of the parameter  $d$ . After that, four PC's were performing the model estimation procedures on all series in one subset each, and the results were sent back to one PC which collected them to perform residual tests. The most time-consuming part of the computation appears to have been the implementation to each series of Hamilton's filter.

$d$	Mean $p_{11}$	Mean $p_{22}$	Mean $\beta_1$	Mean $\beta_2$	Mean $\sigma_{\epsilon,1}^2$	Mean $\sigma_{\epsilon,2}^2$
0.15	0.9699	0.9359	-0.0005	-0.0007	0.2550	0.6112
0.25	0.9741	0.9364	-0.0008	0.0002	0.5164	1.8027
0.35	0.9760	0.9382	-0.0006	-0.0004	0.9371	5.1673
0.45	0.9761	0.9345	0.0010	0.0026	1.3788	14.8580

Table 1: Parameter estimation results from estimated HMM to simulated FIGARCH data.

$d$	Mean Skew	Mean Absolute Skew	Mean Kurtosis	Mean RSS
0.15	0.0016	0.0436	3.2823	898
0.25	0.0032	0.0712	4.1345	2155
0.35	-0.0039	0.1281	5.9987	5176
0.45	0.0112	0.2584	9.9728	13112

Table 2: Diagnostic statistics from HMM-estimation of FIGARCH series.

$d$	$ARCH_{10}$	$Q_{10}$	$BDS_{10}$
0.15	1.000	0.095	1.000
0.25	1.000	0.211	1.000
0.35	1.000	0.477	1.000
0.45	1.000	0.683	1.000

Table 3: Rejection frequencies from residual tests from HMM-estimation of FIGARCH series.

In table 1 the estimated parameters are reported for the estimation of (7) in which  $\beta$  and  $\sigma$  are both switching. We can see that the switching probabilities  $p_{ii}$  are all above 0.9, and that the  $p_{11}$  parameters rise marginally with  $d$ . This means that the conditional probability of remaining in state 1 when starting in state 1 rises, and that there will be fewer regime shifts. This pattern holds also for  $p_{22}$ , except when  $d = 0.45$ .

The columns with the mean of  $\sigma_{\epsilon_i}^2$  present the mean of the estimated variances in state 1 and 2 respectively. We see that the difference between them is more pronounced with larger  $d$ , and also that the variances rise with  $d$ . The series with higher fractional integration parameter  $d$  are expected to have more pronounced volatility clustering, as was indicated in figure 2. This larger amount of volatility clustering when  $d$  is large is assumed to be the reason for the larger difference in estimated volatility between the states.

We turn now to residual analysis. If the model is well specified, the errors from the estimated models are independent and identically normal  $N(0, 1)$  distributed. We analyse first the distribution of the errors, and then perform standard residual tests to see if there is any dependence in the residual series. When we analyse the distribution of the errors, the residual series should have no skew, zero mean and a kurtosis equal to 3. In table 2, residual diagnostics are reported. Each parameter is the average of the parameters calculated from the residual series from the fitting of the HMM model with switching  $\beta_{S_t}$  and variance  $\sigma_{S_t}$ . We present the absolute skew as well as the skew to be able to detect any cancellation effect. If the estimator of the skew is normal distributed, then the absolute value of the skew has a folded distribution. The table indicates that the residual series tend to have a kurtosis larger than 3, which means that the series have fat tails. The kurtosis is more pronounced with larger  $d$ , that is when there is more long memory in the

second moment. There appears also to be some skew, in particular when the fractional integration order  $d$  is larger. This means that the estimated model could be claimed not to fit the data very well. It is possible that estimating a HMM model with three states instead of two would generate residuals with a distribution closer to a normal distribution. The reason for that is that a mixture distribution with three mixture components can fit even fatter tails than a two-state HMM. We have not estimated three-state HMMs, as the computational burden would be prohibitive. In the present experiment, we simulate the data as FIGARCH(0,d,0), and so unless that model is in some sense equivalent to a HMM model, we estimate a misspecified model. Then the result should not be too surprising. The purpose of the simulation experiment is to assess whether there is a risk of mistaking the FIGARCH data for HMM data. Therefore we will, for the sake of comparison, also have to do this same analysis on residuals from estimated FIGARCH models, which we do in section 3.3.

We continue now with analysing the dependence structure in the residual series by performing standard residual tests. In table 3, the rejection frequencies from three standard residual tests are reported. We use them to assess whether the residual series have the property that their entries are iid. The ARCH Lagrange Multiplier (Engle, 1982) test ( $ARCH_{10}$ ) at lag length 10 is a test for remaining conditional heteroscedasticity. We see that the null hypothesis of no remaining heteroscedasticity is rejected in each residual series. Hence, the test lends evidence against the estimated model. The next test, the Ljung-Box (Ljung and Box, 1978) test ( $Q_{10}$ ) is a portmanteau test of no remaining autocorrelation. The rejection frequency appears to depend on the parameter  $d$ , with the rejection frequency rising with  $d$ . This is in line with the results on the distributional fit, in which we noticed that the distribution is more leptokurtic when  $d$  is large. Lastly, the Brock Dechert Scheinkman (Brock et al., 1987, 1996) test ( $BDS_{10}$ ) is sometimes applied to test for general nonlinearity as well as for general mis-specification. The computation of the test statistic was carried out by using a program provided in Kanzler (1999). It was chosen in favour of implementing the C-program by (Brock et al., 1996), as it is numerically more efficient than that one, while implementing the same test. No attempt to use small-sample correction is made as the difference is expected to be small given the length of the estimated series. The rejection frequency is again 1, that is the null hypothesis that the series are iid is rejected for each residual series. The choice of setting the lag lengths and the embedding dimension to ten in the residual tests was made to agree with Shi and Ho (2014a). Choosing a shorter lag length and embedding dimension does not appear to affect the result. We did not try to perform the tests with a very large number of lags because of the computation time that would have been needed for doing so. In figure 5, a panel of simulation convergence plots for the Ljung-Box Q test is presented. It demonstrates that the simulation experiment presented in table 3 appears to have converged. Each plot depicts the cumulative rejection frequency. If the cumulative



frequency appears to stabilise, this is an indication that the simulation has converged. The conclusion is then that the number of simulated series is large enough. If it would not converge, that would mean that the simulation experiment is of little value, perhaps because too few series have been simulated.

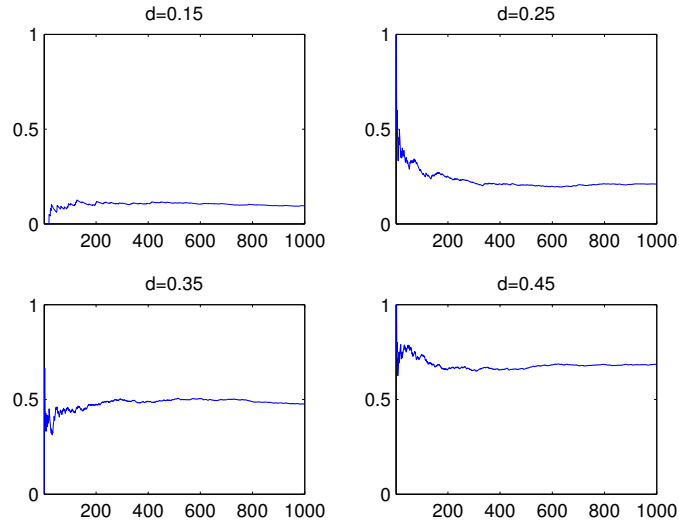


Figure 5: Simulation convergence plot depicting the cumulative rejection frequency of the Ljung-Box  $Q$ -test applied to residuals from HMM-estimation of FIGARCH data.

The overall view from the residual diagnostics in table 2 and from the residual tests in table 3 is that fitting a HMM-model to simulated FIGARCH(0,d,0)-series lead to serious problems that are obvious when inspecting the residual series. The residuals do not have a normal distribution, as they have excess kurtosis. Moreover, the residual tests indicate that there is dependence in the residual series.

We will continue by fitting a FIGARCH-model to the simulated series for comparison.

### 3.3 Parameter Estimation in FIGARCH models

We fit FIGARCH(0,d,0) models to the simulated FIGARCH(0,d,0) data, for the purpose of comparison with the residual diagnostics that is performed on the estimated HMM-models. In particular, we will examine the distributions of the residual series, and perform residual tests to examine the iid-ness of the residual series. The parameter estimation was implemented by the MFE Toolbox of Sheppard (2014), implemented in Matlab. The estimation procedure was less time consuming than the one in which HMM were fitted, and we did not parallelise the code.

The estimated parameters are close to the parameters set in the simulation. In table 4, diagnostic statistics are reported. We see that the skew is small, and in particular smaller than when a HMM model is estimated. There are, however, indications of leptokurtosis

in the residuals, in particular when the fractional integration order is closer to 0.5. The Residual Sum of Squares (RSS) is lower than in the HMM estimation. Since the data is simulated as FIGARCH(0,d,0)-models, the model is not misspecified as was the case in the previous section. The source of deviation from the correct distribution is instead likely to be a finite-sample phenomenon. Each realisation evaluated at  $t$ ,  $r_t$ , is dependent on the entire history of the realisation. But when the model is fitted to finite-sample data, out-of-sample conditional variances are unavailable. In the estimation procedure, they were set to the unconditional sample variance, following Baillie et al. (1996). This effect is stronger when  $d$  is closer to 0.5, as the long-memory effect is more pronounced then. That is in line with the observation that the kurtosis in the residual series is larger when  $d$  is large. This source of error appears to affect also the dependence structure in the residuals, as we will see.

As in section 3.2, we now turn to residual tests to examine the existence of dependence in the residual series. In each of the tests and for each  $d$ , the theoretical rejection frequency of the tests is 0.05 if the data is drawn from a FIGARCH(0,d,0) process. The results are reported in table 5. We see that the ARCH LM test and the BDS test does not always reject the null hypothesis, as was the case when applied to the residuals from the HMM estimations. The rejection frequency of the ARCH LM test is still very high, though. This could be a sign of the difficulty mentioned above of estimating FIGARCH models. It could also indicate that the ARCH LM residual test is of little value when deciding between a HMM and a FIGARCH specification when analysing financial returns data.

$d$	Mean Skew	Mean Absolute Skew	Mean Kurtosis	Mean RSS
0.15	-0.0014	0.0442	3.1695	3404
0.25	-0.0006	0.0520	3.4882	888.1
0.35	-0.0002	0.0659	4.0004	2673
0.45	-0.0006	0.0818	4.5620	9299

Table 4: Diagnostic Statistics from FIGARCH-estimation of FIGARCH series.

$d$	$ARCH_{10}$	$Q_{10}$	$BDS_{10}$
0.15	0.9948	0.0334	0.9864
0.25	0.9532	0.0440	0.8328
0.35	0.8818	0.0680	0.1572
0.45	0.9486	0.0902	0.0772

Table 5: Rejection frequencies from residual tests from FIGARCH-estimation of FIGARCH series.

In the Ljung-Box test, the null hypothesis is rejected less often than is the case when applied to the HMM residuals, and relatively close to the correct frequency of an iid process. The rejection frequency is closest to the correct frequency when  $d$  is between 0.25

and 0.35. The rejection frequency of the BDS test appears to depend on the fractional integration order  $d$ , and it appears to be more reliable when  $d$  is large. When  $d = 0.45$ , the empirical rejection frequency is 0.0772, which is not far from the correct theoretical frequency. As in the previous section, we present simulation convergence plots that demonstrate the convergence of the simulation experiment. The simulation convergence plots for the ARCH LM test is presented in figure 6, the plots for the Ljung-Box Q-test in figure 7, and the plots for the BDS-test in figure 8.

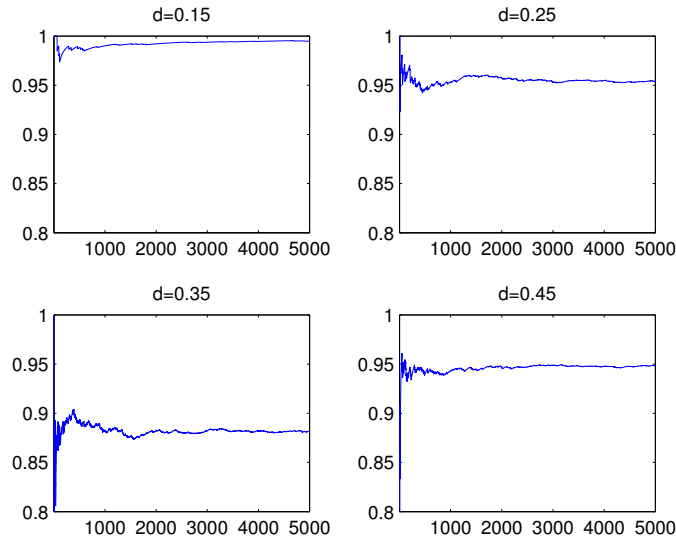


Figure 6: Simulation convergence plot depicting the cumulative rejection frequency of the ARCH LM test applied to residuals from FIGARCH-estimation of FIGARCH data.

Our aim is to evaluate the risk of confusing long memory in the second moment with regime switching in the second moment. When evaluating that risk, the method used by Diebold and Inoue (2001) was to characterise "the finite-sample inference to which a researcher armed with a standard estimator of the long-memory parameter would be led" (p.141). We use the same method, applied to long memory in the second moment as in Shi and Ho (2014b). We argue, based on the simulation results, that such a researcher would see two models whose distribution deviate from the one expected under the respective correct model. That deviation is larger when the estimated fractional integration order  $\hat{d}$  is small. Furthermore, the researcher would conclude that under both models, the residuals are non-iid. That conclusion would follow as it is likely that at least two of the presented residual tests reject its respective null hypothesis of no heteroscedasticity, no autocorrelation or no nonlinear dependence. The probability of rejecting iid-ness in residuals is large when the estimated  $\hat{d}$  is small. This means that our results are in line with the conclusion in Diebold and Inoue (2001), Shi and Ho (2014a) and Shi and Ho (2014b), that long memory and regime switching can be easily confused. We thus add the FIGARCH(0,d,0) model, which has long memory in the second moment to the list of

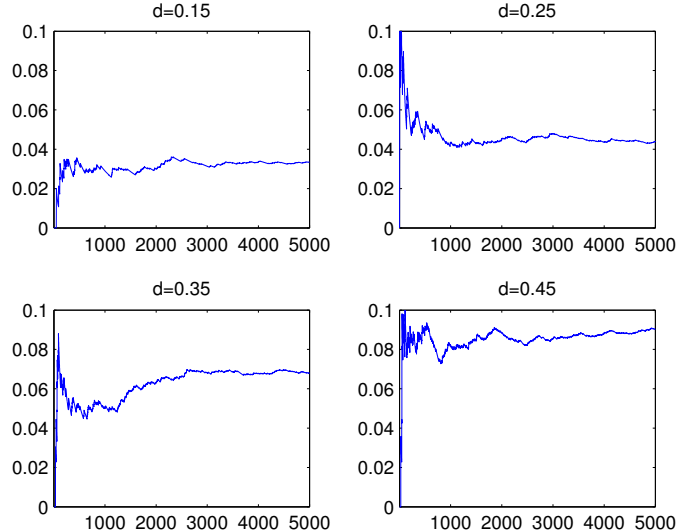


Figure 7: Simulation convergence plot depicting the cumulative rejection frequency of the Ljung-Box  $Q$ -test applied to residuals from FIGARCH-estimation of FIGARCH data.

model specifications for which this empirical feature has been demonstrated. The previous results regarded models with long memory in the first moment, as well as models with long memory in the second moment generated as a t-FIGARCH(1,d,1)-process.

Among the main possible sources of error in the study which restricts the interpretation of the results, we mention again that all data is simulated with  $\omega = 0.1$ . This is not expected to affect the results, but we have not examined that. To examine that, a replication of the simulation study for different values of  $\omega$  could be performed. Similarly, we have fixed the length of the simulated series to 2500. A replication of the simulation study with other lengths of the series would make the results more widely applicable.

While the above results support the risk of confusing long memory and regime switching when using standard estimation procedures, we agree with Perron and Qu (2010) and Shi (2015) and argue that this need not preclude the possibility of developing testing procedures that can disentangle the two phenomena. We propose based on our simulation results that there is a situation in which the Ljung-Box residual test could lend evidence to a FIGARCH model specification in favour of a HMM specification. That is the case when the estimated  $\hat{d} \geq 0.35$  in the FIGARCH model specification and the null hypothesis of the Ljung-Box residual test is not rejected when applied to the FIGARCH residuals, while it is rejected when applied to the HMM-residuals. The test is then a conservative one, built on the results from the present simulation study. In the case that the estimated  $\hat{d} \leq 0.25$ , the test give no guidance however. If the rejection frequency depends monotonically on the fractional integration order  $d$ , it is possible that extending the present simulation experiment to the case  $d = 0.30$  would extend the parameter space in which the above testing procedure would be applicable. In a situation similar to the

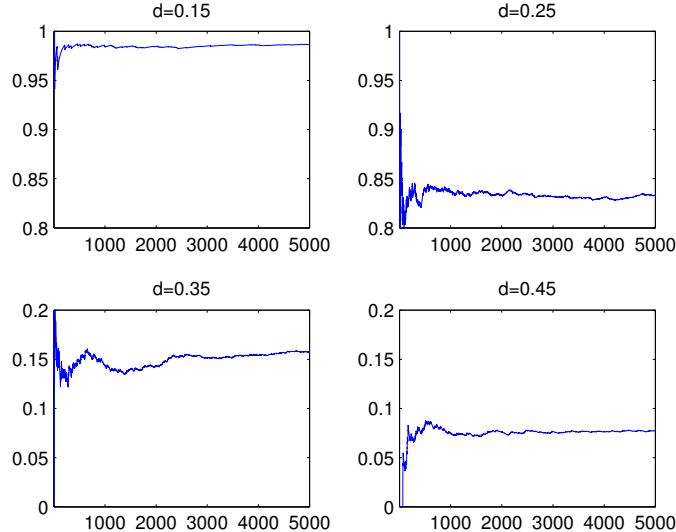


Figure 8: Simulation convergence plot depicting the cumulative rejection frequency of the BDS-test applied to residuals from FIGARCH-estimation of FIGARCH data.

one just described, the BDS-test could be used to lend evidence to the FIGARCH model. These observations could be used to formulate tests that can be used to distinguish long memory from regime switching in the second moment, as in Shi (2015). We also see that the ARCH LM-test might be unreliable. This could be an artifact pointing towards simulation problems, for example stemming from the truncation of operator in (4). If that is so, the reason for it should be examined and software procedures should be corrected. Otherwise, the ARCH LM-test should be used with care when examining models with long memory.

## 4 Conclusion

In an influential paper by Diebold and Inoue (2001) it was argued that there is a clear risk of confusing long memory and regime switching using standard estimation procedures and tests. We have performed a simulation study which extends the models for which this has been demonstrated to data simulated as a FIGARCH(0,d,0)-process with normal distributed error terms, which is a model of long memory in the second moment. Previously, this has been demonstrated by simulating data with regime switching in the first moment, by simulating data with long memory in the first moment, as well as by simulating data with long memory in the second moment generated as a t-FIGARCH(1,d,1)-process. We furthermore show that the theoretical models studied are relevant for applications in modelling of financial returns data since they both satisfy many of the stylised facts that have been observed by Cont (2001) and Rydén et al. (1998). Furthermore, we identify

properties of the tests used in the simulation study that can possibly serve as a starting point for finding an estimation procedure which can be used to distinguish long memory and regime switching, similarly to Perron and Qu (2010) and Shi (2015).

## References

- P Artzner, F Delbaen, J-M Eber, and D. Heath. Coherent measures of risk. *Mathematical Finance*, 9(3):203–228, 1999.
- RT Baillie, T Bollerslev, and HO Mikkelsen. Fractionally integrated generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 74(1):3 – 30, 1996. ISSN 03044076.
- Tim Bollerslev. Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 314(3):307–327, 1986.
- Tim Bollerslev. A conditionally heteroskedastic time series model for speculative prices and rates of return. *The review of economics and statistics*, pages 542–547, 1987.
- W.A. Brock, W.A. Dechert, and J.A. Scheinkman. A test for independence based on the correlation dimension. Mimeo, 1987.
- W.A. Brock, W.D. Dechert, J.A. Scheinkman, and B. LeBaron. A test for independence based on the correlation dimension. *Econometric Reviews*, (15):197–235, 1996.
- Rama Cont. Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative Finance*, 1(2):223 – 236, 2001. ISSN 14697688.
- Francis X. Diebold and Atsushi Inoue. Long memory and regime switching. *Journal of Econometrics*, 105(Forecasting and empirical methods in finance and macroeconomics): 131 – 159, 2001. ISSN 0304-4076.
- Randal Douc, François Roueff, and Philippe Soulier. On the existence of some ARCH( $\infty$ ) processes. *Stochastic Processes and their Applications*, 118:755 – 761, 2008. ISSN 0304-4149.
- Robert F. Engle. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, (4):987–1007, 1982. ISSN 00129682.
- James D. Hamilton. Rational-expectations econometric analysis of changes in regime. *Journal of Economic Dynamics and Control*, 12(2/3):385, 1988. ISSN 01651889.
- James D. Hamilton. A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica*, (2):357, 1989. ISSN 00129682.
- J. R. M. Hosking. Fractional differencing. *Biometrika*, (1):165, 1981. ISSN 00063444.
- John Hull and Alan White. The pricing of options on assets with stochastic volatilities. *The Journal of Finance*, 42(2):281–300, 1987. ISSN 1540-6261. doi: 10.1111/j.1540-6261.1987.tb02568.x.

- Y. Hwang. Asymmetric long memory GARCH in exchange return. *Economics Letters*, 73(1):1–5, 2001. ISSN 01651765.
- Ludwig Kanzler. Very fast and correctly sized estimation of the BDS statistic. Online paper, 1999. URL <http://ssrn.com/abstract=151669>.
- R. Kiliç. Conditional volatility and distribution of exchange rates: GARCH and FI-GARCH models with NIG distribution. *Studies in Nonlinear Dynamics and Econometrics*, 11(3), 2007. ISSN 10811826.
- G. M. Ljung and G. E. P. Box. On a measure of lack of fit in time series models. *Biometrika*, (2):297–303, 1978. ISSN 00063444.
- Benoît Mandelbrot. The variation of certain speculative prices. *The Journal of Business*, (4):394, 1963. ISSN 00219398.
- Peter Nystrup, Henrik Madsen, and Erik Lindström. Stylised facts of financial time series and hidden Markov models in continuous time. *Quantitative Finance*, 2015. doi: 10.1080/14697688.2015.1004801. Forthcoming.
- Marcelo Perlin. MS regress - the Matlab package for Markov regime switching models. Mimeo, 2014. URL <http://dx.doi.org/10.2139/ssrn.1714016>.
- P. Perron and Zhongjun Qu. Long-memory and level shifts in the volatility of stock market return indices. *Journal of Business and Economic Statistics*, 28(2):275 – 290, 2010.
- Esther Ruiz and Ana Perez. Asymmetric long memory GARCH: A reply to Hwang’s model. *Economics Letters*, 78(3):415 – 422, 2003. ISSN 01651765.
- Tobias Rydén, Timo Teräsvirta, and Stefan Åsbrink. Stylized facts of daily return series and the hidden Markov model. *Journal of Applied Econometrics*, (3):217, 1998. ISSN 08837252.
- Kevin Sheppard. Mfe toolbox, version 1.08. Online, 2014. URL <http://www.kevinshppard.com>. Downloaded on November 13, 2014.
- Y. Shi. Can we distinguish regime switching from long memory? A simulation evidence. *Applied Economics Letters*, 22(4):318–323, 2015. ISSN 14664291.
- Yanlin Shi and Kin-Yip Ho. Long memory and regime switching: A simulation study on the Markov regime-switching arfima model. *Forthcoming*, xx(xx):xx, 2014a.
- Yanlin Shi and Kin-Yip Ho. Long memory and regime switching in the second moment: A simulation study. Online Publication, 2014b. URL <http://ssrn.com/abstract=2469028>.



Albert Nikolaevich Shiryaev. *Probability (2nd Ed.)*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 1995. ISBN 0-387-94549-0.

Maryam Tayefi and T.V. Ramanathan. Maximum likelihood estimation of stationary univariate fractionally integrated time series models. *Austrian Journal of Statistics*, 41 (3):175 – 196, 2012.