Markov Chain Switching model applied on entropy index based on Systemic Risk Measures

MSc Econometrics Economics & Finance
Ca' Foscari University of Venice

Master thesis 2015

Student name: Imane EL idrissi
Student number : 855749
Thesis supervisor : Prof. Roberto Casarin

Acknowledgements

The master's thesis in front of you is the result of months of labor and quite a bit of hurdles. Difficulties which I managed to overcome with the help of my supervisor Prof. Roberto Casarin. I would like to thank him ever so much for his patience and good advice. I found our meetings truly inspiring and he always gave me an energy boost to continue working and further improving my thesis.

I would also like to thank Prof. Monica Billio and prof. Eric Girardin for their encouraging words during my masters at Aix Marseille and Ca'Focsari Universities.

Lastly, I am very grateful to have attended the MEEF program. The courses that I took challenged me and proved enormously helpful towards my professional career.

In addition, I am utterly grateful for the unquestionable love and moral support given to me by my family and some friends for their endless support and faith in me. Especially in times when I had to relieve my feelings, they were always there for me to listen and cheer me up.

Imane ELidrissi
June 2015

Abstract

The aim of this thesis is to identify entropy regimes for the European financial market. I applied Markov switching model to the entropy series estimated on three different systemic risk measures, $\Delta CoVaR$, Marginal Expected Shortfall (MES) and finally network connectedness In-Out measure. The approach used to estimate those models is Expectation Maximization. I considered two and three states. In order to select the model I used (AIC) and (BIC) criteria

Finally since the entropy measures exhibits a unit roots I investigate the role of the unit root in the detection of the entropy regimes.

Keywords: systemic risk; Systemic risk measure; $\Delta CoVaR$, MES, Dynamic Causality, network connectedness; Shannon entropy (1984); Rényi entropy (1960); Tsallis entropy (1988); Markov switching model; transition probability; AIC ; BIC; EM

Contents

Introduction

The financial crisis of 2007-2008 has pushed concerns about systemic risk and its measurement at the forefront of both academic research and supervisory policy agenda. In particular, ongoing work by the Basel Committee and the Financial Stability Board striving to set new regulatory requirements for Systemically Important Financial Institutions (SIFI) requires that an agreement can be reached on which characteristics make a financial institution more prone than others to be severely hit by system-wide shocks (systemic resilience or participation) or to propagate such shocks to other institutions, thereby amplifying their overall impact (systemic contribution)[1].

New mathematical and computational tools have allowed researchers to work on undiscovered areas and to create new measures and approaches to systemic risk measurement. In particular, systemic risk analysis has improved a lot using extended versions of VaR based measures as $\Delta CoVaR$, defined as the difference between the VaR of the financial system conditional on the institution being under distress and the VaR of the financial system conditional on state of that institution. Another relevant measure is the marginal expected shortfall (MES) proposed by Acharya et al. (2010), it considers the average return of each firm during the 5% worst days for the market. Additionally to these measures Diebold and Yilmaz (2009) introduced a new approach of connectedness which called In-Out measure, it is an international spillover index based on assessing shares of forecast error variation in different locations (firms, markets, countries,etc) due to shocks arising elsewhere.

In the same way, few years ago new econometrics measures appeared considering the connectedness among financial institutions, one of the new measures is called Dynamic Causality index introduced by Gemantsky, Lo and Pelizzon and also (Billio et al 2012) which is capturing the degree of interconnectedness by looking at Principal Components and Granger Causality relations between returns and the international spillover index. Thus it is interesting to compare this new approach to understand the systemic risk with some previous categories of measures in order to identify the best risk measure that could capture possible missed elements of the contagion effect observed during financial crises and also to characterize network structure. These measures consider the financial system as a “portfolio” of institutions, where the shocks of the market price could impact the others. These measures are constructed primarily to capture the contribution to systemic risk of individual institutions. In particular, they capture the relation between the distress of each individual firm and the distress of the whole system.

Finally, my study relies on the use of entropy applied to a feature distribution estimated on the market, such as the cross-sectional systemic risk measures introduced above $\Delta CoVaR,$ MES and network connectedness In-Out measure. Indeed, the change of entropy built on these measures could reveal signs of changes on systemic risk.

[1] Analytically, one may want to distinguish between situations where bank A reacts more than others to an exogenous shock and situations where Bank A is a source or an amplifyer of endogenous systemic events. Both dimensions of systemic importance are in practice clearly inter-related. The participation vs contribution approach was proposed by Drehman and Tarashev (2010).

Econometric tool model applied in this thesis is the Markov switching model, which is an important reference point since the work of Hamilton (1989, Econometrica). These models have been increasingly used in financial time-series econometrics thanks to their ability to capture some key features, such as heavy tails, volatility clustering, and mean reversion in asset returns [see Cecchetti and al. (1990), Pagan and Schwert (1990), Turner and al. (1989), Gray (1996), Hamilton (1988), and Timmermann (2000), among others]. In contrast to linear models those assume stationary distributions (such as ARIMA models), regime-switching models are based on a mixture of parametric distributions whose mixture probabilities depend on unobserved state variable(s). A key difference between the various regime-switching models lies in the stochastic structure of the state variables *S*. For instance, the state of the unobserved process can be modeled by a discrete time/discrete space Markov chain, which can has either fixed or time-varying transition probabilities, or by an independent stochastic state variable.

In this analysis I focus on univariate mean/intercept and variance switching models only considering the cases of two and three states with no time-varying transition probabilities. I compare three entropy indexes Shannon (1984), Rényi, (1960) and Tsallis (1988) applied on the three cross section distribution of risk measures mentioned before, ΔCOVAR, MES and In-Out.

These popular models are applied for the entropy indexes associated with each risk measure obtained from the available total number of financial institutions included in the European market for the period January 02nd 1986 to May 12th 2014, the data corresponding to the three entropy indexes. In order to compare the goodness of fit of the models we compute the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC). In the models estimations I use the MS_Regress Matlab Tool-Box developed by Marcelo Perlin (Perlin, 2014) for estimating Markov Chain Switching Models (MCSM) through the maximum likelihood method and Hamilton´s filter.

This thesis is organized as follows, after this introduction, in section 1 and 2, I will offer a literature review covering past studies into systemic risk and systemic risk measures, which offer a suitable backdrop to the results of this paper. In section 3, I will discuss the entropy measures applied on the risk measures introduced before focusing on their differences and characterizations. I follow up by revising the approach of Markov Chain Switching Models and its importance in section 4. Before introducing and discussing empirical results in section 6 and 7 for data with unit root and without unit root, a detailed description of the estimation used in the empirical results is given on section 5. The intention of this is to give the reader a factsheet with which to study the empirical results of section 6 and 7. Finally, in section 8, I provide a conclusion of the outcomes of this study. As the study involves a substantial amount of empirical work, the main body of this thesis covers and discusses the most important elements thereof. The empirical appendix on section 9 offers detailed explanations about the MATLAB code implemented, general summary and statistics.

1. Systemic Risk

1.1 Definition and literature review.

There is no accord regarding the concept of financial stability and systemic risk. The materialization of systemic risk during the recent global financial crisis proved that the financial safety net and financial institutions significantly underestimated it. Systemic risk turned out to be much more than just the composition of individual types of risks affecting financial institutions. Whereas liquidity risk, credit risk, operational risk, etc. can be directly attributed to a given institution, systemic risk can only be attributed indirectly. Before the global financial crisis those types of risk were frequently considered separately. However, the interaction (correlation) between them leads to undesired and unexpected costs and consequences and when aggregated to systemic risk. Systemic risk evolves along with the development and growth of financial markets, regulations and collective behavior of market participants and it may be prompted by regulatory arbitrage. It is useful to explore the Systemic Risk definition. As Patrick Liedtke discusses, "*what is truly remarkable is that at this point in time no definition of systemic risk exists that would be both fully convincing and generally accepted*"

1. Clearly it represents an important aspect to deal with in order to identify an appropriate measure. The European Central Bank defines systemic risk as "the risk that the inability of one institution to meet its obligations when due will cause other institutions to be unable to meet their obligations when due. Such a failure may cause significant liquidity or credit problems and, as a result, could threaten the stability of or confidence in mar-kets"

2. In this sense, the Contagion Effect that characterizes financial crises could be associated with one of the mechanisms by which systemic risk is observed. Of course, a good measure of systemic risk has a relevant role for policy makers, as Acharya comments "the need for economic foundations for a systemic risk measure is more than an academic concern as regulators around the world consider how to reduce the risks and costs of systemic crises"

3. As a result, an appropriate systemic risk measure should include both theoretical and practical relevance.

There are at least three approaches that can be used to assess the build-up of systemic risk. The first focuses on monitoring traditional indicators of financial soundness or stability in order to assess broad developments in the financial system; the second focuses on measuring interconnectedness between financial institutions; and the third focuses on changes in the behavior of prices of financial assets.

2. Systemic Risk Measures

In this section I introduce three cross sectional distribution of marginal systemic risk measures first Marginal expected Shortfall (MES), Delta Conditional Value-at-Risk ($\Delta CoVaR$), and network connectedness called In-Out measure.

2.1 Systemic risk measure: Marginal Expected Shortfall (MES)

MES is a measure of the sensitivity of a financial firm to systemic risk .Systemic risk is defined as "the risk that the intermediation capacity of the entire financial system is impaired" (Adrian & Brunnermeier p.1). The first way to look at this risk is to examine the extent to which an individual bank or institution is affected by a systemic crisis. To this end, Marginal Expected Shortfall (MES) as introduced by Acharya et al. (2010) is considered. It is the average return of an individual institution during the 5% worst days of the market:

$$MES_{5\%}^{i} = \mathrm{E}\left[\frac{w_1^i}{w_0^i} - 1 | I_{5\%}\right]$$

In which

$MES_{5\%}^{i}$ = Marginal Expected Shortfall during the 5% worst trading days on the market

$\frac{w_1^i}{w_0^i} - 1$ = Return of institution *i*

$I_{5\%}$ = The 5% worst outcomes at the market

MES measures the loss of an individual Institution when the entire market is doing poorly. A more negative value for MES indicates more systemic risk. From now on, a higher value of MES is interpreted as a more negative value for MES. Thus a higher value for MES implies more systemic risk.

2.2 Systemic risk measure: $\Delta CoVaR$:

The second way to view systemic risk is to measure the contribution of an individual institution *i* to the overall systemic risk of the system. For this purpose, $\Delta CoVaR$ is used as discussed in Adrian and Brunnermeier (2011). *CoVaR* is the conditional Value at Risk, in which "Co" stands for conditional, contagion, or co-movement. *CoVaR* of the institution *i* relative to the system is the VAR of the whole financial sector conditional on institutions under distress. Furthermore, $\Delta CoVaR$ is the difference between the *CoVaR* of the financial system conditional on institution *i* being in distress and ΔCoVaR of the financial system conditional on institution *i* operating in its median "normal" state, (Brunnermeier, Dong, and

Palia, 2012). As deduction ΔCoVaR, captures the marginal contribution of a particular institution i (in a non-causal sense) to the overall systemic risk.

$$\Delta CoVaR_{q,t}^{i} = \Delta CoVaR_{q,t}^{system\,|\,i} - \Delta CoVaR_{q,t}^{system\,|i\,,median}$$

Denote $\Delta CoVaR_{\mathrm{q,t}}^{\mathrm{i}}$ as the contribution of Institution i to the systemic risk of the entire system at time t. Denote $\Delta CoVaR_{q,t}^{system\,|\,i}$ as the Value at Risk of the entire system conditional on bank *i* being in distress at time *t*. It is the q % Value at Risk of the entire system conditional on institution *i* operating at its VaR level. Denote $\Delta CoVaR_{q,t}^{system\,|i\,,median}$ as the Value at Risk of the entire system conditional on bank *i* operating at its median state at time *t*. As one can conclude from subscript *t*, the terms are time varying, which implies that the model is a dynamic conditional model instead of a stable unconditional model.

There are several advantages to the $\Delta CoVaR$ measure. First, while $\Delta CoVaR$ emphasizes on the contribution of each institution to overall system risk, traditional risk measures focus on the risk of individual institutions. Further, it is general enough to study the risk spillovers from institution to other across the whole financial network. Δ for example captures the increase in risk between two institutions, when the first falls into distress. To the extent that it is causal, it captures the risk spillover effects that institution i causes on institution j. Of course, it can be that institution i distress causes a large risk increase in institution j, while institution j causes almost no risk spillovers over institution i. The most common method to test for volatility spillover is to estimate a multivariate GARCH processes.

2.3 Systemic risk measure: network conectedness "In − Out"

Systemic risk highlights on contagion, spillover effects and co-movement in financial markets, some measures of systemic risk are based on principal components analysis (PCA) or Granger-causality test. For instance absorption ratio based on PCA was used for systemic risk measurement as proposed by Kritzman et al. In the same vein, Diebold and Yilmaz (2014) propose several connectedness measures built from variance decompositions, and they argue that these piece of variance decomposition provides a natural and insightful measure of connectedness among financial asset returns and volatilities. They also show that variance decomposition defines weighted, directed networks, so that their connectedness measures are intimately-related to key measures of connectedness used in the network literature. ‘Building on these insights, we track both average and daily time-varying connectedness of major U.S. financial institutions' stock return volatilities in recent years, including during the financial crisis of 2007-2008. On the other hand, Billio et al. (2012) suggested several measures of systemic risk focused on the connectedness of the financial institutions to quantify the interrelationship between the monthly returns of hedge funds, banks, brokers and insurance companies based on principal Component analysis (PCA) and Granger-causality tests. They introduced an indicator Number of Connections (*NC*) calculated from Granger-causality test to measure the degree of systemic risk. When applying PCA, the basic idea is that systemic risk is getting higher when the largest eigenvalue increases explaining most of variation of the data. When applying Granger-causality test. The first step is to estimate a Generalized Autoregressive Conditional Heteroskedasticity model for the returns using a structure GARCH (1, 1) and conditional on the system information that is developed over the sigma-algebra б:

$$I_{t-1}^{s} = б(\{\{R_{\tau}^{i}\}_{\tau=-\infty}^{t-1}\}_{i=1}^{N}$$

So this structure allows identifying the relationships among institutions that are included in the network. Applying the Granger-causality test they work on the directionality of the relationships between the elements in the system. Now, they define the following indicator of causality:

$$(j \rightarrow \mathrm{i}) = \begin{cases} 1, & if\ j\ Granger\ causes\ i, \\ 0, & otherwise \end{cases}$$

And also when $(j \rightarrow \mathrm{i}) = 0$. This indicator is used to define the connections between N financial institutions. The degree of Granger causality (DGC) among all N (N-1) pairs of N financial institutions is given by the formula:

$$DGC \equiv \frac{1}{N(N-1)} \sum_{\mathrm{i}=1}^{\mathrm{N}} \sum_{\mathrm{j} \neq \mathrm{i}} (j \rightarrow \mathrm{i})$$

As a result, when DGC exceeds a threshold *K* the systemic risk increases and it is possible to count the *Number of Connections* that affect a particular institution (*IN)* and are affected by the same institution (*OUT).* So, when systemic risk is raising the level of system interconnectedness also increases. Therefore, the total number of connections is given by:

$$\#In + Out: (j \leftrightarrow S)\big|_{DGC \geq K} = \frac{1}{2(N-1)} \sum\nolimits_{\mathrm{i} \neq \mathrm{j}} (i \rightarrow j) + (j \rightarrow i)\big|_{DGC \geq K}$$

3. Entropy index

3.1 Introduction

"Entropy" this word appeared in 1865 when the German physicist Rudolf tried to rename the irreversible heat loss, what he previously called "equivalent-value". This name was chosen for the reason that in Greek, en+tropein" means "content transformative" or "transformation content" .since this years the entropy has an important application in thermodynamics. Being defined as the sum of "heat supplied" divided by "temperature", it is central to the Second Law of Thermodynamics. Using the Conrad definition "*In our probabilistic context entropy is viewed as a measure of the information carried by the probability distribution, with higher entropy corresponding to less information (more uncertainty, or more of lack of information)"*

Entropy can be used and measured in many other fields than thermodynamics. One of the important applications of entropy in information theory is Shannon entropy. In finance as well the application of entropy can be considered as the extension of the information entropy. It is an important tool in asset pricing and portfolio selection.

All the above papers recognize that entropy could be a good measure of risk; however it seems to be difficult to use this measure. In the systemic risk context considering the financial system behavior as random process where the values of the risk measure associated with each institution included in are realizations of this process, it is possible to estimate an *entropy* function for some specific period and then analyze the performance of this entropy index over the time. Thereby, comparison between entropy indexes associated with different risk measures could be a good method to understand systemic risk dynamics. My main motivation is to compare tree different entropies indexes, each one based on tree different systemic risk measures, $\Delta CoVaR$, MES, In-Out in order to select the best model to identify and forecasts financial crises.

3.2 Concepts of Entropy Used in Finance

3.2.1 The Shannon Entropy (1984)

The Shannon entropy of a probability measure *p* on a finite set *X* is given by:

$$E_S(p) = -\sum_{i\in x} p_i \log p_i \qquad (1)$$

Where $\sum_{i=1}^{n} p_i = 1$, $p_i \geq 0$ and $0 \log 0 = 0$. And the base of the logarithm is 2.

When dealing with continuous probability distributions, a density function is evaluated at all values of the argument. Given a continuous probability distribution with a density function *f*(*x*), we can define its entropy as:

$$E = -\int_{-\infty}^{+\infty} f(x)\, log(x)\, dx \qquad (2)$$

Where $\int_{-\infty}^{+\infty} f(x)dx=1$ and $f(x) \geq 0$.

3.2.2 The Rényi (1960) and Tsallis (1988) Entropy

For any positive real number α. $0 < \alpha < \infty$ The Rényi and Tsallis entropy of order α, of a probability measure *p* on a finite set *X* is defined to be.

$$E_R(p) = \frac{1}{1-\alpha} log\left(\sum_{i\in x} p_i^{\alpha}\right) \qquad (3)$$

$$E_T(p) = \frac{1}{\alpha-1}(1 - \sum_{i\in x} p_i^{\alpha}) \qquad (4)$$

We need $\alpha \neq 1$ to avoid dividing by zero, but l'Hôpital's rule shows that the Rényi entropy approaches the Shannon entropy as α approaches *1*

$$lim_{\alpha\to 1} S_R(p) = -\sum_{i\in x} p_i\, log\, p_i$$

Maszczyk and Duch (2008) showed that, the Shannon entropy is a special case of the Rényi and Tsallis entropies. Specifically, according to the value of α, the measures in Equations (3) and (4) attribute roughly weight to the tails of the distribution. At a first glance, the main difference between Shannon and Rényi entropies is the placement of the logarithm in the expression. In Shannon entropy (1), the probability mass function weights the log (p_i) term, whereas in Rényi entropy the log is applied for the total summation that involves the α power of the probability mass function. A large positive α value implies this measure is more sensitive to events that occur often, while for a large negative α value shows more sensitiveness to the events which happen seldom.

In order to compare further with Shannon entropy let me rewrite Rényi entropy as

$$E_R(p) = \frac{1}{1-\alpha} log\left(\sum_{i \in x} p_i^{\alpha}\right) = - log \left(\sum_{i \in x} p_i^{\alpha}\right)^{\frac{1}{\alpha-1}}$$

$$= - log \left(\sum_{i \in x} p_i p_i^{\alpha-1}\right)^{\frac{1}{\alpha-1}} \quad (5)$$

In (5) the probability mass function p_i also weights a term that now is the $\alpha - 1$ power of the probability mass function. At a deeper level, Rényi entropy measure is much more flexible due to the parameter α enabling several measurements of uncertainty (or dissimilarity) within a given distribution, the higher the α and the less the distribution entropy far from the uniform. In other words the farther a distribution is from the uniform, the thinner its tails are.

While, the Tsallis entropy assigns less importance to randomness that is it penalizes uniformity in the distribution. This entropy behaves in the same way of Shannon; there is only difference in the magnitude. When α increases or decreases the magnitude also increases or decreases.

4. Markov Chain Switching Regime

4.1 Glossary

- Regime-Switching Model

  Is a parametric model of a time series in which parameters are allowed to take on different values in each of given fixed number of regimes

- Markov Chain (MC)

  Is a process that consists of a finite number of regimes, where the probability of moving to a future regime conditional on the present regime is independent of past regime

- Markov-Switching Model (MSM)

  Is a regime switching model in which the shifts between the states evolve depending on an unobserved MC.

- Regime-Switching Model

  Is a parametric model in which parameters are allowed to take on different values in each of given fixed number of regimes (states).

- Transition Probability

  The probability of switching from state *j to* state *i*

- Filtered Probability of a Regime

  The probability that the unobserved Markov Chain (MC) for a Markov-switching model is in a particular regime in period t, conditional on observing sample information up to period T.

- Smoothed Probability of a Regime

  The probability that the unobserved MC for a Markov-switching model is in a particular regime in period T, conditional on observing all sample information.

4.2 Definition and Importance of Regime switching models

Regime-switching models have been discussed more than 50 years ago. Early econometrics focused on a simple model that incorporates a single non-recurring structural break. These time-series models in which parameters are allowed to take different values in each fixed number of "regimes." A stochastic process supposed to have generated the regime shifts is included as part of the model, that allows for model-based forecasts which incorporate the possibility of future regime shifts. In some special situations the regime in operation at any point in time is directly observable. Generally the researchers must conduct inference about which regime the process was in at past points in time.

Describing changes in the dynamic behavior of macroeconomic and financial time series has been the main purpose of using these models in the applied econometrics literature. Regime-switching models could be divided into two categories, "threshold" models and "Markov-switching" models, which are used in my thesis. The principal difference between these approaches is in how the advancement of the state process is modeled. The first models, introduced by Tong (1983), suppose that regime shifts are triggered by the level of observed variables in relation to an unobserved threshold. The second models, introduced to econometrics by Goldfeld and Quandt (1973), Hamilton (1989), suppose that the regime shifts evolve depending on a Markov chain. Regime-switching models became increasingly popular tool of modeling for applied work, these models provided an alternative and important approach to understand and interpret how an economy's growth rate changes over time. An important development in this regard is the Hamilton (1989) to model and identify the phases of the US business cycle. There is other application which includes modeling regime shifts in time series of inflation and interest rates (Garcia and Perron, 1996).High and low volatility regimes in equity returns (Hamilton and Susmel, 1994; Hamilton and Lin, 1996) and time variation in the response of economic output to monetary policy actions (Lo and Piger, 2005; Kaufmann, 2002).

<u>4.3 Markov chain Switching</u>

There is huge interest in modeling dynamic behavior of macroeconomic and financial time series. One of the challenges for this analysis is that these time series usually undergo changes in their behavior over reasonably long sample periods. The change usually occurs in the form of a "structural break", in which there is a shift in the behavior of the time series due to some permanent change in the economy's structure. The change in behavior might also be temporary, as in the case of wars or "pathological" macroeconomic episodes such as economic depressions, hyperinflations, or financial crises. These shifts might be both temporary and recurrent, in that the behavior of the time series might cycle between regimes. For example, as indicated by early studies of the business cycle, the behavior of economic variables changed dramatically in expansions vs. recessions.

The constant parameter time series models might be inadequate to describe the potential for shifts in the behavior of economic time series and their evolution. As a result, recent decades have seen extensive interest in econometric models designed to incorporate parameter variation. One approach to describing this variation, denoted a "regime-switching" model in the following, is to allow the parameters of the model to take on different values in each of some fixed number of regimes, where, in general, the regime in operation at any point in time is unobserved by the econometrician. The process that determines the arrival of new regimes is assumed known, and is incorporated into the stochastic structure of the model which allows the econometrician to draw inference about the regime that is in operation at any point in time, as well as form forecasts of which regimes are most likely in the future.

Quandt (1958) first tried to estimate the parameters of this linear regression system; he supposed that such a system follows two separate regimes. Similar to the model in (1) and (2)

$$y_1 = \mu_1 + \varepsilon_t \qquad \varepsilon_t \sim N\ (0,\sigma_1^2) \qquad (1)$$

$$y_2 = \mu_2 + \varepsilon_t \qquad \varepsilon_t \sim N\ (0,\sigma_2^2) \qquad (2)$$

Denote y_t as return series, and μ_i and σ_i^2 define the mean and variance in period *i (*further $\mu_1 > \mu_2$ or $\sigma_1 < \sigma_2$ is assumed). He assumed a point in time at which the system switches the regime and the data is described by a different regression equation. Using maximum likelihood estimation, Quandt was capable to infer the corresponding turning point and to determine the regression parameters. Despite, his model considers only the possibility of a single switch and it was only tested in a hypothetical sampling experiment.

To allow studies of multiple regime-switches Quandt (1972) prolonged his earlier model and applied it to the US housing market between June 1959 and November 1969, by introducing the λ-method, where λ and1 – λ are unknown probabilities for the observed data points being driven by the regime 1 or 2, as result he got a significant results for the two regime.

Goldfeld and Quandt (1973) introduced the first Markov-switching model after relaxing the assumption of constant regime probabilities. They assumed that there is dependence of the current regime on its preceding state through a Markov chain process. To capture the state dependency they introduced a transition probability matrix, which governs the transitions across states. When they applied τ -method to the same data sample as Quandt (1972) they obtained also similar results as for the λ-method. All these applied works shows that these models became more sophisticated and complex because of the data availability and computational improvements..

5. Estimation

A general Markov Switching model can be estimated with two different ways first by Maximum likelihood or by Bayesian inference (Gibbs-Sampling). This thesis focuses and explains the first method.

This estimation precedes a recursive filter and numerical maximization by a method by an EM algorithm, which is appropriate for maximizing likelihood with unobserved variables or missing observations. I describe these procedures in this section.

5.1 Specification

I define the model structure that I will apply in this estimation. I am working on intercept and variance switching models, for two-states the model is given by:

$$E_t = \mu_{S_t} + \varepsilon_t$$

$$\varepsilon_t \sim \begin{cases} N(0, \sigma_1^2) & if\ S_t = 1 \\ N(0, \sigma_2^2) & if\ S_t = 2 \end{cases}$$

Where E_t represents one entropy index for each of the risk measure in the period t and μ_{S_t} and $\sigma_{S_t}^2$ represent its mean and variance. I assume that ε_t and μ_{S_t} follow a distribution that depends on a latent process (variable) S_t, this variable follows a Markov Chain process. At each point in time, the process S_t is in one of two regimes, which we Indicate by $S_t = 1$ and $S_t = 2$.

In both regimes, the entropy (E_t) follows a normal distribution, though with different means and different variances. I use the function *f* to denote the normal probability density function.

$$f(E_t) = \frac{1}{\sqrt{2\pi\sigma_{St}^2}}\ exp\left(-\ \frac{(E_t - \mu_{St})^2}{2\sigma_{St}^2}\right)$$

And the log likelihood of this model is given by:

$$ln\,L = \sum_{t=1}^{T} ln\left(\frac{1}{\sqrt{2\pi\sigma_{St}^2}}\ exp\left(-\ \frac{E_t - \mu_{St}}{2\sigma_{St}^2}\right)\right) \qquad (1)$$

If all of the states of the world were know, that is, the values of S_t are available, then estimating the model by maximum likelihood is straightforward. All I need is to maximize equation (1) as a function of parameters μ_1 ,μ_2 ,σ_1^2 and σ_2^2 . But is not the case of a Markov Switching model, because the states of the world are unknown. So in this case it is necessary to change the notation of the likelihood function. By considering $f(E_t \mid S_t = j, \Theta)$ as the likelihood function for state j conditional on a set of parameters (Θ), then the full log likelihood function of the model is given in the equation (2)

$$lnL = \sum_{t=1}^{T} ln\ \left(f(E_t \mid S_t, \Theta)\ \Pr(S_t = S_{t-1})\right) \qquad (2)$$

The latent process S_t follows a first order Markov Chain, this means that the probability for regime 1 to occur at time *t* only depends on the regime at time *t-1* . I denote these Transition probabilities by

$$p_{ij} = Pr\ [S_t = i \mid S_{t-1} = j]$$

The p_{ij} term represents the probability of switching from state *j* to state *i.* The transition probabilities for the departure states *j* should add up to one, i.e., $p_{11} + p_{21} = 1$ and $p_{12} + p_{22} = 1$. So, for a binary process S_t, I have two free parameters. p_{11} and p_{22}.The transition matrix is defined by:

$$P = \begin{pmatrix} p_{11} & p_{21} \\ p_{12} & p_{22} \end{pmatrix} = \begin{pmatrix} p_{11} & 1 - p_{11} \\ 1 - p_{22} & p_{22} \end{pmatrix}$$

To estimate the parameters of the regime switching models using a maximum likelihood approach. As with other conditional models such as ARMA- or GARCH-models, the likelihood function will take a conditional form, too. I gather the parameters of the model in a vector $\Theta = (\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, p_{11}, p_{22})$

5.2 The Hamilton filter

In order to maximize the Log-Likelihood function is necessary to estimate Pr $(S_t = j)$ using the Hamilton's filter that allows estimating these probabilities using the available information in the period *t-1* denoted byΨ_{t-1}. In this sense, the Log-Likelihood function will be estimated as a function of the parameters and the available information, then we have:

$$lnL = \sum_{t=1}^{T} ln \sum_{j=1}^{2} (f(E_t \mid S_t = j, \Theta) \Pr(S_t = j \mid \Psi_t)$$

Four steps are needed to realize the algorithm:

1. Establish a guess of stating probabilities in *t=0* of each state, $\Pr(S_0 = j)$ for *j=1,2* which satisfy the relation

$$\Pr(S_0 = 1 \mid \Psi_0) = \frac{1 - p_{1,1}}{2 - p_{1,1} - p_{2,2}}$$

$$\Pr(S_0 = 2 \mid \Psi_0) = \frac{1 - p_{2,2}}{2 - p_{1,1} - p_{2,2}}$$

2. Prediction of regime probabilities i.e. In *t=1* compute the probabilities for each state given Ψ_0 and the transition matrix, where

$$\Pr(S_t = j \mid \Psi_{t-1}) = \sum_{i=1}^{2} p_{j,i} (\Pr(S_{t-1} = i \mid \Psi_{t-1}))$$

3. Calculation of filtered probabilities through update i.e Update the probability of each state with the new information from time *t*. This can be done using the parameters of the model in each state, $\mu_1, \mu_2, \sigma_1^2, \sigma_2^2$ and the transition probabilities $p_{1,1}$and $p_{2,2}$ for computing the likelihood function in each state over the period *t*. Next step, update the probability of each state given the new information as follows:

$$\Pr(S_t = j \mid \Psi_t) = \frac{f(E_t \mid S_t = j, \Psi_{t-1}) \Pr(S_t = j \mid \Psi_{t-1})}{\sum_{j=1}^{2}(f(E_t \mid S_t = j, \Psi_{t-1}) \Pr(S_t = j \mid \Psi_{t-1})}$$

4. Now move to *t+1* and repeat the steps 2 and 3 until *t=T*, completing all the available data in the *T* periods. As a result, we obtain all the state probabilities across the sample.

5.3 Maximization

After computing the states probabilities by applying Hamilton's filter, it is possible to maximize the log likelihood function using a numerical method. In this case the numerical method used in MS_Regress_fit.m consists on applying four algorithms included in fmincon function in MATLAB.

1. SQP Algorithm

2. Interio-point Algorithm

3. Active-set Algorithm

4. Trust-Region-Reflective Algorithm

6. Results

6.1 data

I consider the daily closing price for the stocks of the European financial institutions from 2^{nd} January 1986 to 12^{th} may 2014. The data has 7398 observations. According to the dynamics of system the number of institutions is changing across the time, beginning with 30 institutions and ending with 203.

To estimate systemic risk measures I used a rolling window approach (e.g., see Zivot and Wang J.(2003), Billio et al.(2012), Diebold and Yilmaz(2014)) with a window size of 252 daily observations, which corresponds approximately to a year of daily observations.

6.2 Descriptive statistics

In this section I study the mean and variance behavior of each entropy index related with each risk measure $\Delta \boldsymbol{CoVaR}$, MES and INOUT, in order to identify different regimes of entropy and multimodal entropy distribution.

1. **ΔCoVaR**

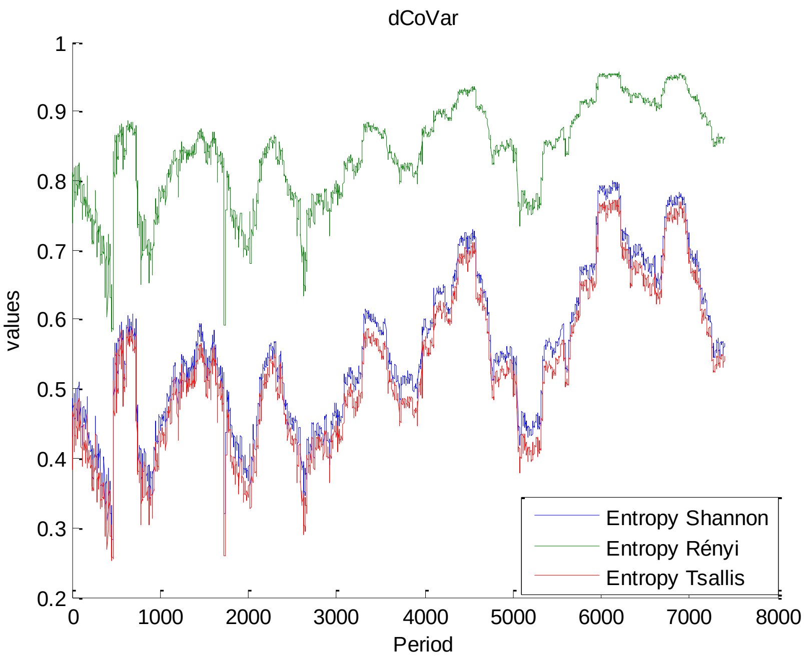


Figure 1: The entropy indexes for ΔCoVaR measure

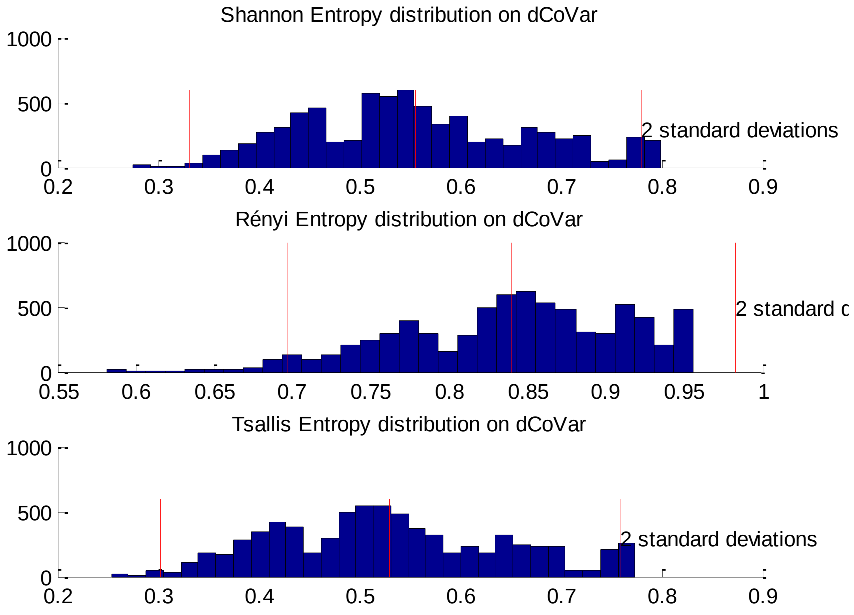


Figure 2: The entropies distributions for $\Delta CoVaR$ measure

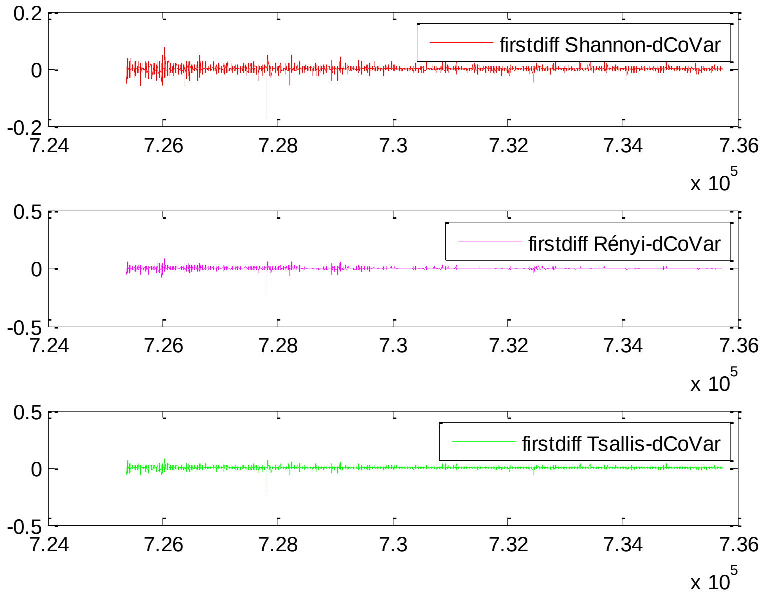


Figure 3: First difference entropies indexes for ΔCoVaR measure

For the cross-sectional systemic risk measure ΔCoVaR, the Figure 1 shows the plots of the three entropies Shannon, Tsallis and Rényi. The two first entropies behave in the same way, the difference remains in the magnitude. The Rényi behaves differently. Using this entropy could change totally the estimation of MSCM, specifically the presence of the states. I will have more information by analyzing the histograms and the first difference included in the Figure 2 and 3.

In order to get some conclusion about the state of the mean and of the variance, as the histograms show the mean is not representative because the entire histograms exhibit clearly multimode, so taking the mean like a normal model is not working properly for this data. Therefore I need to use models which account for multimodality such as mixture normal or Markov Switching, are more suitable, for these data.

Even more, the first difference series shows the presence of at least two states for the variance behavior. As a result, taking in account these facts I consider that a three states switching Markov chain could be the best model to approximate the entropies indexes. I will compare these expectations with the estimation results.

At the end my expectation for the entropy related to this systemic risk measure, is a Markov Chain Switching Model three states for the Rényi entropy, and two states for the Shannon and Tsallis.

2. Marginal Expected Shortfall (MES)

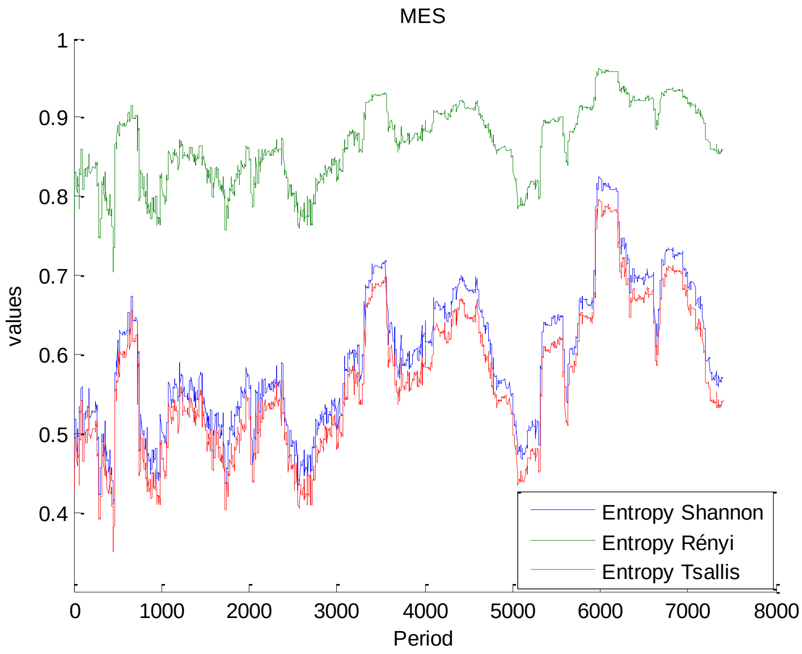


Figure 4: The entropy indexes for MES measure

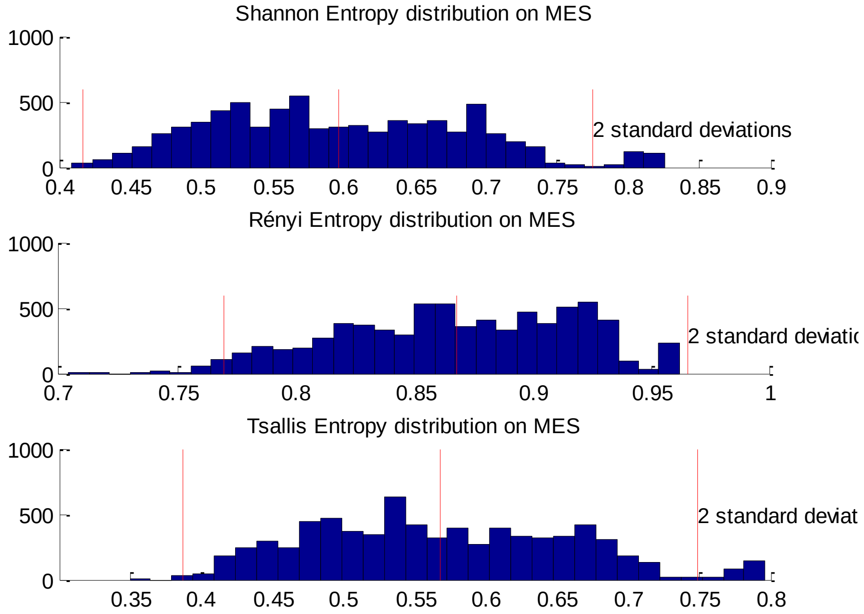


Figure 5: The entropies distributions for MES measure

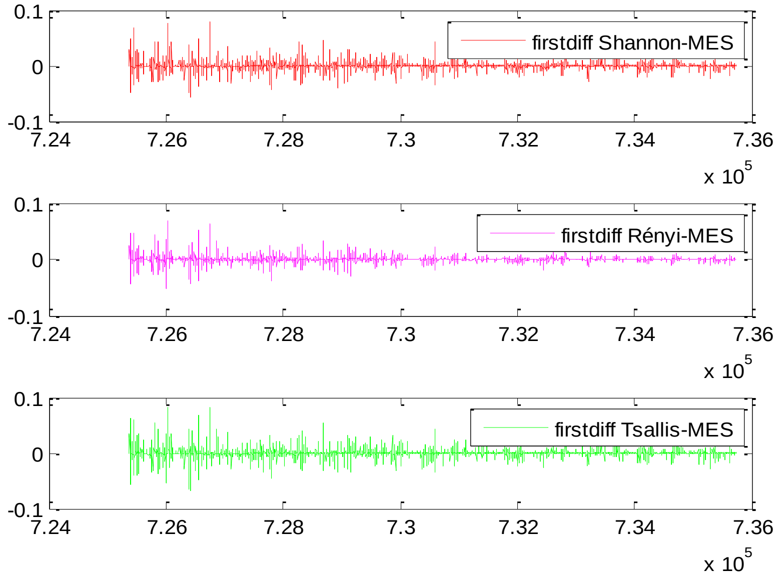


Figure 6: First difference entropies indexes for MES measure

From the Figure 4 I have the same conclusions for this measure, the Shannon and Tsallis entropies are similar in the behavior. Furthermore, after the period 3000 the mean increases and also the behavior becomes more volatile. Unlike the Rényi 's entropy, the histograms of Marginal expected shortfall reflect that it could be a Markov chain switching model with two state for the Shannon and three states for the Tsallis and Rényi In addition, the first difference shows that there are at least two states for the variance.

As a result my expectation for the entropy related to MES is a Markov Chain Switching model three states for all the entropies.

3. In-Out

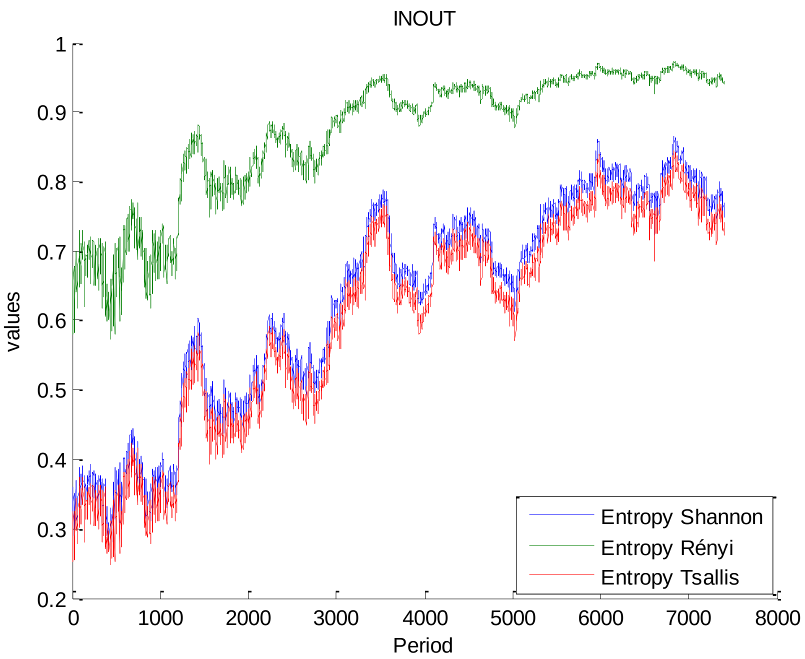


Figure 7: The entropy indexes for In-Out measure

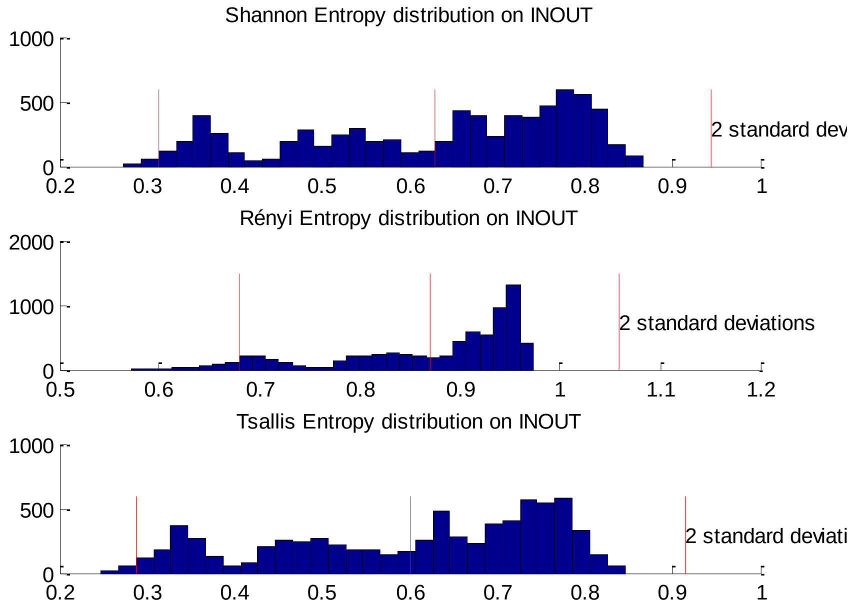


Figure 8: The entropies distributions for In-Out measure

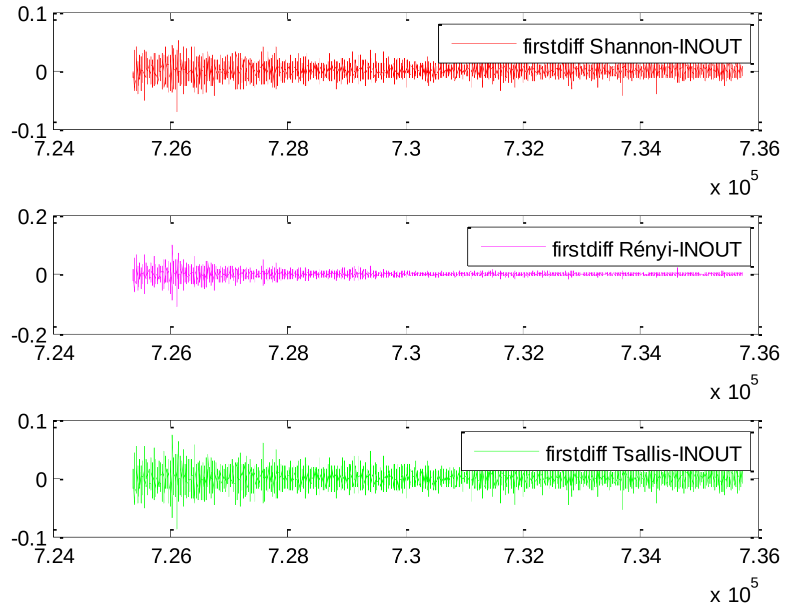


Figure 9: First difference entropies indexes for In-Out measure

The entropies of In-Out measure are increasing with time and tending to the Rényi entropy, as result I have bigger variance and a higher mean over the time. But for the histograms I remark a clearly a different behavior for the Rényi entropy with a bigger concentration around the mode .however it is possible to have three states for all of the entropies. In addition the first difference for the Shannon entropy is more even, in this case it will be difficult to identify the states but for the other entropies it has at least to states for MCSM.

Additionally when I plot the normalized entropy against the non-normalized, I remark that for the Shannon the normalization does not affect the pattern, I have only the change in the scales, because this entropy was normalized by using this formula (1) where k is the bins. However checking the magnitude the entropies before and after the normalization the Rényi entropy becomes the highest one, it means that the normalization has a bigger effect over Shannon and Tsallis than Rényi.

$$\text{Entropy Shannon\_n} = \frac{\boldsymbol{Entropy\ shannon\ non\ normalized}}{\boldsymbol{Log\ k}} \qquad (1)$$

The expected result for the entropy based on In-Out systemic risk measure is three states for all the entropies.

<u>6.3</u> <u>Estimation results</u>

I apply Markov chain switching model to entropy indexes, using the estimation method outlined in the previous section, for two states and three states. The data consists of three entropy indexes each one associated to one of these risk measures $\Delta\boldsymbol{CoVaR}$, MES and In-Out for the period mentioned before for two cases non-normalized and normalized.

The thesis includes the estimation of 36 models. In order to compare the fitness of the models I compute the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC).

$$AIC = -2\ ln\ (L) + 2\ ln\ (k)$$

$$BIC = -2\ ln\ (L) + k\ ln\ (n)$$

Where L is the likelihood evaluated at the mode of the parameters, *n* the number of the observations and k is the number of parameters in the model. BIC compares the (negative) likelihood, but penalizes for increased model complexity. For small sample sizes, BIC tends to choose less complex models, but the series in the empirical study have many observations, so this is not an issue.

After checking the estimation results I can identify some issues. Table 1 summarizes the log likelihood values evaluated at the parameter modes obtained for the different states. As expected, in most of the cases the log likelihood increases with the number of states in the

MCSM, with a big jump for the MES. According to the BIC values, a switching model with three states is the strongest candidate. I examine the results from these models further.

Tables 2-3 and 4 contain the values of (BIC) and (AIC) estimated for each entropy in two and three states MCSM. First of all I focus on the normalized case, for the Shannon and the Tsallis entropy, the best measure is $\Delta CoVaR$ for three states, unlike the Rényi entropy the best one is MES also in three states. Similarly the non-normalized indexes have the same results.

One the other hand when I compare the criteria for the normalized and non-normalized indexes it is clear that the models for the normalized indexes show a better performance than the non-normalized. In fact sometimes the difference between both is 29 times for the Shannon and Tsallis entropies. However in the Rényi case there are no significant differences, this could be associated to its own formula. As result I can say that the normalization is a good tool to improve the quality of the estimation in the Shannon and Tsallis entropies, but not in the case of the Rényi entropy.

When I analyze the models associated to In-Out measure, for all the entropies the best result is the three states and also the normalized case is better than the non-normalized. As result MCSM three states for normalized Rényi entropy index based on this measure has the lowest value for both criteria.

It is important to underline that when I compare the models related to Shannon entropy, the ratio for the (AIC) between the best model and the worst is 4.42 times, 1.43 times for the Rényi and 4.08 times for the Tsallis, as a conclusion is less sensible to the use of different risk measures, in other words this entropy has a lower model risk unlike the other entropies. Last important remark, this ratio is very high in the case of Shannon entropy for the non-normalized data (17.02) times, in the same why, by analyzing the (BIC) criteria I got similar results. In the case of the Shannon and Tsallis the normalization has a positive effect in reducing model risk but no a significant results for the Rényi.

		Shannon entropy Index		*Tsallis entropy Index*		*Rényi entropy Index*	
	logliklihood	*non-Norm*	*Norm*	*non-Norm*	*Norm*	*non-Norm*	*Norm*
2 S	$\Delta CoVaR$	-8306.2301	5701.4137	10723.1721	13060.9947	-5737.2494	5559.2982
	MES	-2040.9473	11966.2589	13994.4332	13651.2607	436.631	11733.1787
	In-Out	-10821.0986	3186.5452	11522.4526	13860.2752	-8083.1864	3213.3613
3 S	$\Delta CoVaR$	-623.6275	13384.0163	10723.0509	13061.0641	1784.8115	13081.3592
	MES	-1078.0593	12929.5845	16378.6258	18716.4484	436.6299	11733.1776
	In-Out	-4903.11	9104.5338	14295.3194	16633.142	-1095.6283	10200.9194

Table 1: The likelihood evaluated at the parameter modes obtained for different states

	$\Delta CoVaR$			MES			In-Out		
2 S	NORM	BIC	-1.1349e+04	NORM	BIC	-2.3879e+04	NORM	BIC	-6.3196e+03
		AIC	-1.1391e+04		AIC	-2.3921e+04		AIC	-6.3611e+03
	Non_NORM	BIC	1.6666e+04	Non_NORM	BIC	4.1353e+03	Non_NORM	BIC	2.1696e+04
		AIC	1.6624e+04		AIC	4.0939e+03		AIC	2.1654e+04
3 S	NORM	BIC	-2.6661e+04	NORM	BIC	-2.5752e+04	NORM	BIC	-1.8102e+04
		AIC	-2.6744e+04		AIC	-2.5835e+04		AIC	-1.8185e+04
	Non_NORM	BIC	1.3542e+03	Non_NORM	BIC	2.2630e+03	Non_NORM	BIC	9.9131e+03
		AIC	1.2713e+03		AIC	2.1801e+03		AIC	9.8302e+03

Table 2: (BIC) and (AIC) values for the Shannon entropy index in two and three states

	ΔCoVaR			MES			In-Out		
2 S	NORM	BIC	-2.6069e+04	NORM	BIC	-2.7249e+04	NORM	BIC	-2.7667e+04
		AIC	-2.6110e+04		AIC	-2.7291e+04		AIC	-2.7709e+04
	Non_NORM	BIC	-2.1393e+04	Non_NORM	BIC	-2.7935e+04	Non_NORM	BIC	-2.2991e+04
		AIC	-2.1434e+04		AIC	-2.7977e+04		AIC	-2.3033e+04
3 S	NORM	BIC	-2.6015e+04	NORM	BIC	-3.7326e+04	NORM	BIC	-3.3159e+04
		AIC	-2.6098e+04		AIC	-3.7409e+04		AIC	-3.3242e+04
	Non_NORM	BIC	-2.1339e+04	Non_NORM	BIC	-3.2650e+04	Non_NORM	BIC	-2.8484e+04
		AIC	-2.1422e+04		AIC	-3.2733e+04		AIC	-2.8567e+04

Table 3: (BIC) and (AIC) values for the Tsallis entropy index in two and three states

	ΔCoVaR			MES			In-Out		
2 S	NORM	BIC	-1.1065e+04	NORM	BIC	-2.3413e+04	NORM	BIC	-6.3733e+03
		AIC	-1.1107e+04		AIC	-2.3454e+04		AIC	-6.4147e+03
	Non_NORM	BIC	1.1528e+04	Non_NORM	BIC	-819.8082	Non_NORM	BIC	1.6220e+04
		AIC	1.1486e+04		AIC	-861.2620		AIC	1.6178e+04
3 S	NORM	BIC	-2.6056e+04	NORM	BIC	-2.3359e+04	NORM	BIC	-2.0295e+04
		AIC	-2.6139e+04		AIC	-2.3442e+04		AIC	-2.0378e+04
	Non_NORM	BIC	-3.4627e+03	Non_NORM	BIC	-766.3522	Non_NORM	BIC	2.2982e+03
		AIC	-3.5456e+03		AIC	-849.2598		AIC	2.2153e+03

Table 4: (BIC) and (AIC) values for the Rényi entropy index in two and three states

After the estimation of the models, comparing these results with my expectations, I got the same conclusions for all the entropies based on In-Out measure; it means that the best model is MCSM with three states. For the Rényi and Shannon entropy based on MES measure the efficient models are three states MCSM, which agrees with my expectation, additionally the model related to Tsallis index is a two state MCSM which is opposite to my beliefs. On the other side for all the entropies related to $\Delta CoVaR$ measure, the expectations are opposite to the results, notice that for the Rényi index the difference between the states is very small. Now it is relevant to analyze the results of the best model of each entropy index in the normalized case in terms of the switching process and the level of the states.

The transition matrix of the Shannon entropy based on $\Delta CoVaR$ has1´s in all the diagonal showing than there is a strong evidence of switching between the three states. However, the transition matrix has zeros out of the diagonal, showing a low efficiency of the Markov Chain process, which implies that there is no relation between the state in one period and the state of the next period. This is the transition matrix associated with this model:

$$\boldsymbol{P} = \begin{pmatrix} 1.00 & 0.00 & 0.00 \\ 0.00 & 1.00 & 0.00 \\ 0.00 & 0.00 & 1.00 \end{pmatrix}$$

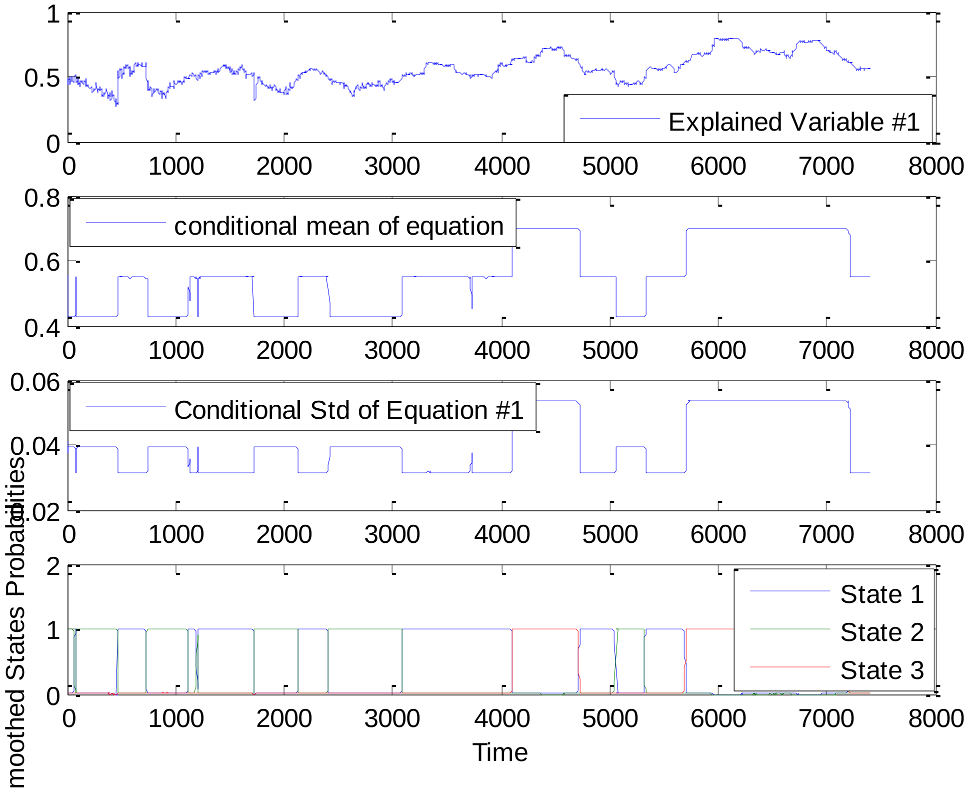


Figure 10: Smoothed states Probabilities for Shannon entropy based on $\Delta CoVaR$

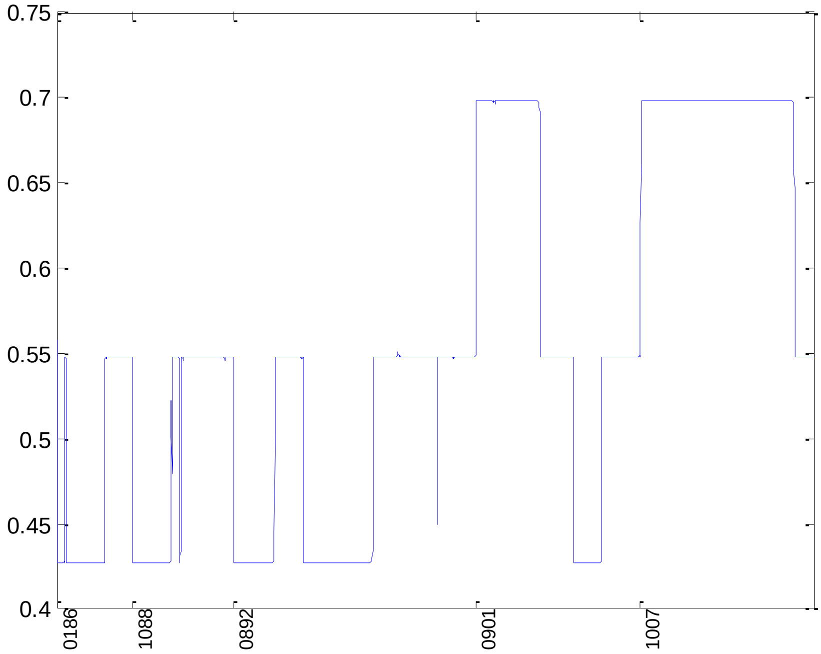


Figure 11: Conditional mean of Shannon entropy based on $\Delta CoVaR$

Even more the level of the entropy is higher in the state three with 0.6992 followed by state one 0.5484 and state two 0.4267. In particular the state three has two relevant period of crises especially around the observations 4000 (crises of 2001) and before the period 6000 (crises of 2008), for this model the crises finish in 2013.in this case I can say that the state two reflects no crises period and the state one could be considered as the bullish market before the crises. Finally it is important to remark that the expected duration for the state three is 1.4 years.

As regards the analysis of conditional mean related to this entropy index, the order is reverse for the lowest and mid-level because the lowest level of standard deviation corresponds to mid-level of entropy. And lowest level of entropy correspond midlevel of standard deviation, but when the level of entropy is high also the level of Standard deviation is high, when both are high the probability of systemic risk is high and that corresponds to period crises in 2001 and 2008 (see the figure 11), while the state one reflect in the case the period before the crises. Table 5 summarizes for each state the level and the value of both conditional mean and conditional standard deviation.

	Cond-mean	levels	Cond-std	levels
state1	0.5484	M	0.000988	L
state2	0.4267	L	0.001533	M
state3	0.6992	H	0.002844	H

Table 5: levels of both Cond-mean and the Cond-standard deviation with their values

For the Rényi entropy based on MES measure I also obtain some interesting results. , the three-state model has values different of zero in three of the six transitions probabilities out of the diagonal. However, the probabilities related to transition from the first and the second states to the third state are equal or close to zero indicating low relation between the states for two different periods. But the probability of switching from state three to state two is 15%, so there is a significant relation between these states, this is the transition matrix:

$$\boldsymbol{P} = \begin{pmatrix} 1.00 & 0.01 & 0.00 \\ 0.00 & 0.84 & 0.00 \\ 0.00 & 0.15 & 1.00 \end{pmatrix}$$

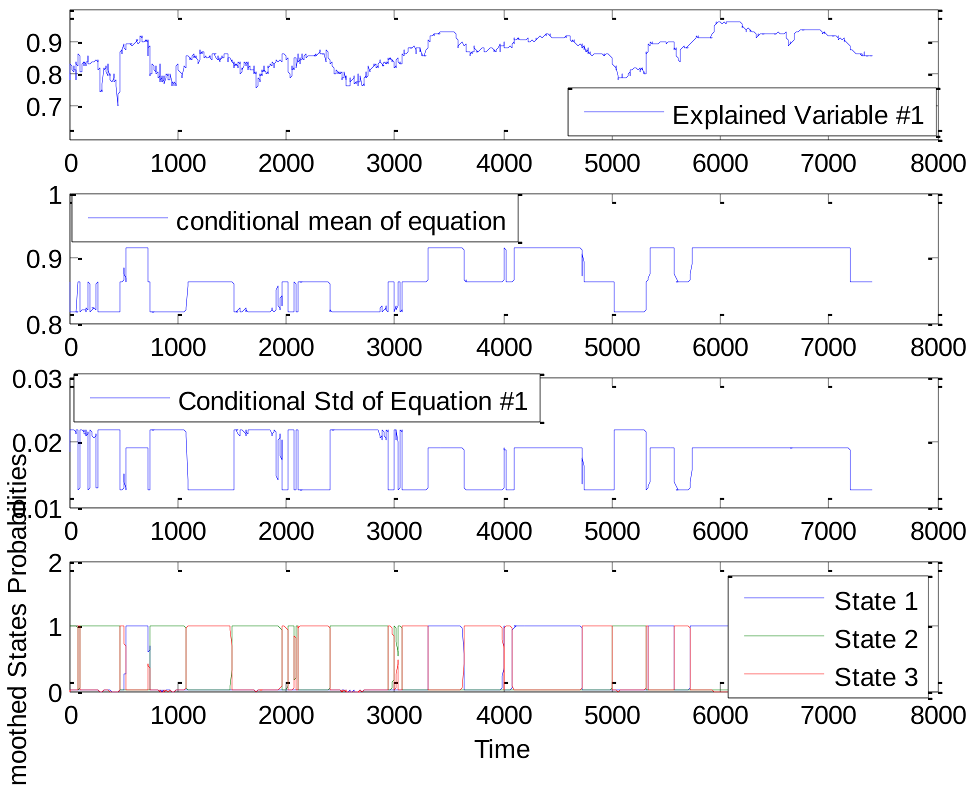


Figure 12: Smoothed states Probabilities for Rényi entropy based on MES

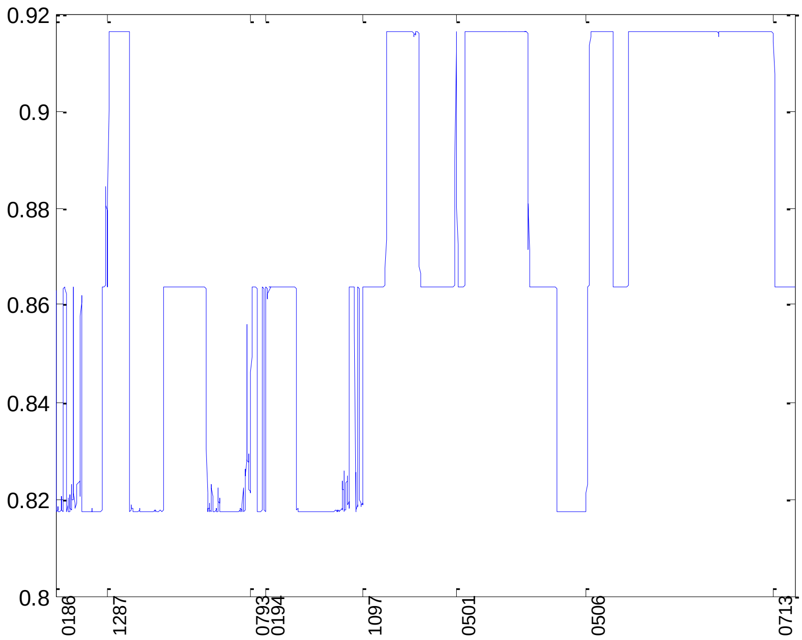


Figure 13: Conditional mean of Rényi entropy based on MES

Here the level of the entropy is increasing from the state two to state one as follow, 0.8083, 0.8636, and 0.9167. This model it does not recognize the crisis period because it includes the crises before the 4000 and 5000 period and after 1000 period as a conclusion I have a lot of small crises with higher frequency, and in particular this model it doesn't recognize the crises of 2008. Simultaneously the expected duration for this state is 5723.12 (23years).

Following the idea of analyzing the level of entropy in this case the lowest level of conditional mean corresponds to high level of standard deviation, high conditional mean corresponds to mid-level of standard deviation and finally mid-level of conditional mean corresponds to lowest level of standard deviation. The state one has the highest level of entropy and it appears in these years: at the end of 1987, after May 2001 and also after May 2008 but it last till 2013 with small jump to state three in 2009 which lasts a few months.

	Cond-mean	levels	Cond-std	levels
state1	0.9167	H	0.000363	M
state2	0.8083	L	0.000563	H
state3	0.8636	M	0.000157	L

Table 6: levels of both Cond-mean and the Cond-standard deviation with their values

For the In-Out measure the model with the best performance was the Rényi with a three states, for this case I have a probability of 3% to switch from state two to one and also for switching from three to two, I have no significant result between these two states and the state one.as show the transition matrix

$$\boldsymbol{P} = \begin{pmatrix} 1.00 & 0.03 & 0.00 \\ 0.00 & 0.97 & 0.03 \\ 0.00 & 0.00 & 0.97 \end{pmatrix}$$

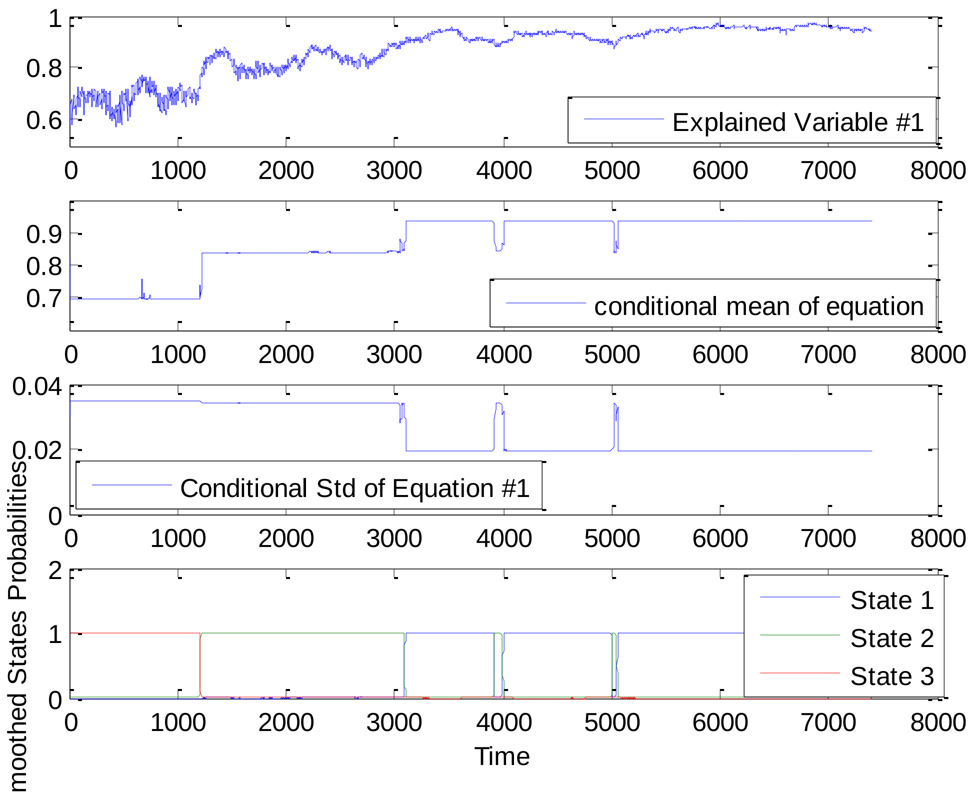


Figure 14: Smoothed states Probabilities for Rényi entropy based on In-Out

For this analysis I have a decrease of the entropy level from the state one to three, with 0.9378 for the state one and 0.8367 for state two and finally 0.6893 for the state three.in this case I have only the state one from 3000 period till the end with the appearance of small switches between the state two and state one, this model does not recognize properly the crises especially the crises of 2008. This model could be showing that the connectedness of the system has increased across the time making more difficulties to identify the crises using this risk measure.

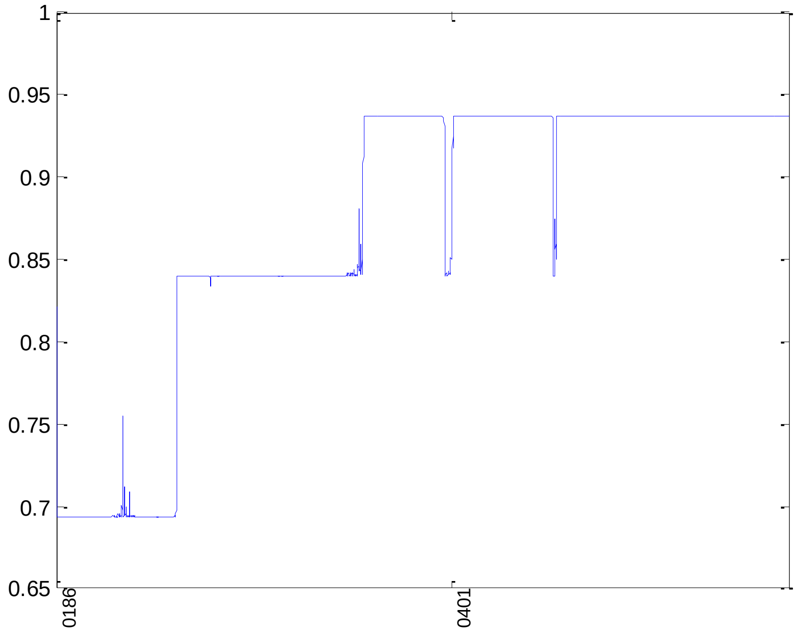


Figure 15: Conditional mean of Rényi entropy based on In-Out

	Cond-mean	levels	Cond-std	levels
state1	0.7338	H	0.002239	L
state2	0.6325	M	0.000292	M
state3	0.4457	L	0.018075	H

Table 7: levels of both Cond-mean and the Cond-standard deviation with their values

For the last model the level of entropy is different than the previous models because the lowest and the mid-levels of the conditional mean corresponds in the same order the high and the mid-levels of the conditional standard deviation, but as shown in the Figure 15 the difference between this two levels is not significant, as regard the high level of entropy it is clearly that the standard deviation has the lowest level. In this model since April 2001 till the end of the series the state one covers this period which recognizes the dot-com bubbles and the global financial crises. Additional remark in this period there is a jump but no change in the regime that could be possible because the conditional mean is not the value of the mean in the regime S_t but is an average with weights given the smoothed probability, consequently the probability of the state one is changing but it continue to be the highest of the states.

As a conclusion of this section I identify the ΔCoVaR three states Shannon entropy model as the best one for the identification of the crises, in spite of the relation between the states is not a strong as it is reflected in its transition matrix. On the other hand the MES and In-Out measure were not able to identify both properly the global financial crises even if the transition matrix shows a significant switching, however for the In-Out the switching process could be affected by the presence of positive trend of the connectedness of the system, thus it could be interesting to estimate the models after removing the trend.

7. MSCM Removing the trend

7.1 Tests

Let E_t be the observed time series which contain a unit root, in order to check the existence of these roots, I will apply two tests:

The ADF test and the PP test the following way to proceed can be determined beforehand:

- I Apply the Phillips-Perron (PP) which assesses the null hypothesis of a unit root in a univariate time series, Values of h equal to 1 indicate rejection of the unit-root null in favor of the alternative model. A value of h equal to 0 indicates a failure to reject the unit-root null.

- I Apply the Augmented Dickey Fuller (ADF) test to check the null hypothesis of unit root existence. If the null hypothesis is rejected, the conclusion is there is no unit root (stationary),

After the ADF test I got h = 0 for the all entropies with a pValue around *0.56* which indicate a failure to reject the null hypothesis, it means a failure to reject the existence of unit root. Similarly the PP rejects null hypothesis h = 0, as result all the entropies have a unit root. (See table 1)

In conclusion as all the entropies indexes are not stationary, all of them have at least two states in the Markov Chain process. In order to identify clearly the states I removed the trend by running a linear regression of each entropy index against the time and I applied the MCSM on the residuals of this regression, so this procedure will not affect the states of the original index as it is not changing the non-stationary pattern

before removing the trend		h_ADFtest	pValue	h-PPtest	pValue
ΔCoVaR	Shannon entropy	0	0.5147	0	0.5147
	Rényi entropy	0	0.5720	0	0.5720
	Tsallis entropy	0	0.4778	0	0.4778
MES	Shannon entropy	0	0.5877	0	0.5767
	Rényi entropy	0	0.6506	0	0.6033
	Tsallis entropy	0	0.5580	0	0.5580
In-Out	Shannon entropy	0	0.5767	0	0.5767
	Rényi entropy	0	0.6033	0	0.6033
	Tsallis entropy	0	0.4863	0	0.4863

Table 8: ADF and PP tests on entropies indexes before removing the trend

7.2 Descriptive statistics

The residuals look much better after removing the trend. Figures16-18 and Table 6 show the residuals for the all entropy indexes and the results of the Tests after removing the trend. First of all the indexes are moving around zero, also there is a change for the Rényi entropy that now behaves similar to the others unlike the descriptive analysis seen above.

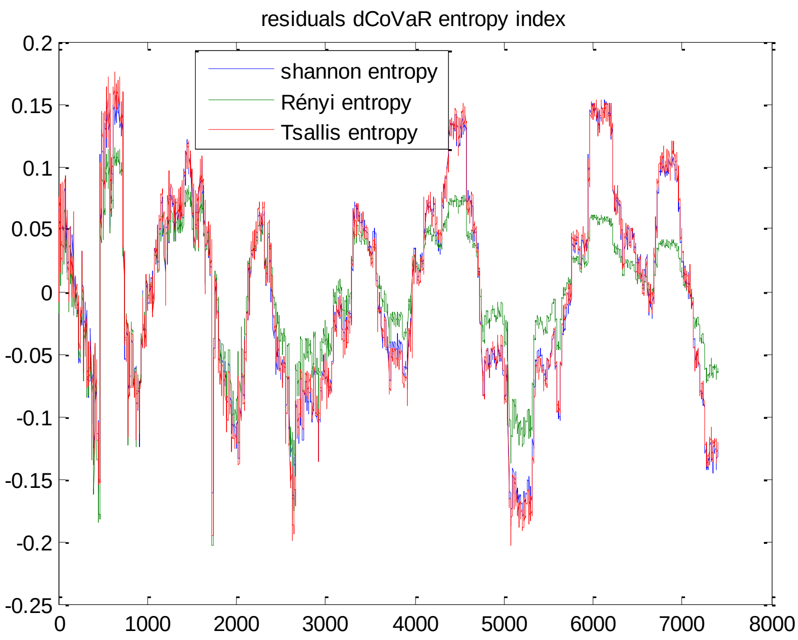


Figure 16: Residuals of entropy indexes based on ΔCoVaR

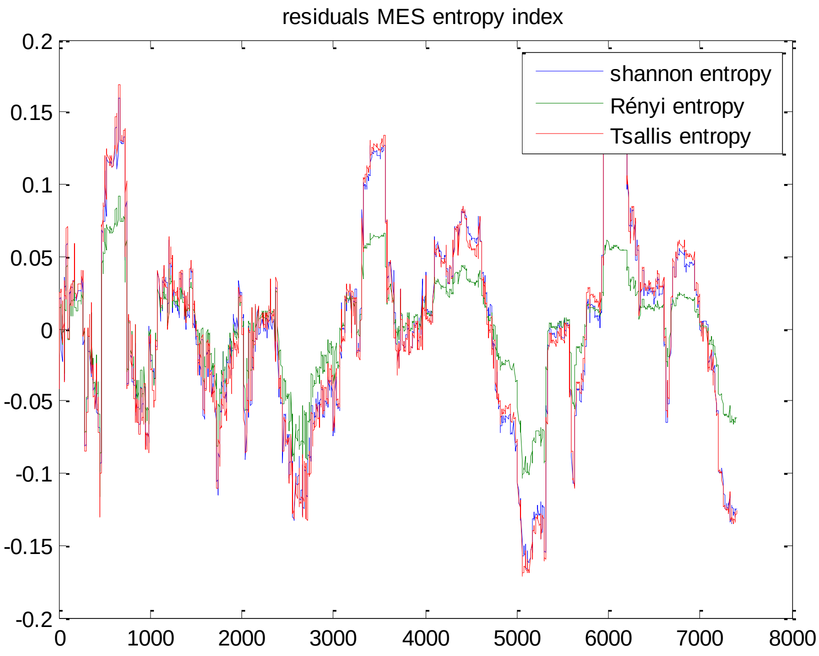


Figure 17: Residuals of entropy indexes based on MES

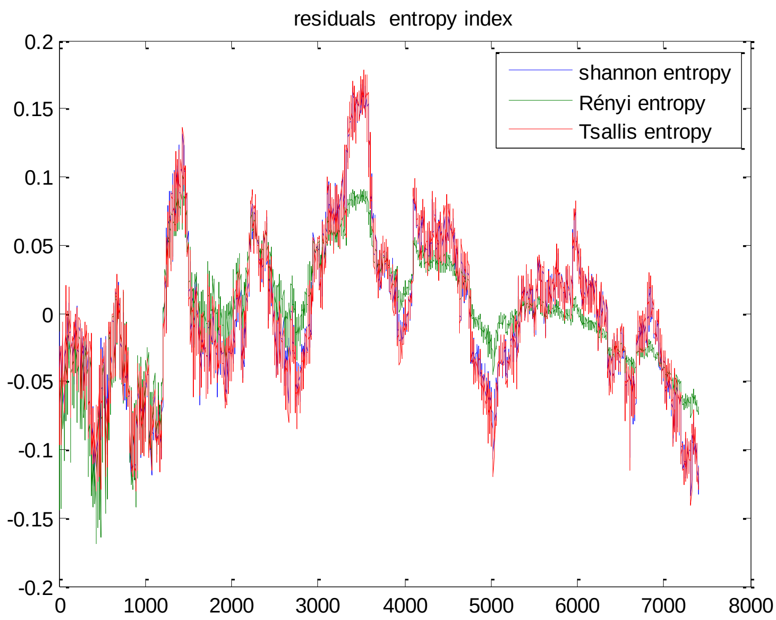


Figure 18: Residuals of entropy indexes based on In-Out

after removing the trend		ADFtest	pValue	PPtest	pValue
ΔCoVaR	Shannon entropy	1	1.0e-03	1	1.0e-03
	Rényi entropy	1	1.0e-03	1	1.0e-03
	Tsallis entropy	1	1.0e-03	1	1.0e-03
MES	Shannon entropy	1	0.0036	1	0.0036
	Rényi entropy	1	0.0010	1	0.0010
	Tsallis entropy	1	0.0015	1	0.0015
In-Out	Shannon entropy	1	1.0e-03	1	1.0e-03
	Rényi entropy	1	1.0e-03	1	1.0e-03
	Tsallis entropy	1	1.0e-03	1	1.0e-03

Table 9: ADF and PP tests on entropies indexes after removing the trend

As result the detrend has a positive effect on the entropies related In-Out measure in the sense that the series behaves with more stability that allows to identify the mean states, additionally the Tsallis is more volatile than the others.

7.3 Estimation Result

Knowing that the normalization increases the quality of the estimation I will just focus on the normalized indexes. Using the criteria (BIC) and (AIC) to compare the efficiency the models I got the following results:

For Shannon and Tsallis entropies based of In-Out measure in three states were the best models, for the Rényi entropy based on MES still the best model as the previous case (trend included). In this sense I can say that the detrend method is a good tool to clean the In-Out measure in order to increase its efficiency detecting the financial crises using the connectedness approach. Tables 7, 8 and 9 show the results.

	NORMALIZED	MEASURE	BIC	AIC
2 states	Shannon Entropy	ΔCoVaR	-19500	-19541
		MES	-24485	-24526
		In-Out	-25611	-25653
3 states	Shannon Entropy	ΔCoVaR	-20586	-20669
		MES	-26239	-26322
		In-Out	-26537	-26620

Table 10: (BIC) and (AIC) values for the residuals of Shannon entropy index in two and three states

	NORMALIZED	MEASURE	BIC	AIC
2 states	Renyi Entropy	ΔCoVaR	-26045	-26086
		MES	-32441	-32483
		In-Out	-30719	-30761
3 states	Renyi Entropy	ΔCoVaR	-27461	-27544
		MES	-33639	-33722
		In-Out	-30666	-30749

Table 11: (BIC) and (AIC) values for the residuals of Rényi entropy index in two and three states

	NORMALIZED	MEASURE	BIC	AIC
2 states	Tsallis Entropy	ΔCoVaR	-19019	-19061
		MES	-23848	-23890
		In-Out	-25290	-25331
3 states	Tsallis Entropy	ΔCoVaR	-20076	-20159
		MES	-25334	-25417
		In-Out	-26142	-26225

Table 12: (BIC) and (AIC) values for the residuals of Tsallis entropy index in two and three states

Now I will be focus on the results of the best model of each entropy indexes. I start by analyzing the Shannon entropy

$$\boldsymbol{P} = \begin{pmatrix} 0.99 & 0.00 & 0.01 \\ 0.00 & 0.99 & 0.00 \\ 0.01 & 0.01 & 0.98 \end{pmatrix}$$

The transition matrix in this case has a probability of 1% of switching from the state three to one and also of switching from state three to state one, but the other probabilities are null except the diagonal. Even more the smoothed states probability shows the presence of blocks of blank, in this case it is difficult to identify which state represent the crises. The state two shows the highest level of the standard deviation which corresponds to the lowest level of entropy, the state one in this model represents the highest regime. (See table 13 and Figure 13). Between April 2006 and January 2008 the switching was only between state one and three, this happened also between September 2008 and May 2010, the period of financial crises. In this case it is difficult to identify the dot-com bubble; it could be related with the contagion effect that usually has strong changes during the financial crises. But also this systemic risk measure could be sensible to other situation that affects the connectedness of the system like wars or political events. In contrast with the case including the trend the expected duration of the regimes is shorter. For example the Expected duration of Regime #2: 170.79 time periods.

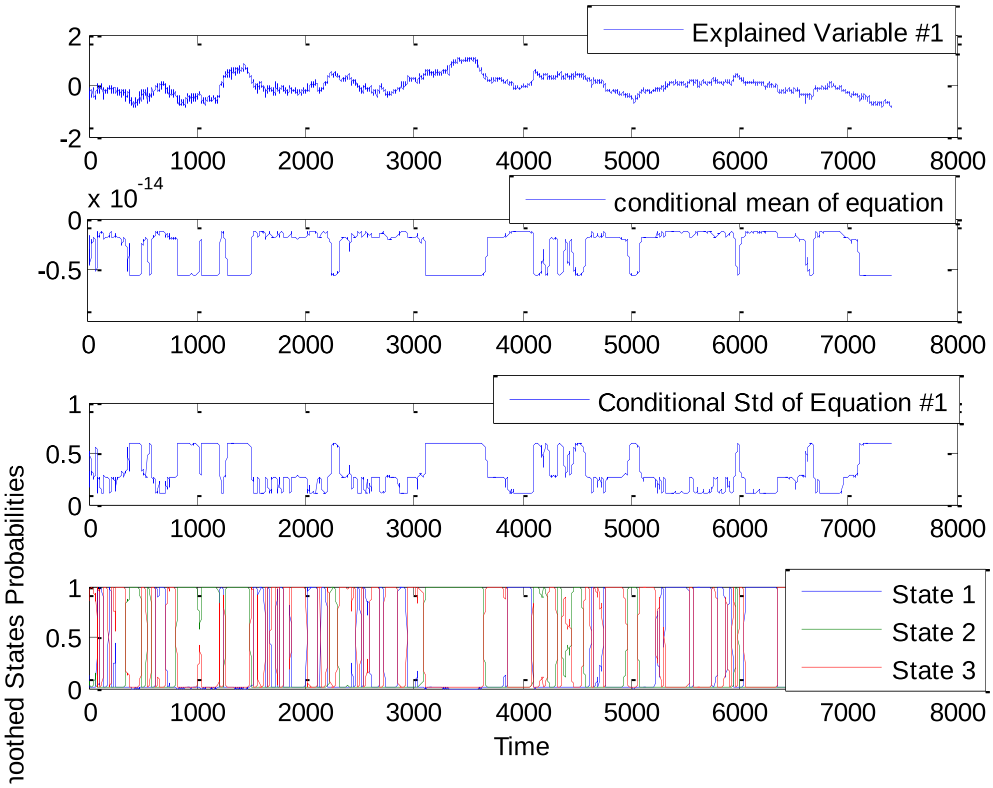


Figure 19: Smoothed states Probabilities for Shannon entropy based on In-Out

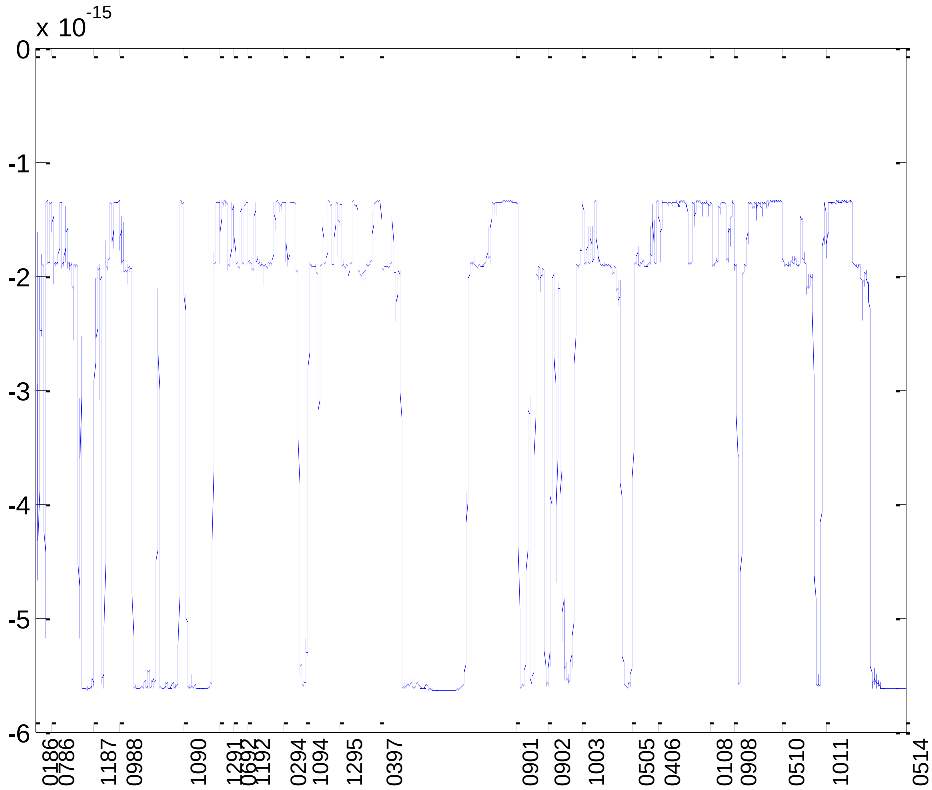


Figure 20: Conditional mean of residuals Shannon entropy based on In-Out

	Cond-mean	levels	Cond-std	levels
state1	0.0000	H	0.000234	L
state2	-0.0000	L	0.008169	H
state3	-0.0000	M	0.001583	M

Table 13: levels of both Cond-mean and the Cond-standard deviation with their values

The Tsallis entropy also identifies that the model of the In-Out as best the one. The transition matrix and the probabilities are the following:

$$\boldsymbol{P} = \begin{pmatrix} 0.98 & 0.01 & 0.00 \\ 0.02 & 0.98 & 0.01 \\ 0.00 & 0.01 & 0.99 \end{pmatrix}$$

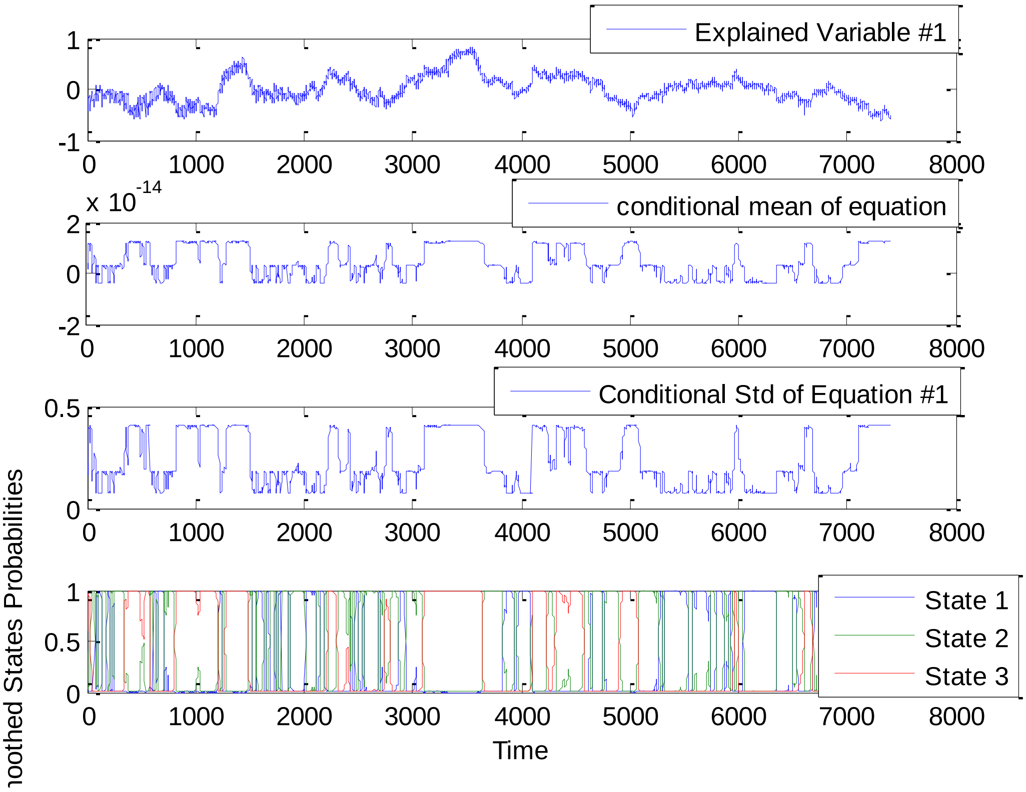


Figure 21: Smoothed states Probabilities for Tsallis entropy based on In-Out

As the transition Matrix shows, the probability of switching between the states is not significant. Focusing on the smoothed states probabilities, the switching process shows several blank that represents the presence of a crises especially after 3000 (crises 2001) and 6000 (crises 2008), also I can say that the states three represents a mid-level of entropy and high volatility with some persistence before 2001, this state has the longest expected duration: 180.92 time periods

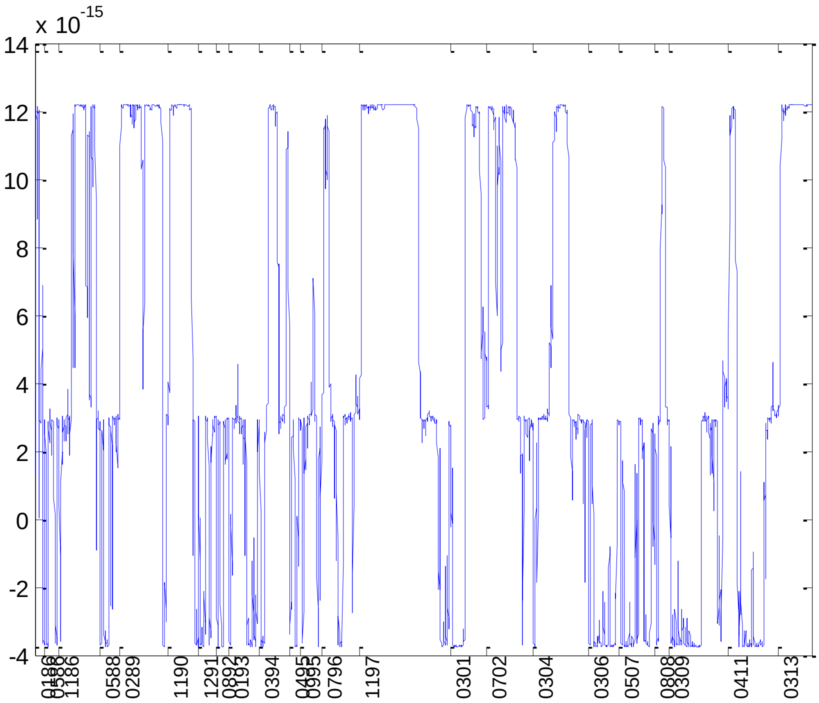


Figure 22: Conditional mean of residuals Tsallis entropy based on In-Out

Analyzing the conditional mean, I remake that the level of entropy and the level of standard deviation behave in the same way. In 1990, 1994, 1998 and 2008 the level of entropy is high before the dot-com bubble and before the global financial crises of 2008 but there is a jump from the lowest state to the highest in 2008.

Finally in the case of the Rényi entropy the best model was the MES for a three states as in the previous case the element out of the diagonal in transition matrix are not significant

$$\boldsymbol{P} = \begin{pmatrix} 0.98 & 0.00 & 0.01 \\ 0.02 & 0.99 & 0.00 \\ 0.00 & 0.01 & 0.99 \end{pmatrix}$$

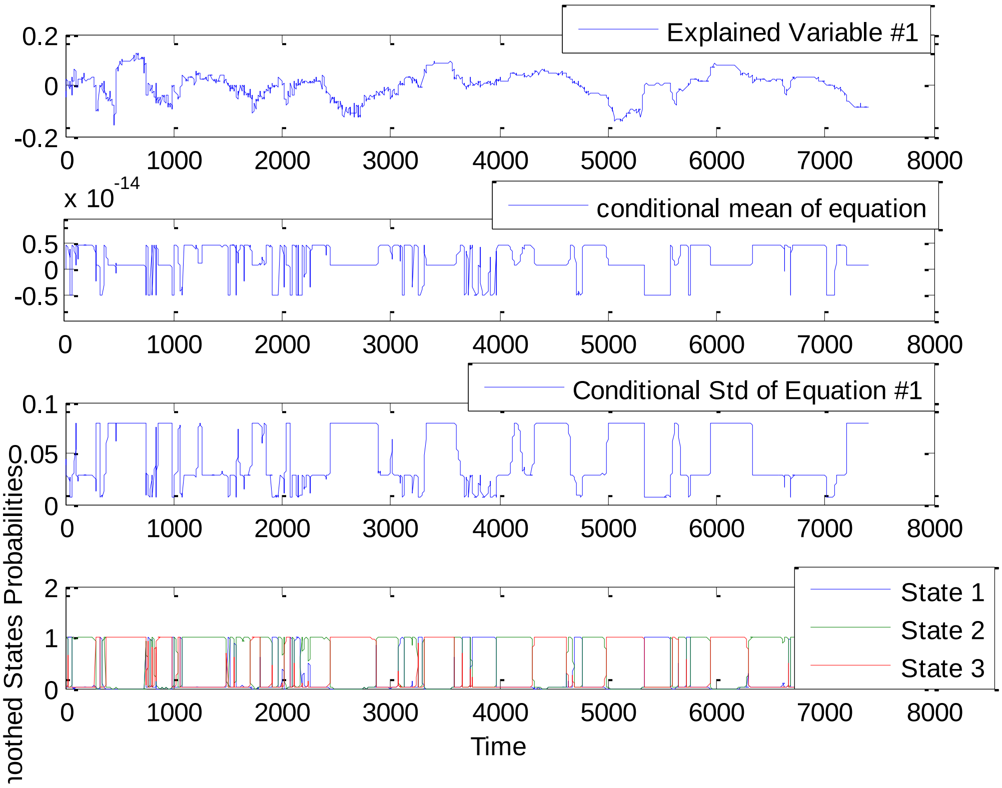


Figure 23: Smoothed states Probabilities for Rényi entropy based on MES

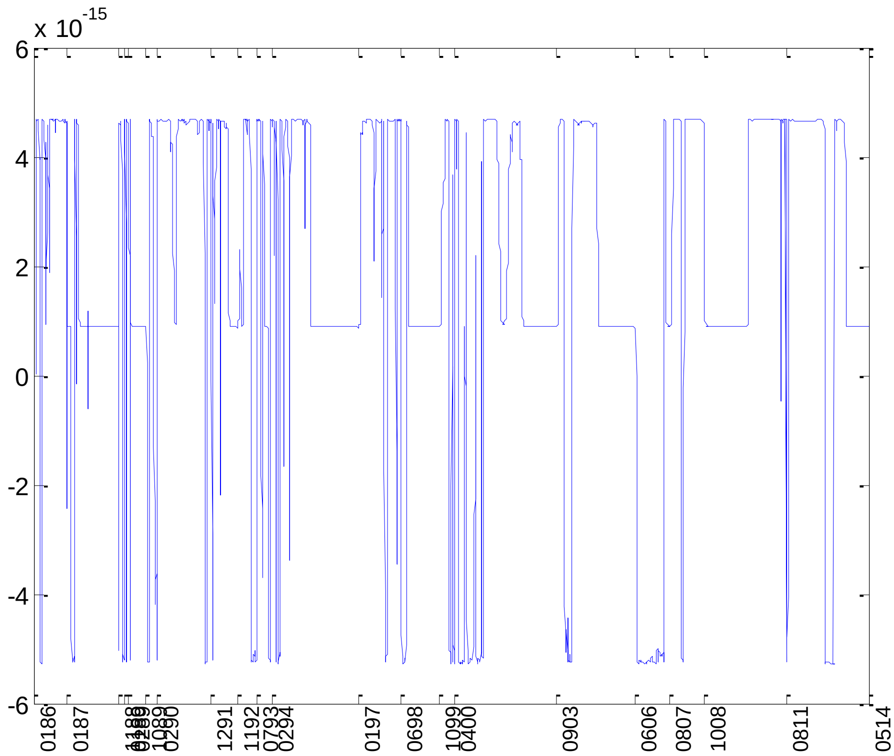


Figure 24: Conditional mean of residuals Rényi entropy based on MES

	Cond-mean	levels	Cond-std	levels
state1	0.0000	L	0.000234	L
state2	-0.0000	H	0.001472	M
state3	-0.0000	M	0.007721	H

Table 14: levels of both Cond-mean and the Cond-standard deviation with their values

The smoothed states probabilities for this entropy index show a persistent switching between the states that makes a difficult to detect the financial crises periods, this is represented by the presence of several blank in the switching process plot. As a conclusion comparing this model with the results in presence of the trend, this model is worst because it can not identify which state represents the financial crises; in fact the Rényi entropy based on the MES is more efficient without removing the trend.

Following the idea of analyzing the level of entropy in this case I have reverse results between the highest and mid-level of conditional mean and conditional standard deviation, but for the lowest they are the same, in this case the period of high entropy represented by state two. In some periods there only a switching between low and high state, for example between 1992-1994, 1997-1998, 1999-2000 and finally 2006-2007 which complicate the identification of which state corresponds to the crises.

As result of this part all the figures of conditional mean related to these series shows a high number of small jumps between the states which sometimes makes the identification of the crises periods complicated unlike the conditional means in the case of presence of trend the states lasted more longer and most of the time it was a state which could be considered as the bullish market before the crises. Additionally the detrend method was useful for the In-Out measure.

.

8. Conclusions

- Using Markov Chain Switching approach is useful because it is possible to identify the different regimes of market entropy.

- The estimation of the MCSM shows that is difficult to determine the probability of transmission between the states, this is represented by the presence of values close to one in the diagonal of transmission matrix of most of the models. The normalization improved the quality of the results, especially for the Shannon and Tsallis entropy but not for the Rényi entropy, it could be related to its formula.

- All the criteria show that the MCSM three states is more efficient than two states that implies the existence of one transition period between the crises and the stable period.

- The model with the best performance for the Shannon Rényi and Tsallis entropy are in order the $\Delta CoVaR$, MES and $\Delta CoVaR$. However in the case of Rényi and Tsallis the models were not able to recognize the crises period of 2001 and 2008, as result the entropy Shannon based on $\Delta CoVaR$ was the best one.

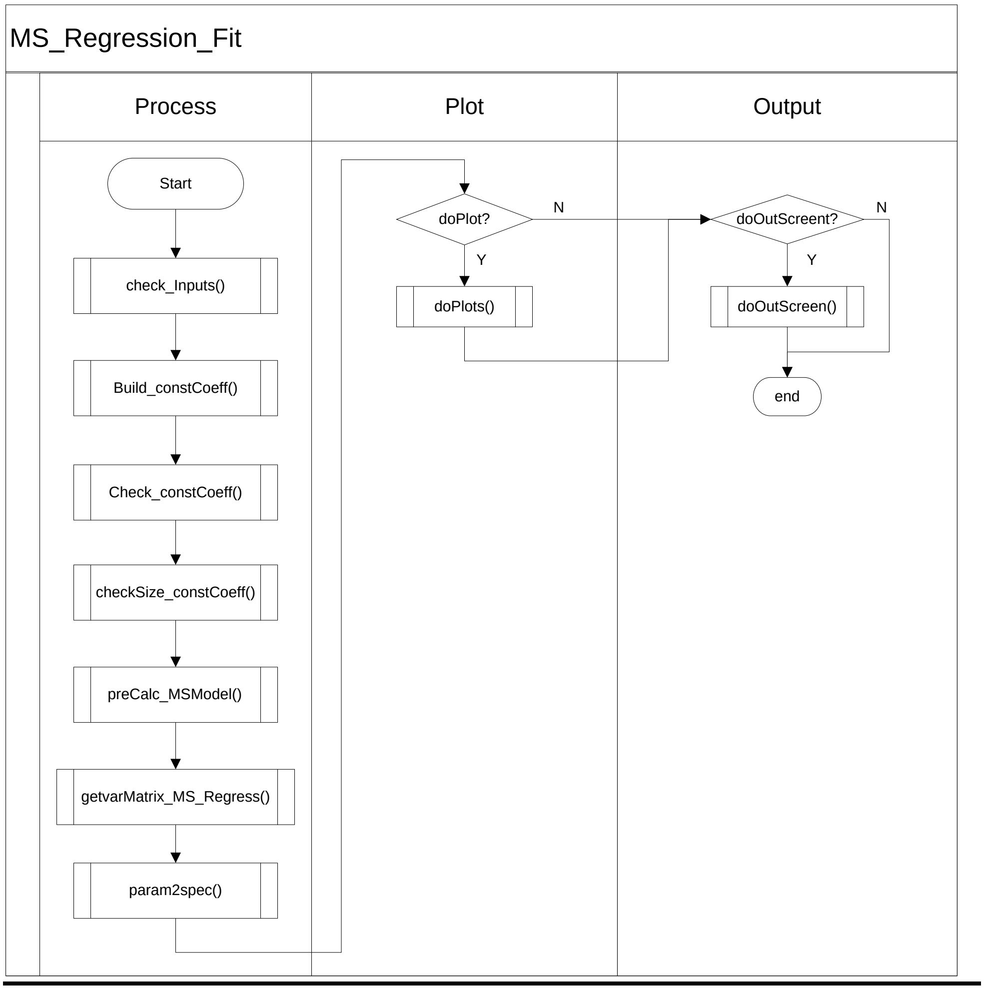
- After removing the trend the models with the best performance for the Shannon Rényi and Tsallis entropy are in order the In-Out, MES and In-Out.

- Analyzing the In-Out measure it is remarkable to see a positive trend for the whole analysis period. The detrend method increased the efficiency of this measure in sense that after applying this method it became the best measure for the Shannon and Tsallis entropy. The In-Out measure is more sensible to other situation that affects the connectedness of the system different than financial crises.

9. Appendix

I provide a brief description of the MATLAB library used for the estimation of the MCSM. I refer reader to Marcelo Perlin paper MS_Regress - The MATLAB Package for Markov Regime Switching Models, (version: April 19, 2015) for further details. The flow is explained based on the functions as the following flowchart.

.



To understand how this flowchart works, I will explain part by part in this chapter.

9.1 The process part

Before reaching the calculation, I need to provide some inputs:

Functions	**Explanation**
check_Inputs()	This function will check every given input from the outline.m file. It is started from the number of input function and then the availability of its variables in one object, which is advOpt, whether or not they are empty. The explanation of these variables itself is given below : - advOpt.distrib It defines a certain distribution to be used in the maximum likelihood calculation. - advOpt.std_method It defines the method to be used for the calculation for the calculation of the standard errors of the estimated coefficients. - advOpt.useMex It defines whether to use the mex version of Hamilton's filter in the calculation of likelihood function. - advOpt.diagCovMat It defines the use of a diagonal matrix for sigma (covariance matrix) in a multivariate estimation - advOpt.printOut Flag for printing out to screen the model in the end of estimation - advOpt.printIter Flag for printing out numerical iterations of maximum likelihood estimation - advOpt.doPlots Flag for plotting fitted conditional standard deviations and smoothed probabilities in the end of estimation. - advOpt.optimizer it defines which Matlab's optimizer to use in the maximum likelihood estimation of the model.
Build_constCoeff()	This function is to make the constants : - advOpt.constCoeff.covMat it defines covariance matrix of innovations - adv.constCoeff.nS_Param it defines for all switching coefficient at indep matrix (also chosen with input), in this case we use vector (1,1), which is the first order is switching mean and switching variance, and these are already given from the input. - adv.constCoeff.S_Param it defines for All switching coefficient at indep matrix. - adv.constCoeff.p it defines transition matrix
Check_constCoeff()	This function is to check whether the constants generated by Build_constCoeff() are not empty.
CheckSize_constCoeff()	This function use to check the size for every variable in

	constCoeff structure
preCalc_MSModel()	This is pre-calculation for switching model.
getvarMatrix_MS_Regress()	This function is to calculate standard errors of MS_Regress_Fit. The calculation (approximation) of the first and second derivative of the likelihood function is done by a two side finite differences method. The four methods for the calculation of the covariance matrix were implemented here : - Using the second partial derivatives - Using first partial derivatives (outer product matrix). In this function also does the MS_Regression_Lik that will be explained later in this table.
param2spec()	This function will check some parameter specification for the spec_out
MS_Regression_Lik	All of the models are estimated through this function using maximum likelihood. There are 3 outputs given by this function, which are : - sumlik Negative sum of log likelihood for fmincon. In addition, it minimizes the function. - Output This is structure variable which contains of - Coeff It is a structure variable with the same structure as consCoeff in Build_constCoeff(). - filtprob the filtered probabilities of regimes (iterated over the states (columns)) - LL Final log likelihood of model - k Number of states - param all estimated parameters in vector notation - S Switching flag control (iterating over equations (cell)) - advOpt - A structure variable with the same structure as in check_inputs(). - logLikeVec it is a logarithm of the function to calculate for maximum likelihood, in vector column

9.2 Plots part

Every calculation has done. It continues to the next step, doPlots(). The second part will do plotting some graph from its calculation. Each graph will include: plot of the inputs given, such as ΔCoVAR, MES, and In-Out respectively and plot of conditional mean followed by plot of conditional standard of Entropy. Finally the plot of the smoothed states probabilities. The best of those plots are already shown in the estimation part.

9.3 Summary result part

All results that been calculated will be summarized through doOutScreen(). From this part, it will show on command window, but only local variables that can be accessed in MS_Regression_Fit. Hence, even if we run the function outside, it will not show anything. Having finished doing calculation through this process, all variables will be saved in to this spec_output structure variable:

The content of Structure Variables	Explanation
SsmoothProb	Smoothed probabilities of regimes (iterated over the states (columns))
nObs	Number of Observations (rows) in the model
nEq	Number of dependent variables
Number_Parameters	Number of estimated parameters
advOpt.distrib	The type of distribution used for this calculaton, in this case is Normal Distribution.
advOpt.std_method	The same explanation in check_Inputs()
Coeff_SE	A structure with all standard errors of coefficients (same field as Coeff)
Coeff_pValues	A structure with all parameter's p-values (same fields as Coeff)
AIC	Akaike information criteria of the estimated model
BIC	Bayesian information criteria for estimated model

9.4 MATLAB code

```
clc;
clear all;
close all;
clf;

load('entropy_final.mat')

% specify beginning and ending as string variables
dateBeg = '725374'; % day, month, year: ddmmyyyy
dateEnd = '735731'; % day, month, year: ddmmyyyy

% dynamic assignment to end of period
dateBeg = datestr(725374,'ddmmyyyy') % today as first date
dateEnd = datestr(735731,'ddmmyyyy') % today as last date

% Take the data from normality entropy
for i = 1:3,
    dCoVaRentropy(:,i) = dCoVaRentropy_n(:,i);
    MESentropy(:,i)     = MESentropy_n(:,i);
    INOUTentropy(:,i) = INOUTentropy_n(:,i);
end

% The name of Entropy Formula of entropy
methods     = ['Shannon';'Rényi  ';'Tsallis'];
cell_mthds = cellstr(methods);
type        = ['dCoVar';'MES   ';'INOUT '];
cell_ent    = cellstr(type);

%-------------------------Plot the Entropy--------------------------
-------
for j = 1:3,
    plotEntropy
end


% Descreptive statistics----------------------------------------------
--------------------------------------------------------------------
-------------------
for j = 1:3  % Type : dCoVaR,MES,INOUT
    entropy_Data = 0;
    for i =1:3, % Methods : Shannon,Renyi,Tsallis
        change_entropy;
        figure(j+3);
        subplot(3,1,i);
        [meanRet1(j,i),stdDev1(j,i)] =
entropyCalc(entropy_Data,cell_mthds(i),cell_ent(j));
    end
```

```
end


%------------------------Plot First Difference---------------------
------%
for j = 1:3, % Type : dCoVaR,MES,INOUT
    entropy_data = 0;
    figure(j+6);
    hold on
    for i=1:3, % Methods : Shannon,Renyi,Tsallis
        plotfirstdiff;
    end
end


for i=1:3, % Methods : Shannon,Renyi,Tsallis
    [h_adfC(i),pValue_adfC(i)]    =
adftest(dCoVaRentropy_n(:,i),'alpha',0.05)
    [h_adfM(i),pValue_adfM(i)]    =
adftest(MESentropy_n(:,i),'alpha',0.05)
    [h_adfI(i),pValue_adfI(i)]    =
adftest(INOUTentropy_n(:,i),'alpha',0.05)
    [h_ppC(i),pValue_ppC(i)] =
pptest(dCoVaRentropy_n(:,i),'alpha',0.05)
    [h_ppM(i),pValue_ppM(i)] =
pptest(MESentropy_n(:,i),'alpha',0.05)
    [h_ppI(i),pValue_ppI(i)] =
pptest(INOUTentropy_n(:,i),'alpha',0.05)
end


%------------------------------------------------------------------
--------------------------------------------------------------------
-------------------
T = 7398;
t =[1:T]'
for j = 1:3,  % Type : dCoVaR,MES,INOUT
    for i = 1:3, % Methods : Shannon,Renyi,Tsallis
        change_entropy;
        [r(j,i),m(j,i),b(j,i)] = regression(t,entropy_Data,'one')
        switch j
            case 1
                e1(:,i) = entropy_Data - m(j,i)*t-b(j,i);
                [h1,pValue1]= adftest(e1(:,i),'alpha',0.05)
            case 2
                e2(:,i) = entropy_Data - m(j,i)*t-b(j,i);
                [h2,pValue2]= adftest(e2(:,i),'alpha',0.05)
            otherwise
                e3(:,i) = entropy_Data - m(j,i)*t-b(j,i);
                [h3,pValue3]= adftest(e3(:,i),'alpha',0.05)
        end
    end
end
```

```
% for j = 1:3,
%     plot(e1)
%     legend('shannon entropy','Rényi entropy','Tsallis entropy')
%
% end


for i=1:3, % Methods : residuals Shannon,Renyi,Tsallis
    [h_adfC(i),pValue_adfC(i)]    = adftest(e1(:,i),'alpha',0.05)
    [h_adfM(i),pValue_adfM(i)]    = adftest(e2(:,i),'alpha',0.05)
    [h_adfI(i),pValue_adfI(i)]    = adftest(e3(:,i),'alpha',0.05)
    [h_ppC(i),pValue_ppC(i)] = pptest(e1(:,i),'alpha',0.05)
    [h_ppM(i),pValue_ppM(i)] = pptest(e2(:,i),'alpha',0.05)
    [h_ppI(i),pValue_ppI(i)] = pptest(e3(:,i),'alpha',0.05)
end


% plot(e1)
% legend('shannon entropy','Rényi entropy','Tsallis entropy')
% title('residuals dCoVaR entropy index')
% plot(e2)
% legend('shannon entropy','Rényi entropy','Tsallis entropy')
% title('residuals MES entropy index')
% plot(e3)
% legend('shannon entropy','Rényi entropy','Tsallis entropy')
% title('residuals  entropy index')


% ESTIMATION---------------------------------------------------
------------------------------------------------------------------
--------------------
i = 2; %number of the column related to each entropy index (Shannon,
Rényi,Tsallis)
load 'entropy_final.mat'
logRet.CoVaR     = dCoVaRentropy(:,i);
logRet.CoVaRNorm = dCoVaRentropy_n(:,i);
logRet.MES       = MESentropy(:,i);
logRet.MESNorm   = MESentropy_n(:,i);
logRet.INOUT     = INOUTentropy(:,i);
logRet.INOUTNorm = INOUTentropy_n(:,i);
dep.CoVar        = logRet.CoVaR;
dep.CoVaRNorm    = logRet.CoVaRNorm;
dep.MES          = logRet.MES;
dep.MESNorm      = logRet.MESNorm;
dep.INOUT        = logRet.INOUT;
dep.INOUTNorm    = logRet.INOUTNorm;

constVec = ones(length(dep.CoVar),1);
indep = constVec;
k=2;
S=[1,1];
advOpt.distrib='Normal';
advOpt.std_method=1;
```

```
hold on
%Calling fitting function
display('*****************************dCoVar Entropy Computation********************************')
figure(1)
set(gcf,'NumberTitle','off')
set(gcf,'name','dCoVar Entropy Computation')
Spec_out_dCoVar     = MS_Regress_Fit(dep.CoVar,indep,k,S,advOpt)
display('**********************dCoVar Entropy Normalized Computation****************************')
figure(2)
set(gcf,'NumberTitle','off')
set(gcf,'name','dCoVar Entropy Normalized Computation')
Spec_out_dCoVarNorm = MS_Regress_Fit(dep.CoVaRNorm,indep,k,S,advOpt)
display('*******************************MES Entropy Computation*********************************')
figure(3)
set(gcf,'NumberTitle','off')
set(gcf,'name','MES Entropy Computation')
Spec_out_MES        = MS_Regress_Fit(dep.MES,indep,k,S,advOpt)
display('***********************MES Entropy Normalized Computation******************************')
figure(4)
set(gcf,'NumberTitle','off')
set(gcf,'name','MES Entropy Normalized Computation')
Spec_out_MESNorm    = MS_Regress_Fit(dep.MESNorm,indep,k,S,advOpt)
display('****************************INOUT Entropy Computation**********************************')
figure(5)
set(gcf,'NumberTitle','off')
set(gcf,'name','INOUT Entropy Computation')
Spec_out_INOUT      = MS_Regress_Fit(dep.INOUT,indep,k,S,advOpt)
display('***********************INOUT Entropy Normalized Computation****************************')
figure(6)
set(gcf,'NumberTitle','off')
set(gcf,'name','INOUT Entropy Normalized Computation')
Spec_out_INOUTNorm  = MS_Regress_Fit(dep.INOUTNorm,indep,k,S,advOpt)
hold off
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%
[StaMa,StaInd]=max(Spec_out_dCoVarNorm.smoothProb');
dt=datestr(dates,'ddmmyyyy');
df=((StaInd(2:end)-StaInd(1:end-1))~=0)';
T=size(dates,1);
nn=(1:T)';
%stairs(StaInd');
plot(Spec_out_dCoVarNorm.condMean');
xlim([1,T]);
%ylim([-0.1,3.1]);
%yticks([1 2 3]);


dtc=[dt(1,:);dt(df,:);dt(end,:)];
```

```
dn=size(dtc,1);
Xt=[nn(1);nn(df,1);nn(T)];
dtc=dtc(1:4:end,:);
Xt=Xt(1:4:end,:);
set(gca,'XTick',Xt);

Xl=[1,T];
ax = axis;     % Current axis limits
axis(axis);     % Set the axis limit modes (e.g. XLimMode) to manual
Yl = ax(3:4);  % Y-axis limits

% Remove the default labels
set(gca,'XTickLabel','')
% Place the text labels
t =
text(Xt,Yl(1)*ones(1,length(Xt))*1.00001,num2str(dtc(:,[3,4,7,8])));
set(t,'HorizontalAlignment','right','VerticalAlignment','top', ...
      'Rotation',90,'Fontsize',8);
% Get the Extent of each text object.  This
% loop is unavoidable.
for i = 1:length(t)
  ext(i,:) = get(t(i),'Extent');
end
% Determine the lowest point.  The X-label will be
% placed so that the top is aligned with this point.
LowYPoint = min(ext(:,2));
% Place the axis label at this point
XMidPoint = Xl(1)+abs(diff(Xl))/2;
tl = text(XMidPoint,LowYPoint,'', ...
          'VerticalAlignment','top', ...
          'HorizontalAlignment','center');

%set(gca,'XTickLabel',num2str(dtc(:,[3,4,7,8])),'Fontsize',8);

% ----------------residuals ESTIMATION------------------------------
-------
i = 3
logRet.CoVaRNorm = e1(:,i);
logRet.MESNorm   = e2(:,i);
logRet.INOUTNorm = e3(:,i);
dep.CoVaRNorm     = logRet.CoVaRNorm;
dep.MESNorm       = logRet.MESNorm;
dep.INOUTNorm     = logRet.INOUTNorm;

constVec = ones(length(dep.CoVaRNorm),1);
indep = constVec;
k = 3;
S=[1,1];
advOpt.distrib='Normal';
advOpt.std_method=1;
hold on
%Calling fitting function
display('*******Residuals dCoVar Entropy Normalized
Computation**********')
```

```
figure(10)
set(gcf,'NumberTitle','off')
set(gcf,'name','dCoVar Entropy Normalized Computation')
Spec_out_dCoVarNorm = MS_Regress_Fit(dep.CoVaRNorm,indep,k,S,advOpt)
display('*******Residuals MES Entropy Normalized
Computation*************')
figure(11)
set(gcf,'NumberTitle','off')
set(gcf,'name','MES Entropy Normalized Computation')
Spec_out_MESNorm     = MS_Regress_Fit(dep.MESNorm,indep,k,S,advOpt)
display('******Residuals INOUT Entropy Normalized
Computation************')
figure(12)
set(gcf,'NumberTitle','off')
set(gcf,'name','INOUT Entropy Normalized Computation')
Spec_out_INOUTNorm   = MS_Regress_Fit(dep.INOUTNorm,indep,k,S,advOpt)
hold off
```

10.References

- Acharya, V. V., L. H. Pedersen, T. Philippon, and M. P. Richardson (2010). Measuring systemic risk. Working Paper, New York University.

- Adrian, T. and M. K. Brunnermeier (2011). Covar. Staff Report, Federal Reserve Bank, New York (348).

- Billio, M., M. Getmansky, A. W. Lo, and L. Pelizzon (2012). Econometric measures of connectedness and systemic risk in the finance and insurance sectors. Working Papers, Department of Economics, Ca' Foscari University of Venice (21).

- Brunnermeier, M.K., Dong, G., & Palia, D. (2012). *Banks' Non-Interest Income and Systemic Risk.* (AFA 2012 Chicago Meetings Paper).

- Cecchetti, S.G., Lam, P., Mark, N.C. (1990). Mean reversion in equilibrium asset prices. *American Economic Review* 80, 398-418.

- Getmansky, Mila, Andrew Lo, and Igor Makarov, 2004, "An Econometric Analysis of Serial Correlation and Illiquidity in Hedge-Fund Returns", *Journal of Financial Economics* 74, 529–609

- Diebold, F. and Yilmaz, K. (2014). On the Network Topology of Variance Decompositions: Measuring the Connectedness of Financial Firms. *Journal of Econometrics*, 182(1):119–134.

- Diebold, F.X., Yılmaz, K. (2009). Measuring financial asset return and volatility spillovers, with application to global equity markets. *Economic* J. 119, 158–171.

.

- Garcia, R. and P. Perron (1996). An analysis of the real interest rate under regime shifts, *Review of Economics and Statistics*, 78, 111-125.

- Goldfeld, S.M. and R.E. Quandt (1973). A Markov model for switching regressions, *Journal of Econometrics*, 1, 3–16.

- Gray, S.F. (1996). Modeling the conditional distribution of interest rates as a regime-switching process, *Journal of Financial Economics*, 42, 27–62.

- Hamilton, J.D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle, *Econometrica*, 57, 357-384.

- Hamilton, J.D. and G. Lin (1996). Stock market volatility and the business cycle, *Journal of Applied Econometrics*, 11, 573–593.

- Hamilton, J.D. and R. Susmel (1994). Autoregressive conditional heteroscedasticity and changes in regime, *Journal of Econometrics*, 64, 307–333.

- Kaufmann, S., 2002, Is there an Asymmetric Effect of Monetary Policy Over Time? A Bayesian Analysis using Austrian Data, *Empirical Economics* 27, 277-297.

- Lo, M. and J. Piger, (2005), "Is the Response of Output to Monetary Policy Asymmetric? Evidence from a Regime-Switching Coefficients Model," *Journal of Money, Credit and Banking* 37, 865-887.

- Maszczyk, T. and W. Duch (2008). Comparison of Shannon, Renyi and Tsallis entropy used in decision trees. *In Artificial Intelligence and Soft Computing*– ICAISC 2008, pp. 643–651. Springer.

- Pagan, A., Schwert, G.W. (1990). Alternative models for conditional stock volatility. *Journal of Econometrics* 45, 267-90.

- PERLIN Marcelo. (2015) MS_Regress -*The MATLAB Package for Markov Regime Switching Models.*

- Quandt, R. E. (1958). The Estimation of the Parameters of a Linear Regression System Obeying Two Separate Regimes. *Journal of the American Statistical Association*, 53(284), 873-880.

- Quandt, R.E. (1972). A new approach to estimating switching regressions, *Journal of the American Statistical Association*, 67, 306–310.

- Rényi, A. (1960). On measures of information and entropy. *In Proceedings of the 4th Berkeley Symposium on Mathematics, Statistics and Probability*, pp. 547–561. Also available at http://digitalassets.lib.berkeley.edu/math/ucb/text/math_s4_v1_article-27.pdf

- Shannon, C. E. (1948, July). A mathematical theory of communication. *The Bell System Technical Journal*, 27(3), 379–423.

- Tong, H. (1983) *Threshold Models in Non-linear Time Series Analysis*, Springer Verlag, New York.

- Tong, H. (1990). Non-linear Time Series: *A Dynamical System Approach*, New York, NY: Oxford University Press.

- Tsallis, C. (1988). Possible generalization of Boltzmann-Gibbs statistics. *Journal of Statistical Physics*, 52(1-2), 479–487.

- Turner, C. M., Startz, R. and C.R. Nelson, (1989), A Markov Model of Heteroskedasticity, Risk, and Learning in the Stock Market, *Journal of Financial Economics*, 25, 3-22.

- Timmermann, A., 2000. Moments of Markov switching models. *Journal of Econometrics* 96, 75-111.

- Zivot, E. and J., W. (2003). Modelling Financial Time Series with S-Plus R. Springer.