FERNANDO FERNANDES NETO

Essays on long memory processes

São Paulo

2017

FERNANDO FERNANDES NETO

Essays on long memory processes

Tese apresentada à Escola Politécnica da Universidade de São Paulo para obtenção do título de Doutor em Ciências

Orientador:
Prof. Dr. Claudio Garcia

São Paulo

2017

FERNANDO FERNANDES NETO

Essays on long memory processes

Tese apresentada à Escola Politécnica da Universidade de São Paulo para obtenção do título de Doutor em Ciências

Área de Concentração:
Engenharia de Sistemas

Orientador:
Prof. Dr. Claudio Garcia

São Paulo

2017

Este exemplar foi revisado e corrigido em relação à versão original, sob responsabilidade única do autor e com a anuência de seu orientador.

São Paulo, ______ de ____________________ de __________

Assinatura do autor: ________________________

Assinatura do orientador: ________________________

Catalogação-na-publicação

Fernandes Neto, Fernando
Essays on long memory processes / F. Fernandes Neto -- versão corr. -- São Paulo, 2017.
139 p.

Tese (Doutorado) - Escola Politécnica da Universidade de São Paulo. Departamento de Engenharia de Telecomunicações e Controle.

1.Análise de séries temporais 2.Contaminação 3.Crise econômica 4.Risco 5.Sistema Financeiro I.Universidade de São Paulo. Escola Politécnica. Departamento de Engenharia de Telecomunicações e Controle II.t.

RESUMO

O presente trabalho tem como objetivo discutir os principais aspectos teóricos ligados à ocorrência dos processos de memória longa e sua respectiva aplicação em economia e finanças. Para discutir os principais aspectos teóricos da sua ocorrência, recorre-se primeiramente à abordagem de sistemas complexos e fenômenos emergentes, tendo em vista que muitos destes são irredutíveis computacionalmente, ou seja, o estado atual do sistema depende de todos os estados anteriores, tal que, qualquer mudança nos instantes iniciais deve causar significativa diferença nos estados posteriores. Em outras palavras, há uma persistência da informação – conceito este intimamente ligado à memória longa. Portanto, com base em simulações de sistemas complexos computacionais, três fatores (podendo haver outros mais) foram relacionados ao surgimento de processos de memória longa: heterogeneidade dos agentes, ocorrência de grandes desvios do equilíbrio do sistema (em consonância com as respectivas leis do movimento de cada sistema estudado) e a complexidade espacial (que deve influenciar na propagação da informação e na dinâmica competitiva dos agentes). Em relação à aplicação do conhecimento, primeiro é reconhecido que os fatores explicativos para o surgimento de processos de memória longa são inerentes a estruturas/características de mercados reais e que é possível identificar potenciais fatos estilizados, ao filtrar as componentes de memória longa de séries temporais – grande parte da informação presente nas séries é função da estrutura de autocorrelação que advém das especificidades de cada mercado. Com base nisso, nesta tese foi desenvolvida uma nova técnica de estimação de contágio de risco, que não necessita intervenções adicionais, tendo em vista a identificação prévia de potenciais fatos estilizados em diferentes economias, utilizando as séries filtradas de variância condicional, tal que a partir destas séries filtradas é calculada uma correlação com horizonte móvel de observações entre choques (picos de risco) de curto prazo livres de possíveis reações causadas por idiossincrasias de cada mercado. Posteriormente, com base na identificação dos episódios ligados à Crise do Subprime de 2007/2008 nos Estados Unidos e seu respectivo contágio para a economia brasileira, filtrou-se a variância condicional dos ativos brasileiros (que é uma medida de incerteza), objetivando-se eliminar os eventos de contágio e, consequentemente, foi feita uma projeção contrafactual da evolução da economia, caso os episódios da crise não tivessem ocorrido. Com base nestes dados e com uma análise da tendência evolutiva da economia brasileira no período anterior à crise, constatou-se que 70% da crise econômica vivenciada no Brasil no período pós-2008 é decorrente de falhas na condução da política macroeconômica e somente 30% decorre dos efeitos do cenário externo na economia.

Palavras-chave: Atividade econômica. Contágio de risco. Crise do Subprime. Incerteza. Memória longa. Séries temporais. Sistemas complexos.

.

ABSTRACT

The present work aims at discussing the main theoretical aspects related to the occurrence of long memory processes and its respective application in economics and finance. In order to discuss the main theoretical aspects of its occurrence, it is worth starting from the complex systems approach and emergent phenomena, keeping in mind that many of these are computationally irreducible. In other words, the current state of the system depends on all previous states, in such a way that any change in the initial configuration must cause a significant difference in all posterior states. That is, there is a persistence of information over time – this is a concept directly related to long memory processes. Hence, based on complex systems simulations, three factors (possibly there are many others) were related to the rise of long memory processes: agents' heterogeneity, occurrence of large deviations from the steady states (in conjunction with the motion laws of each system) and spatial complexity (which must influence on information propagation and on the dynamics of agents competition). In relation to the applied knowledge, first it is recognized that the explanatory factors for the rise of long memory processes are common to the structures/characteristics of real markets and it is possible to identify potential stylized facts when filtering the long memory components from time series – a considerable part of information present in time series is a consequence of the autocorrelation structure, which is directly related to the specificities of each market. Given that, in this thesis was developed a new risk contagion technique that does not need any further intervention. This technique is basically given by the calculation of rolling correlations between long memory filtered series of the conditional variances for different economies, such that these filtered series contain the stylized facts (risk peaks), free from possible overreactions caused by market idiosyncrasies. Then, based on the identification of risk contagion episodes related to the 2007/2008 Subprime Crisis in the U.S. and its respective contagion to the Brazilian economy, it was filtered out from the conditional variance of the Brazilian assets (which are an uncertainty measure) aiming at eliminating the contagion episodes and, consequently, it was made a counterfactual projection of what would have happened to the Brazilian economy if the risk contagion episodes had not occurred. Moreover, in conjunction with the evolutionary trend of the Brazilian economy prior to the crisis, it is possible to conclude that 70% of the economic crisis posterior to the 2008 events was caused by macroeconomic policies and only 30% is due to the effects of risk contagion episodes from the U.S.

Keywords: Economic activity. Risk contagion. Subprime Crisis. Uncertainty. Long memory. Time series. Complex systems.

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS AND ACRONYMS

ACF	Autocorrelation function
ARIMA	Autoregressive integrated moving average model
ARMA	Autoregressive moving average model
BEKK	Baba-Engle-Kraft-Kroner GARCH model
DCC	Dynamic conditional correlation model
DCLMF	Dynamic correlation based on long memory filter
DSGE	Dynamic stochastic general equilibrium model
GARCH	General autoregressive conditional heterokedastic
GDP	Gross domestic product
GPH	Geweke and Porter-Hudak statistical model
GVAR-VOL	Global vector auto regression with volatility
PACF	Partial autocorrelation function
PID	Proportional-integral-derivative control
VAR	Vector autoregressive model
VECH	Column vector representation of multivariate processes
VIX	Implied volatility index over S&P 500 options

TABLE OF CONTENTS

1 INTRODUCTION

1.1 PROLOGUE

Long Memory processes are not new at all, but yet, this field has a wide set of problems to be explored and understood. Furthermore, it has deep implications for the human beings organized as a society.

Stories like the “seven years of famine” and the “seven years of prosperity”, as prophesied by Joseph to the Pharaoh (BÍBLIA, 2002), which turned out to occur later, are interesting examples of the effects of strong serial correlations, as hypothesized in the pioneer works of Hurst (1951) in hydrology, while studying the Nile River, and narrated by Beran (1994).

Then, more than three and a half millennia after such stories, strong serial correlation effects continue to generate some sort of crisis and famine, scaring the societies in form of inflation related problems. Emerging countries like Brazil, Turkey, Russia and Argentina suffer from high inflation, which impacts directly in the economic growth and increases wealth inequality (that can be interpreted as some sort of “famine”); and rich countries are suffering from the lack of economic activity, that can be observed in their respective deflation rates, also pointing out to a kind of “famine”. Both seem to exhibit strong serial autocorrelation (Long Memory).

Nonetheless, it is worth mentioning that several sets of natural phenomena are tangent to this subject. From wheat yield variation across space to turbulence are just a few to mention, as can be seen in Beran (1994). Consequently, it is clear that phenomena with long-range dependency impact human beings in a wide extent. Given that, a brief introduction to long memory processes is made in Chapter 2 of this Thesis.

But, still, despite the fact that this field has been widely studied in econometrics (which is a branch of statistics focused on economics and social science problems), discussing its respective applications and its emergence in terms of stochastic properties (as discussed in Chapter 3 of this PhD thesis), few works, if none, have aimed to describe the emergence of Long Memory Processes as a characteristic of Complex Systems, which is a relatively new interdisciplinary branch of knowledge.

Thus, the main purpose of the Chapter 3 of this Thesis is to discuss the connection between long memory processes and complexity, in order to better understand how this kind of processes emerge; and potential gains with the usage of modeling techniques that takes into account long-range dependency.

For the sake of motivation, it is worth mentioning that social sciences, in general, rest over main concepts such as homogeneity between individuals, some sort of equilibrium assumptions (or at least an implied competition or game), stability and global interactions among the agents in their respective development.

This affirmation can be easily verified while observing the construction of the most popular macroeconomic/financial models, voting models, demographic models and so on, as can be seen in Epstein (2006).

Moreover, in their respective stochastic part (when it exists), usually the researchers consider a Gaussian world, or some sort of well behaved (and known) distribution, usually employed to simplify the mathematical treatment of the problems.

Hence, we can verify that despite the huge success in the development of the social sciences as a whole along the 20th century, they were not able to predict the most recent economic crisis, social disturbances or to explain the growing inequality around the world, probably due to these mentioned simplifications.

On the other hand, despite all this knowledge about these limitations, most of the professionals (or even academics) have been using these traditional tools and continue to evaluate the problems the same old way they always did. Financial market practitioners still use Black-Scholes propositions to price their options; private bankers use Capital Asset Pricing Model for their equities; Central-Bankers use dynamic stochastic general equilibrium (DSGE) models for their Macroeconomic forecasts; and social scientists use simple voting models to predict election pools.

One of the main alternatives that have been under heavy development over the two past decades is the complexity theory, which consists of building computational-based models that include several agents or particles to interact, satisfying a previously given set of rules that governs their respective state transitions. These simulations can rely on a set of interconnected (non) linear differential (difference) equations – the system dynamics approach; or on a lattice based world, where agents/particles interact spatially and locally – cellular automata / agent based models, as can be seen in Wolfram (2002).

While the first approach has been successful in bringing more complexity to the traditional modeling, such as chaotic and turbulent behavior as consequence of the coupling of several phenomena, the second one has been catching attention of several researchers due to its ability to generate complex behavior based on very simple computational rules. For example, the European Central Bank currently conduces and supports the EURACE experiment, aiming to reproduce computationally the micro and macrostructures of the European Economy.

According to Hon (2009), this late modeling strategy has shown very interesting results and also provided new viewpoints in Physics, Chemistry, Computation and all sorts of phenomena in which small perturbations in microscopic scale generates huge disturbances in macroscopic scale.

Moreover, as can be seen in Goldreich (2012), it provides an interesting framework to understand the concepts of representation as well as concepts like randomness, knowledge, interaction, secrecy and learning, which are very important in all sorts of problems.

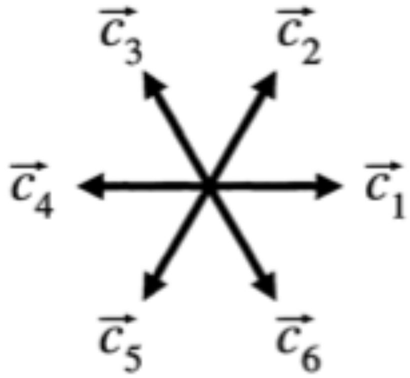
Consequently, very simple models can be implemented to study very complex behaviors that are widely found in nature. Also, it is worth mentioning that, according to Epstein (2006), "(…) even perfect knowledge of individual decision rules does not allow us to predict macroscopic structure". This enlightens the importance of techniques for mapping micro-to-macro relationships.

It is also important to notice that studying/computing the dynamics of highly non-linear systems is not trivial, at least when employing analytical tools and when combining non-linear differential equations with stochastic processes. So it is pretty much common to approximate complex non-linear systems to linearized versions, which may not exhibit the same features as the original systems. And the great advantage of using computational complex systems is the possibility to achieve these very complex behaviors based on simple mathematical rules.

For example, to better understand the concept of exploring computational complex systems in order to study complex phenomena, the present author easily implemented a turbulent flow model that exhibits high non-linear features. This model is built on top of a simple set of collision rules in a hexagonal lattice and allows one to obtain a very interesting fluid dynamics model capable of displaying turbulence, simulate a diffusion or propagate a wave, following Chopard and Droz (1998), Frisch, Hasslasche and Pomeau (1986) and Salem and Wolfram (1985).

It is somewhat surprising (and not intuitive, at a first glance) how simple individual behavior is able to generate very complex collective behavior, i.e. turbulence. In this case, given a hexagonal lattice, it makes sense considering six possible directions for each single particle, establishing a velocity scheme like in Figure 1.1.

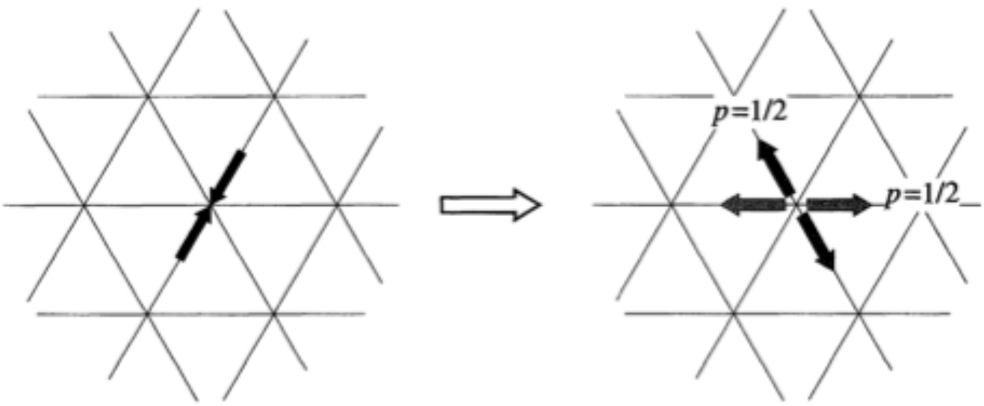
Figure 1.1 – Possible unitary velocities in the lattice-based world.



Source: Chopard and Droz (1998)

Then, whenever two particles collide against each other, they can assume the following directions with a probability of $p = 1/2$, according to Figure 1.2.

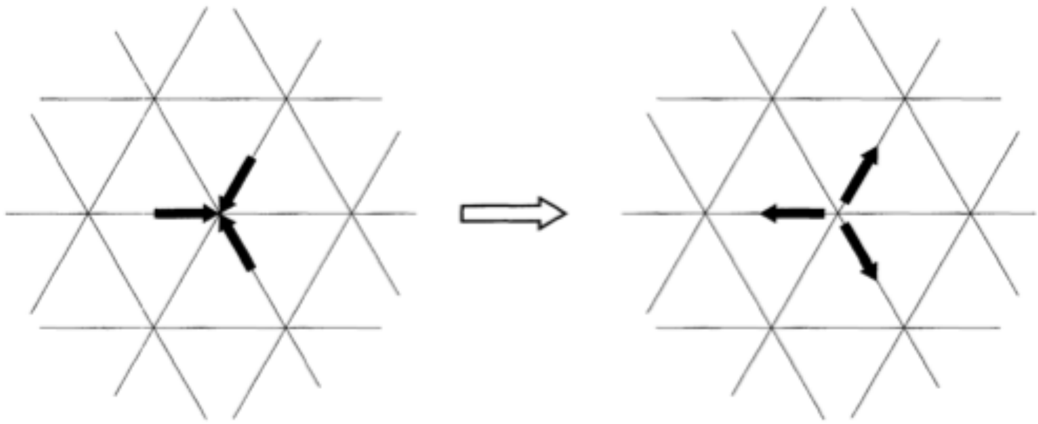
Figure 1.2 – Two-particle collision rules.



Source: Chopard and Droz (1998)

Moreover, whenever three particles collide against each other, their velocities are set according to Figure 1.3.

Figure 1.3 – Three-particle collision rules.



Source: Chopard and Droz (1998)

Hence, given all collisions possibilities, the problem is set up basically by setting each particle resting initially in a hexagonal cell, with its respective velocity randomly assuming one of the possible six directions, according to Figure 1.1, migrating to the next neighbor cell (according to its direction) in the next time step. If there are more than one particle at each hexagonal cell, it is evaluated how many particles are in there and a collision rule is applied according to Figures 1.2 and 1.3. It is also important to notice that there is a simple random behavior implemented in Figure 1.2, in order to guarantee that particles can diverge equally in all possible directions, maintaining the stability of the system in terms of physical statistics. Furthermore, the boundary conditions of this problem imply that if a particle collides against a wall, it simply suffers a 180 degrees rotation.

In order to account for the macroscopic measures, an averaging process is carried out to determine the density and the mean velocity in a specific region of the lattice.

Thus, on top of this simple model, interesting results can be obtained, as in Figures 1.4, 1.5 and 1.6, respectively, turbulence, wave propagation and diffusion processes based on a fluid model.

Figure 1.4 – Turbulent behavior.

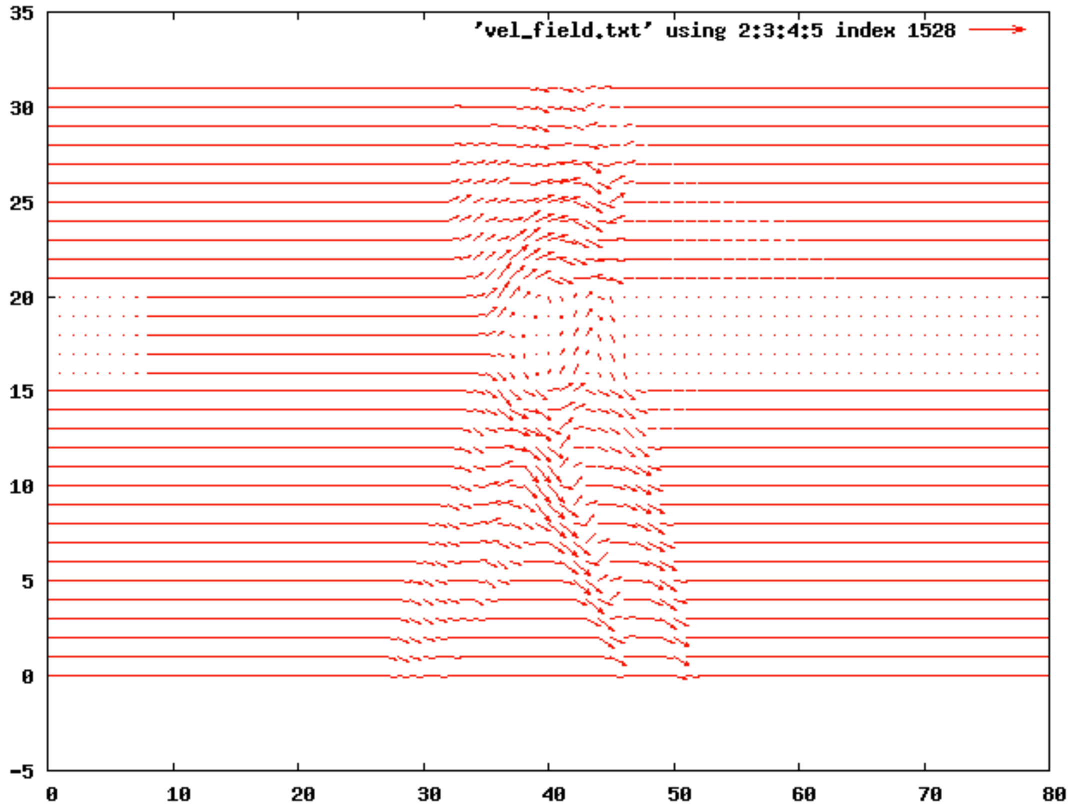


Figure 1.5 – Wave propagation in a fluid.

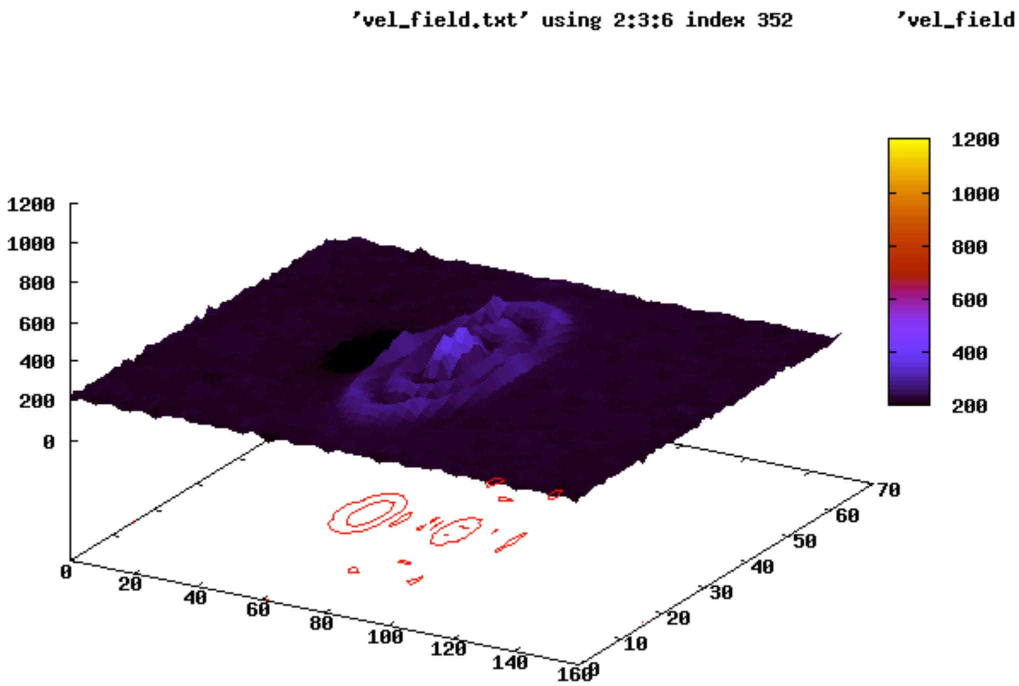
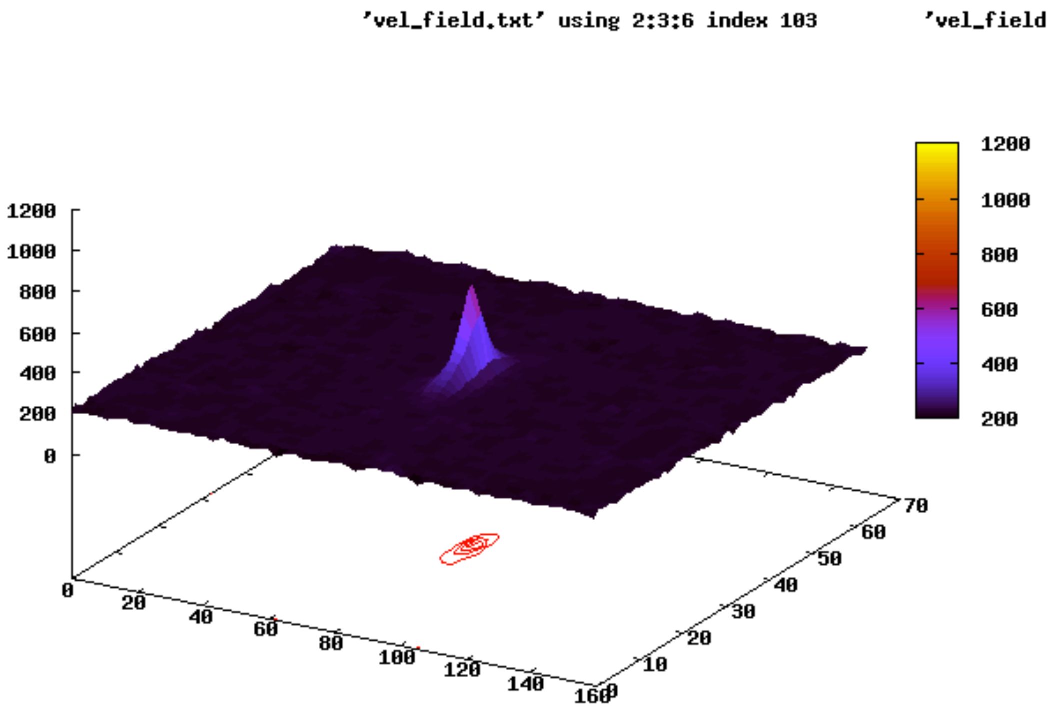


Figure 1.6 – Diffusion process in a fluid.



When analyzing Figures 1.4, 1.5 and 1.6, it is possible to notice the similarity between the plotted results of the same model against the solutions of partial differential equations such as Navier-Stokes, Wave Propagation and Diffusion Processes, respectively. So, instead of analyzing and simulating complicated models based on (partial) differential equations (Equations (1), (2) and (3)), as can be seen, the same results can be obtained by applying a simple set of computational rules.

$$\begin{cases} \rho\left(\frac{\partial u}{\partial t} + u \cdot \nabla u\right) = -\nabla p + \nabla \cdot \left(\mu(\nabla u + (\nabla u)^T) - \frac{2}{3}u(\nabla \cdot u)I\right) \\ \frac{\partial \rho}{\partial t} + \nabla \cdot (\rho u) = 0 \end{cases} \tag{1}$$

$$\frac{\partial f}{\partial t} = \alpha \nabla^2 f \tag{2}$$

$$\frac{\partial^2 f}{\partial t^2} = k^2 \nabla^2 f \tag{3}$$

Thus, complexity emerges. And instead of analyzing complex mathematical equations, one can make use of powerful but still simple computational models to evaluate complex phenomena.

Taking that into account, this set of techniques is finding a new prosperous field: the social sciences. And it seems to be an interesting way to reintroduce one of the main contributions of this PhD Thesis, which is the emergence of long-memory properties as a result of emergent complexity.

But what is an emergent phenomenon?

Epstein (2006) discusses his own *gedankenexperiment*, where people are invited to check which of the following statements they agree:

1. Emergent phenomena are undeducible.
2. Emergent phenomena are unexplainable.
3. Emergent phenomena are unpredictable.
4. Emergent phenomena are irreducible.
5. Emergent phenomena are none of the above.

And, probably, according to the author, the set of responses would be far from uniform, due to the lack of definition of what an emergent phenomenon is. On the other hand, according to Wolfram (2002), there is an interesting connection between emergent and irreducible phenomena, at least, if one follows his characterization of what is a complex cellular automaton. Emergent phenomena seems to be more common in irreducible and complex systems.

Also, according to the same author, the concept of irreducibility goes through the idea of not being able to predict the future states of a system by means of a closed formula, or a differential/difference equation. In an irreducible system it is not possible to predict its future states without passing through all intermediate states. And this is why each calculation step must be studied, because emergent phenomena may arise.

When one remembers the huge lack of prediction capabilities of economic systems outcomes, it seems economics should also be studied from this perspective. (see Epstein, 2006)

In parallel to that, it is shown along Chapter 2 that the main idea of long memory processes comes from fractional derivatives/differences operators, which enable one to model a system with long range dependency. These operators use all past information to calculate the current state of the system.

Overlapping the ideas of long-range dependency and emergent phenomena from irreducible systems, it makes sense investigating complex phenomena that are irreducible (the present state depends on all past information), to investigate whether long memory arises. If there are interesting properties that emerge with complex behavior, such as power-law, self-similarity, chaotic behavior, etc; is long-range dependency one of them?

Based on this insight, the main purpose of Chapter 3 is to shed light on the basic constituents of the mechanics behind long-memory processes, taking as the starting point Agent based Models and Cellular Automata. It is also worth mentioning that Chapter 3 is a result of several discussions in Fernandes and Garcia (2015).

Consequently, by understanding this, it would be possible to have more appropriate insights whether an observed long-memory behavior is spurious or not, based on the constituents of the system, rather than using econometric conjectures about the nature of possible time series components, and then, one of the main assumptions that usually is made in the stochastic part of the systems could be rejected – simple random-walk or white noise processes, which usually is one of the most problematic simplifications made by the most popular models (as discussed in the beginning of this section).

In this case, it is shown that heterogeneity, large deviations of equilibrium points in conjunction with the underlying motion laws and spatial complexity play an important role in the rise of long-range dependency.

When these concepts are transposed to a real society, which is composed by heterogeneous agents, bounded rationality, explicit space, local interactions and non-equilibrium dynamics, it makes sense recognizing that macroeconomic and financial markets data would be subject to the presence of long-range dependency due to the fact that its underlying elements are characterized by such features. About each one of these points and reflecting the main aspects discussed by Epstein (2006), it is easy to verify that:

- **Heterogeneity**: there are different classes in a real society, with different cultural backgrounds and different decision rules among the constituent persons in a country. Moreover, there is an intertemporal heterogeneity, where individual preferences are not fixed for the lifetime of the agent;

- **Autonomy**: there is no central or “top/down” control over individual behavior in reality, or as a central-planner as usually discussed in economic models;
- **Explicit Space**: events transpire on an explicit space, with renewable and non-renewable resources, which implies in a dynamic social network;
- **Local Interactions**: Most day-to-day transactions/interactions are carried out locally, not globally. Agents murder, steal, sell, buy, marry and reproduce themselves locally;
- **Bounded Rationality**: usually agents do not have global information or infinite computational power. Typically, they make use of simple rules based on local information, according to Rubinstein (1998).

As a practical evidence of such statements, it is worth mentioning that Fernandes and Garcia (2016) found that there are strong statistical evidences towards the presence of long-range dependency in the US economic data and it is discussed how the usage of long memory components in stochastic processes modeling can enhance the forecasting capabilities of the main techniques.

Thus, treating long-range dependency, which may be a result of the aspects described here – and discussed in Chapter 2, is a good way to study and identify stylized facts in social/economic data.

Consequently, Chapter 4 of this Thesis is much more focused on the application of mathematical/statistical modeling of such kind of processes, rather than discussing the nature of long-range dependency. Its objective is to discuss how volatility (risk aversion) across different markets exhibits long-range dependency. Furthermore, it shows how long-memory filters can be applied to separate true short-term innovations in the risk aversion from the observed levels of the time series, and finally, how this information can be used to analyze risk contagion across different countries. Nonetheless, it is worth remembering the huge impact that the 2008 crisis had on the world economy, and consequently, this is a theme that a significant number of researchers are paying attention to. So this is definitely a topic of interest.

In this sequence, Chapter 5 of this Thesis can be seen as an application of the technique developed along Chapter 4: given the risk contagion from the American Economy to the Brazilian Economy, how the 2008 Crisis affected the Brazilian GDP? To answer it, long memory filters are applied, in order to filter out what would be the “natural” volatility level and how much of the volatility is a result of risk contagion. And, based on it, calculate the impacts of uncertainty in the economic activity.

1.2 THESIS OBJECTIVES

As can be seen, this PhD Thesis is a very long and interdisciplinary journey trying to relate different fields in order to bring a more insightful view of long-range dependency and fractional differences. Starting with physical problems and going through computational experiments seems more suitable, in order to account for the intangibility of the subject. This is done along the past section and Chapters 2 and 3.

In Chapter 2, the main objective is to understand the motivations of the fractional difference operator, which is the major innovation in long memory processes. To accomplish that, an insight from mechanics is developed (using fractional calculus) and a mathematical interpretation of the fractional derivative is made. Then, the standard fractional difference operator is introduced (following the same insight), and then, the main statistical techniques to identify long memory processes are presented.

In Chapter 3, long-range dependency is discussed as an emergent property of Complex Systems and, also, what kind of constituents of such Complex Systems contribute for the rise of such characteristic. In order to achieve that, all discussed Complex Systems are implemented, employing widely known computer programming languages and further analyzed using the R Statistical Package (in order to detect the presence of long-range dependency).

Moreover, it is important to demonstrate the importance of long-memory models in the real-world problems. Hence, in Chapters 4 and 5 it is demonstrated how long-memory filters can be applied, in order to study how risk contagion effectively works across different countries and different sectors, preceded by the standard long-memory identification tools discussed in Chapter 2.

Finally, in the last chapter (Concluding Remarks), the most important remarks are made with the respective conclusions of the importance of all essays developed in this Thesis.

1.3 THESIS STRUCTURE

This Thesis is organized in the following Chapters:

- This **Introduction** Chapter;
- **Chapter 2**, which briefly presents what is a long memory process (by using fractional calculus as a starting motivation) and the standard identification techniques for these kind of processes;
- **Chapter 3**, which presents interesting connections between Complex Systems, in form of Agent Based Models and Cellular Automata, and Long Memory Processes, in order to understand how long-range dependency can be thought as an emergent property of such systems;
- **Chapter 4**, which presents the importance of considering long-range dependency in the risk contagion analysis across different economies, and how this proposed technique enhances the knowledge about risk contagion in comparison to the traditional techniques;
- **Chapter 5** discusses how long memory filters can be applied, in conjunction with risk contagion tools to understand how risk aversion crisis are spread from the financial sector to the real sectors;
- **Concluding Remarks**;
- **References**.

2 AN OVERVIEW OF FRACTIONAL OPERATORS AND IDENTIFICATION OF LONG MEMORY COMPONENTS

2.1 PROLOGUE OF THE CHAPTER

The first point to be introduced, aiming the better comprehension of the long memory concept, is the fractional differentiation operator, in continuous time processes. This concept is then extended to discrete processes, enabling a better and wider understanding of the fractional difference coefficient. Then, it is presented the main techniques used to assess long memory components in real observed data.

2.2 A MOTIVATION FROM FRACTIONAL CALCULUS

The differentiation operator of any function of *x* can be expressed as:

$$D = \frac{d}{dx} \tag{4}$$

Hence, in terms of this operator, the second order derivative of such function can be expressed as:

$$D^2 = \frac{d^2}{dx^2} \tag{5}$$

Consequently, it is possible to generalize the operator above for the successive derivatives of such function as:

$$D^k = \frac{d^k}{dx^k} \tag{6}$$

where $k \in \mathbb{Z}_+$.

The concept of fractional derivative allows one to generalize this operator, in order to allow operations of the form D^α, where $\alpha \in \mathbb{R}$ being the main motivation for that, the fact that when one applies traditional differentiation operator over a given function, it only considers the local effects of the function in the neighborhood of a given x_0 point, where the function is evaluated.

In order to generalize this fractional differentiation operator, according to Hermann (2014), by induction, it is possible to verify for polynomial functions that:

$$D^a(x^k) = \frac{d^a}{dx^a}(x^k) = x^{k-a} \cdot \frac{k!}{(k-a)!} \tag{7}$$

When one substitutes the factorial function by the Gamma function, which generalizes the factorial function for real numbers, it follows that:

$$D^a(x^k) = x^{k-a} \cdot \frac{\Gamma(k+1)}{\Gamma(k-a+1)} \tag{8}$$

So, for a generic function $f(t)$, as can be seen in Loverro (2004), the first step is considering a linear operator $J(n)$ where:

$$(Jf)(x) = \int_0^x f(t) \cdot dt \tag{9}$$

When one calculates $J^2 f$, one obtains:

$$(J^2 f)(x) = \int_0^x \left(\int_0^t f(s) \cdot ds \right) dt \tag{10}$$

Extending it arbitrally, and using the Cauchy formula for repeated integration (which can be better understood by means of the Laplace Transform), it is possible to obtain:

$$(J^n f)(x) = \frac{1}{(n-1)!} \int_0^x (x-t)^{n-1} \cdot f(t) \cdot dt \tag{11}$$

Consequently, substituting the factorial function by the Gamma function, to allow fractional numbers, one obtains:

$$(J^a f)(x) = \frac{1}{\Gamma(a)!}\int_0^x (x-t)^{a-1} \cdot f(t) \cdot dt \tag{12}$$

Having defined the integral operator J, it is possible to define a fractional derivative operator, which also accommodates non-integer orders. Other authors define other integral operators, as can be seen in Atangana and Secer (2013). The main goal of this demonstration is the construction of the fractional derivative based on the fractional integration operator. As can be seen in Vance (2014), one can build the fractional derivative in several different ways, by nesting the differential and integral operators, such as:

$$\begin{aligned} D^{1.5} &= D^2 J^{0.5} f(t) \\ D^{1.5} &= J^{0.5} D^2 f(t) \end{aligned} \tag{13}$$

Depending on the construction of the integral operator, it is possible to obtain several different formulations. So, on top of the Riemann-Liouville definition of fractional integration operator, it is possible to build the two most popular fractional derivative operators:

$$D^\gamma = D^n J^{\gamma-n} f(t) = \frac{1}{\Gamma(\gamma-n)!}\frac{d^n}{dt^n}\int_0^x (x-t)^{n-\gamma-1} \cdot f(t) \cdot dt \tag{14}$$

which is the Riemann-Liouville fractional derivative, and

$$D^\gamma = J^\gamma D^n f(t) = \frac{1}{\Gamma(\gamma-n)!}\int_0^x (x-t)^{n-\gamma-1} \cdot \frac{d^n}{dt^n} f(t) \cdot dt \tag{15}$$

which is the Caputo fractional derivative operator. Other possible operators can be constructed, depending on the integration operator definition, as can be seen in Atangana and Secer (2013). As pointed out by the same authors, the main issue with the Riemann-Liouville operator is that the fractional derivative of a constant is not equal to zero, while this is true for the Caputo derivative. Moreover, initial and boundary conditions can be included in the Caputo approach, while modeling the problems as fractional differential equations. But, on the other hand, Caputo approach requires that the function must be differentiable, while the Riemann-Liouville does not. Given that, the Caputo operator is the most popular operator in modeling problems.

Yet, from both definitions – and from other possible definitons – it is possible to see that a fractional differentiation corresponds to the convolution of the past states over time, generating a long-range dependency from the previous states over the current state of the system. Hence, it is possible to model a system that does not only depend on the current state, but from combinations between the current state and distant states in space and time, producing a “long memory effect”. Thus, these effects are not (necessarily) resultant from the existence of intelligent beings that preserve the information persistence over the space and time.

As can be seen in Chen, Petras and Xue (2009) and Machado et al. (2010), there are several applications in system dynamics and control engineering, generalizing the traditional proportional-integral-derivative (PID) control to allow fractional integration and derivatives. Other complex models such as electrical circuits with fractance, viscoelasticity problems, electrochemistry and epidemiology problems (DEBNATH, 2003; POOSEH; RODRIGUES; TORRES, 2011), can be modeled using fractional differential equations.

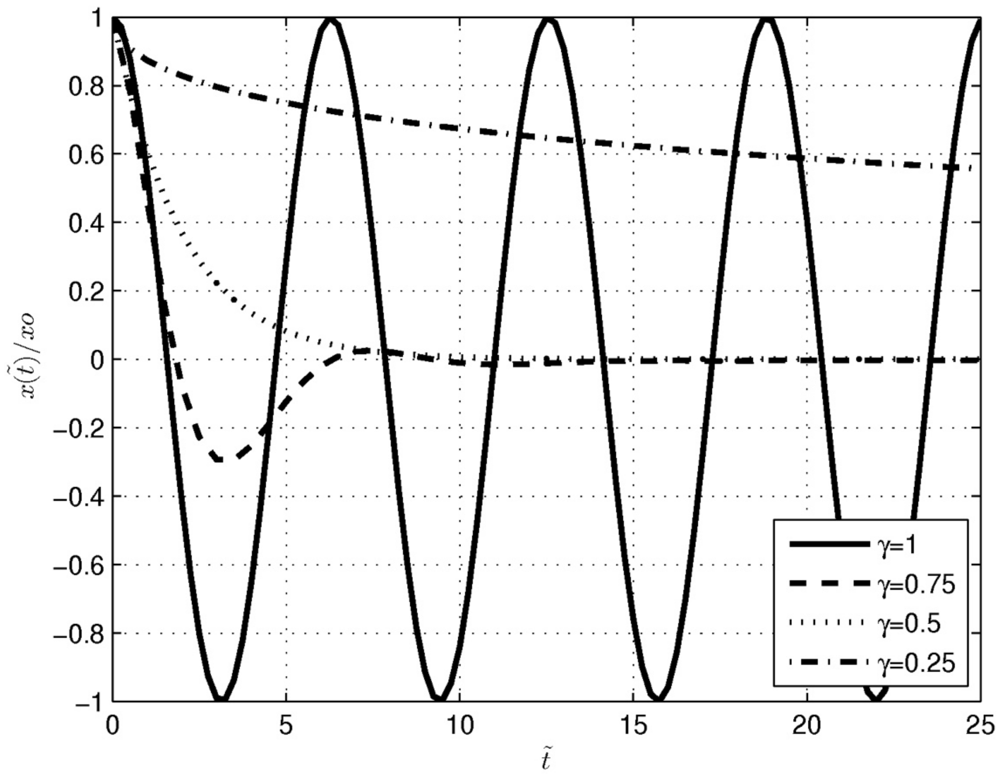
In Gómez-Aguilar et al. (2012), for example, it is possible to verify that building fractional oscillator models allows for new interesting dynamics, such as an intermediate behavior between conservative and dissipative systems. So, if one writes a harmonic oscillator in terms of a fractional differential equation (using the Caputo operator) such as:

$$\frac{d^{2\gamma}x}{dt^{2\gamma}} + 2\omega^2 x(t) = 0 \tag{16}$$

it is possible to see in Figures 2.1 and 2.2 that, depending on the values of γ, completely different behaviors occur in the system, based on the same equation. If $\gamma = 1$, the traditional periodic solution is found. If $\gamma = 0.5$, the regular damped (dissipative) behavior is found. But, depending on the other values of γ, intermediate behavior is found. Based on this equation, it is straightforward to notice that whenever fractional derivatives are used, it is possible to accommodate new dynamics, which cannot be explained by the traditional integer order calculus.

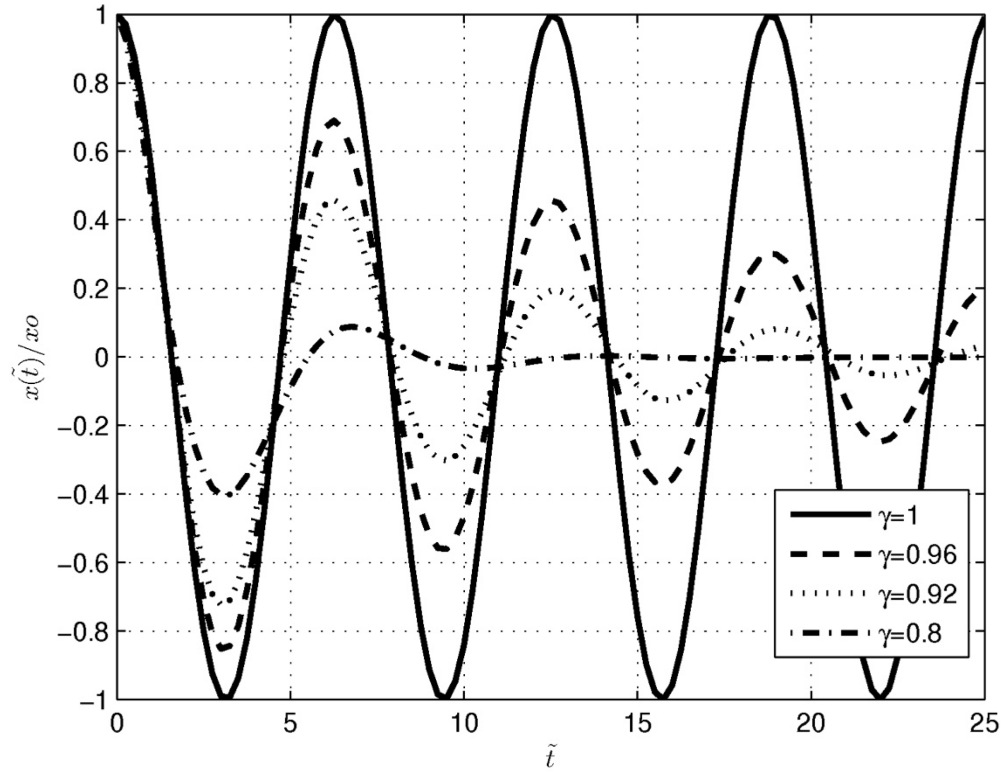
Stronger dependence from initial conditions can be imposed or can be relaxed, according to the nature of the problem. When these systems are sampled over time, resulting in discrete time systems, it is natural to expect that autocorrelation functions should not decay exponentially. It is expected that they must exhibit a slow decay, due to this strong dependence from the initial system conditions. And this is basically why it becomes so attractive modeling discrete systems with fractional integration. In addition to that, it is worth mentioning that the Fractional Brownian Motion, which exhibits long range dependency, can be derived using Fractional Derivatives, as can be seen in Lévy (1953).

Figure 2.1 – Mass-spring system with constants γ = 1, γ = 0.75, γ = 0.5 and γ = 0.25.



Source: Gómez-Aguilar et al. (2012)

Figure 2.2 – Mass-spring system with constants γ = 1, γ = 0.96, γ = 0.92 and γ = 0.8.



Source: Gómez-Aguilar *et al.* (2012)

2.3 LONG RANGE DEPENDENCY IN DISCRETE TIME AND FRACTIONAL DIFFERENCES

From the motivation of long range dependency as a result of several possible phenomena, it is worth extending the same concept to discrete time, beginning from the following definition of a standard difference operator:

$$\delta = (1 - L) \tag{17}$$

where

$$L \cdot y_k = y_{k-1} \tag{18}$$

Following the same procedures of continuous time, expanding the difference operator using Newton binomials and using the Gamma function as a generalization of the factorial operator, one obtains:

$$(1 - L)^d = \sum_{i=0}^{\infty} \frac{\Gamma(d+1)}{\Gamma(i+1) \cdot \Gamma(d-i+1)} \cdot [-L]^{(i)} \tag{19}$$

which is the definition of the fractional difference operator found in Hosking (1981).

Consequently, processes that depend on integer differences have short memory properties, as it happens in the traditional autoregressive integrated moving average (ARIMA) and vector autoregressive (VAR) models, with current states depending only on the most recent observations. On the other hand, if a fractional difference is introduced, these models become able to deal with long-range dependency.

Hence, Long Memory Processes, generally speaking, are stochastic processes that exhibit non-exponential decay in their respective autocorrelation functions, as usually observed in what can be called "short memory processes". Therefore, the main feature of this type of stochastic process is a hyperbolic decay in its respective autocorrelation function, which points out that perturbations occurred far away in time are still able to explain part of the current state of the system.

Consequently, this kind of stochastic process exhibits persistence that is neither consistent with the presence of a unit root nor with its complete absence. Hence, in order to have the necessary flexibility to deal with this apparent dilemma, it is introduced a fractional difference coefficient, which tries to accommodate stochastic processes between those with a unit root process (*d* = 1) and those with no unit root (*d* = 0).

Thus, a long memory stochastic process can be defined as:

$$x_t \cdot (1 - L)^d = \epsilon_t \tag{20}$$

where ϵ_t is a stochastic disturbance (white noise or any other short memory process), L the lag operator, x_t the contemporaneous observation of the stochastic process and d the fractional difference operator. Furthermore, if $(1 - L)^d$ is rewritten as (instead of using the Gamma Function, using a Series Expansion):

$$\begin{aligned}(1-L)^d &= \sum_{k=0}^{\infty}\binom{d}{k}\cdot(-L)^k \\ &= 1-dL-\frac{1}{2}\cdot d\cdot(1-d)\cdot L^2-\frac{1}{6}\cdot d\cdot(1-d)\cdot(2-d)\cdot L^3-\cdots\end{aligned} \tag{21}$$

one is able to verify that whenever d is an integer number, factors in the expression above vanishes according to the multiplication of all factors by zero.

Despite the study of this kind of phenomenon being a relatively old field in Mathematics and Physics, which started to be investigated right at the beginning of the 1950s, by Hurst (1951), followed by Mandelbrot and Wallis (1969) and McLeod and Hipel (1978), among others; only during the 1980s and 1990s the field grew almost exponentially, where the most important works were those written by Hosking (1981), Geweke and Porter-Hudak (1983), Robinson (1994) and Baillie (1996).

The main focus of the pioneer works was the search for empirical evidences of long memory over a different range of problems, while the focus of the posterior works was the mathematical/statistical analysis of such properties, in terms of its stochastic components and the development of proper mathematical tools and statistical tests aiming at the calculation of the fractional difference coefficient.

2.3 METHODS FOR ASSESSING LONG RANGE DEPENDENCY

In order to assess the possibility of having a long range dependency over time, in this Thesis, four statistical methods are used: Geweke and Porter-Hudak (1983) – from now on GPH, Local Whittle Estimator as in Robinson (1995), where these first two are fractional difference coefficient estimators, Modified R/S Statistic Test as in Lo (1991) and V/S Statistic Test following Giraitis et al. (2003). Moreover, in addition to these four statistical methods, a direct graphical comparison can be carried out between the partial and standard autocorrelation functions, in order to detect the presence of long memory components.

As can be seen in Kumar (2014), the GPH estimator is based on the slope of the spectral density function of the fractionally integrated time series around $\lambda \to 0$ through a simple linear regression based on the periodogram. The periodogram is defined as:

$$I(\lambda_k) = \frac{1}{2\pi T} \left| \sum_{j=1}^{N} X_j e^{ij\lambda_k} \right|^2 \tag{22}$$

where $I(\lambda_k)$ is the k^{th} periodogram point and it can be defined as the squared absolute values of the Fourier Transform of the series.

Having calculated the periodogram, the final step is to estimate the fractional difference coefficient d by estimating the following regression, using ordinary least squares:

$$\log\big(I(\lambda_k)\big) = a - d \cdot \log\left(4 \cdot \sin^2\left(\frac{\lambda_k}{2}\right)\right) + \varepsilon_k \tag{23}$$

where $\lambda_k = \frac{2\pi k}{N}$ is a constant and $\varepsilon_k \sim N(0, \sigma_k)$.

On the other hand, the Local Whittle Estimator depends also on the periodogram, as defined right above, but consists of a semiparametric estimate, which is carried out by choosing an appropriate H that minimizes an objective function $R(H)$, as in Künsch (1987):

$$R(H) = \log\left(\frac{1}{m} \cdot \sum_{k=1}^{m} \left\{\lambda_k^{(2H-1)} \cdot I(\lambda_k)\right\}\right) - \frac{(2H-1)}{m} \cdot \sum_{k=1}^{m} \log(\lambda_k) \tag{24}$$

where:

- H is any admissible value;
- $m = \begin{cases} \frac{1}{2} N \; if \; n \; is \; even \\ \frac{1}{2}(N-1) \; if \; n \; is \; odd \end{cases}$;
- $d = H - \frac{1}{2}$, being H the Hurst exponent.

Keeping that in mind, both estimators are applied to generated time series, which result of the simulation of each model explained. Hence, for $0 < d < 1$, the time series displays long memory properties.

Nonetheless, it is also important to notice that despite the fact that ARIMA processes can be seen as truncated fractionally integrated processes, the results of both tests do not suggest fractional difference coefficients where they do not exist.

About the first part of the affirmation above, if a fractionally integrated process can be written as:

$$\begin{aligned} x_t \cdot (1-L)^d &= \sum_{k=0}^{\infty} \binom{d}{k} \cdot (-L)^k \cdot x_t \\ &= \left(1 - dL - \frac{1}{2} d(1-d)L^2 - \frac{1}{6} d(1-d)(2-d) \cdot L^3 - \cdots \right) \cdot x_t \end{aligned} \tag{25}$$

obviously, simple autoregressive processes can be seen as their respective truncations of order *p*.

Moreover, in order to obtain an interesting proof in terms of the statistical tests performance with no further mathematical abstraction, a Monte Carlo based test was carried out (for the sake of this chapter), with 100 simulations for each time series specification with 2000 observations.

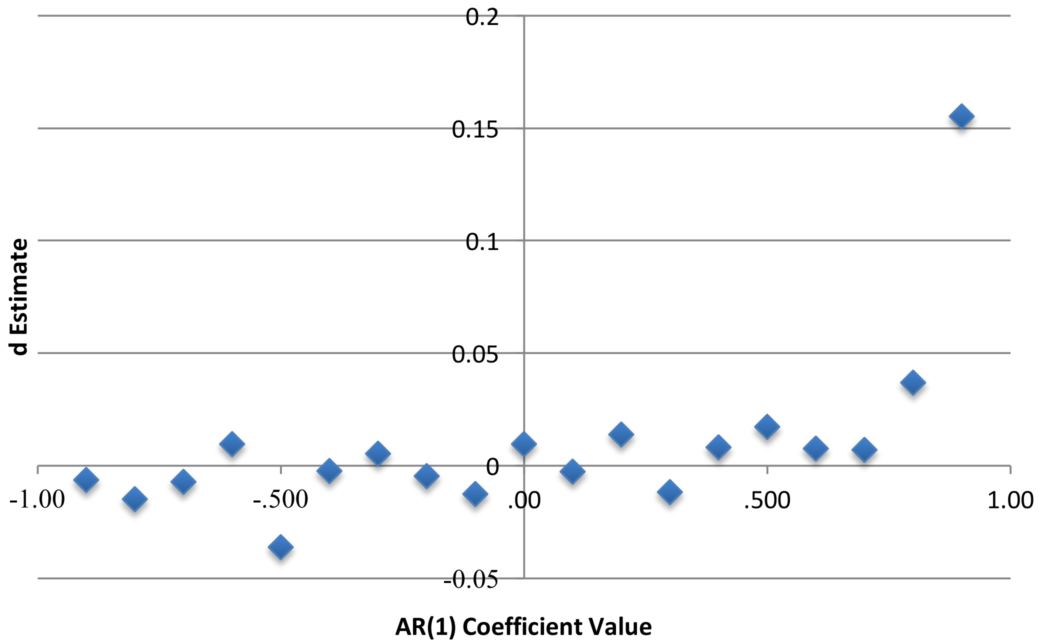
Only first order processes were tested, since none of the 2nd order autoregressive models suggested any spurious fractional difference coefficient. The same occurred for 3rd order processes and so on.

So, basically, it was simulated AR(1) processes with the first order term ranging from $\phi_1 = 0.9$ to $\phi_1 = -0.9$ (using steps equal to 0.1 – 19 different specifications, which were simulated 100 times each), and the results are shown in Figure 2.3, in terms of the suggested fractional difference coefficient. Also, it is important to keep in mind that an AR(1) process (without constant term) can be represented as a model of the form:

$$x_t = \phi_1 \cdot x_{t-1} + \varepsilon_t \tag{26}$$

If $\phi_1 = 1$, the model becomes a pure random-walk process (an ARIMA(0,1,0) model).

Figure 2.3 – Spurious estimates of *d versus* true AR(1) coefficient values.



As can be seen in Figure 2.3, only high positive values of the autoregressive coefficient can produce potential spurious results in terms of the estimate of the fractional difference coefficients. But still, these spurious effects are not significantly large. The main explanation for such occurrence is the fact that as $\phi_1 \to 1$, the model becomes more similar to a random-walk process. Hence, in terms of inference of *d*, it should approximate to 1.

Furthermore, it is also important to mention that both statistical tests produced the same results for sufficient long series, in this case, 2000 observations, and the results did not change if a constant was included.

In order to complement this first analysis, it was simulated a case where $\phi_1 = 0.95$, resulting in $d = 0.3428$. Hence, it is possible to conclude that there is an exponential decay behavior in terms of a spurious fractional difference parameter estimate versus the AR(1) coefficient value.

Hence, only first order autoregressive processes with an AR coefficient close to 1 (larger than 0.9) can lead to possible distortions in the inference of the fractional difference coefficient, which is something that must be taken into account when evaluating data generating processes. In other words, if $\phi_1 > 0.90$, potential spurious long memory can be identified.

It is also important to explain how the Modified R/S Statistic (LO, 1991) and the V/S Test work in order to detect the presence of long memory components. Aiming to explain how the Modified R/S Statistic Test works, it is useful to first discuss the original R/S Statistic developed by Hurst (1951). If one defines the following quantities:

$$Y(n) = \sum_{i=1}^{n} X_i , \qquad n \geq 1 \tag{27}$$

$$S^2(n) = n^{-1} \cdot \sum_{i=1}^{n} \left(X_i - n^{-1} \cdot Y(n)\right)^2 , \qquad n \geq 1 \tag{28}$$

then, the original R/S Statistic can be defined as:

$$\frac{R}{S}(n) = \frac{1}{S(n)} \cdot \left[\max_{0 \leq t \leq n} \left(Y(t) - \frac{t}{n} \cdot Y(n)\right) - \min_{0 \leq t \leq n} \left(Y(t) - \frac{t}{n} \cdot Y(n)\right)\right], \qquad n \geq 1 \tag{29}$$

where

$$E\left[\frac{R}{S}(n)\right] \sim c_1 \cdot n^H, \qquad n \to \infty \tag{30}$$

with c_1 being a positive constant number that does not depend on n and H denoting the Hurst Exponent within the range $]0.5, 1[$, if there is a long memory component. Consequently, as can be seen in Teveroversusky et al. (1999):

> Classical R/S analysis aims at inferring from an empirical record the value of the Hurst parameter in [last equation] for the long-range dependent process that presumably generated the record at hand.

Having this classical procedure in mind, Lo (1991) has identified weaknesses in this standard procedure, where short-memory presence may generate distortions in the detection of the presence of long-range dependency. Thus, in order to take short memory components into account, the author proposed to modify the original R/S Statistic Test, in order to accommodate autocovariance components instead of the simple variance. Furthermore, instead of considering multiple lags, the test focuses only on lag $n = N$, the length of the series.

Thus, remembering that $S^2(N)$ is the sample variance of the series, Lo (1991) proposed the following measure for *S*, which now depends on a parameter q:

$$S_q(N) = \sqrt{S^2(N) + 2\sum_{j=1}^{q}(\omega_j(q) \cdot \gamma_j)} \tag{31}$$

where γ_j is the sample autocovariance at lag j, $\omega_j(q)$ is defined according to:

$$\omega_j(q) = 1 - \frac{j}{q+1} \tag{32}$$

and q is the truncation lag. As can be seen in Teveroversusky et al. (1999):

> the strong dependence between the outcome of the test (based on the test-statistic) and the choice of the truncation lag *q*, with a definite bias toward accepting the null hypothesis of no long-range dependence for large *q*'s.

Hence, the choice of q plays an important role in the sensitivity of the Modified R/S Test. On the other hand, the V/S Statistic Test, developed by Girailis et al. (2003), is less sensitive to the choice of the truncation lag. This test consists of the following statistics:

$$M_N = \frac{1}{S_q^2(N) \cdot N^2}\left[\sum_{k=1}^{N}\left(\sum_{j=1}^{k}(X_j - \overline{X_N})\right)^2 - \frac{1}{N}\left(\sum_{k=1}^{N}\sum_{j=1}^{k}(X_j - \overline{X_N})\right)^2\right] \tag{33}$$

where $S_q(N)$ is the same quantity as defined in Lo's (1991) test. But still, *q* should be defined. If M_N is greater than a critical value, the test suggests the presence of long memory.

In order to establish the truncation lag q for both tests, the present author follow the approach proposed by Lima and Xiao (2004), instead of simply using the formula developed by Andrews (1991), which is the more traditional approach.

The formula developed by Andrews (1991) is given by:

$$q = \left[\left(\frac{3N}{2} \right)^{\frac{1}{3}} \cdot \left(\frac{2\hat{\rho}}{1 - \hat{\rho}^2} \right)^{\frac{2}{3}} \right] \tag{34}$$

where $\hat{\rho}$ is an estimate of the first order autoregressive coefficient.

Lima and Xiao (2004) proposed a different approach, by combining information available in the dataset, as in Andrews (1991) and a correction of bandwidth lag selection based only in the sample size. The main reason for that can be seen in Xiao (2002), where the author demonstrated that Andrews' (1991) formula not only captures the short-range dependence, but can also capture the long-range dependence. Hence, they proposed the following formula (LIMA; XIAO, 2004):

$$q_1 = \left[\left(\frac{3N}{2} \right)^{\frac{1}{3}} \cdot \left(\frac{2\hat{\rho}}{1 - \hat{\rho}^2} \right)^{\frac{2}{3}} \right] \tag{35}$$

$$q_2 = \left[4 \cdot \left(\frac{N}{100} \right)^{\frac{1}{3}} \right] \tag{36}$$

$$q^* = \min(q_1, q_2) \tag{37}$$

where q^* is the optimal choice based on empirical tests.

Consequently, using these four tests and evaluating the Autocorrelation function, it is expected that it is possible to detect the presence of long-range dependency in a robust way, which is, at least, more coherent with empirical simulations.

3 COMPLEX SYSTEMS, REAL MARKETS AND LONG MEMORY: AN INTERESTING CONNECTION

3.1 PROLOGUE OF THE CHAPTER

The aim of this chapter is to establish an interesting connection between the behavior of economic agents and the long memory features, that generally occur in a wide set of time series found in economic/financial problems. It is shown that heterogeneity between agents, large deviations from the equilibrium points (in conjunction with the laws of motion) and spatial complexity are very important in the rise of long memory features, by means of extensive usage of computational multi-agent based models, stochastic analysis and Monte Carlo simulations. Despite the fact that heterogeneity is a widely known characteristic that affects the rise of long memory, the other two factors are not.

Moreover, when a long memory filter is applied over time series with such properties, interesting information can be retrieved.

3.2 A REVIEW OF PREVIOUS CONTRIBUTIONS AND PERSPECTIVE OF THE PRESENT CONTRIBUTION

It is known that most of the papers related to Long Range Dependency published since 1950s focused on: empirical evidences of long range dependency in a wide set of different fields, from Physics to Finance; on propositions that enhance the computational calculations of models which exhibit these stochastic properties (long memory) and on the development of the theoretical mathematical toolset used to build and analyze these models.

However, to the best of the present author’s knowledge, since there are no other works that discuss the aspects of such stochastic property origins by establishing such link with Dynamic Complex Systems, agents’ individual behavior, aggregate behavior and other related fields, integrated with the possibility of doing computational simulations that are able to generate such behavior.

Along the 1980s and 1990s, a set of econometric works was developed, focusing basically on the origins of long memory processes, as a result of the sum of cross-sectional short-memory processes, as can be seen in Granger (1980) and

Zaffaroni (2004); as result of the sum of continuous binary series as in Taqqu, Willinger and Sherman (1997), where the governing cumulative distribution function of the binary states are governed by power-laws; and structural changes/regime switching processes, as discussed in Stock and Watson (1996). Therefore, in order to complement these previous existing works, the author aim to analyze processes with different stochastic properties from those discussed above, with a special focus on the characteristics of the agents' behavior.

Hence, the main idea of this chapter is to discuss possible origins of such phenomena by running computational simulations of the interactions between single individuals (called agents), which produce local and global interactions that are studied in terms of its stochastic properties. Moreover, this chapter aims to show and discuss that long memory properties are not necessarily resultant from long memory behavior of individual agents nor from social/economic frictions. Furthermore, it extends the knowledge of agent behavior complexity (initially restricted to heterogeneity), by adding two other important factors: spatial complexity and motion laws in conjunction with large deviations from equilibria, which seems to extend the previous knowledge about the rise of long memory properties.

The importance of the present work is due to the fact that the incorrect specification of stochastic processes can provide misleading conclusions. Its specification affects the description of the autocorrelation structure of a wide range of problems, such as asset pricing, macroeconomic modeling and other time series phenomena. The misspecification of such features may induce very different results in long term, affecting the way that optimal policy making may be conducted, since these effects last longer than short memory.

To accomplish this goal, in such computational models, the agents must only explicitly have short memory relationships with their respective past states. Thus, it should be possible to show that long memory properties arise not because the agents may have a memory unit which guides them in their respective actions (behavior), as one may think in terms of traders pricing an asset according to the present and their perception of a fair price based on their long experience; but as a result of the aggregate behavior of them, as a consequence of the complexity emergence, pointing back to the seminal works of Mandelbrot and Wallis (1969) and establishing an interesting link with the growing field of Complexity, as in Wolfram (2002), Monteiro (2011, 2014), among others.

Consequently, the behavior of agents in such systems would be somewhat affected by disturbances occurred in a far past, but not explicitly derived of individual long memory behavior, which affects directly the development of optimal control policies for such kind of systems.

Keeping that in mind, three different computational models are presented and simulated in this chapter, showing that long range dependency may simply arise from the interactions between the agents, establishing what can be called “long memory emergence”.

On the other hand, none of these models were developed for this work. Their respective authors separately made them for specific purposes and that is why the present author have decided for such strategy (of picking models made by third parties). Instead of building models (which usually takes a considerable amount of time to make them work properly) that might contain biases in terms of finding such long memory properties – as a consequence of the present chapter idea – they were chosen, simulated (in their respective platforms) and analyzed using the R Statistical Package.

In the three next sections, each one of these three models are presented, simulated and discussed in terms of the stochastic properties found over the results obtained, while pointing out possible reasons for such results, as a consequence of agents heterogeneity, local interactions and spatial complexity. After that, it is presented a final section containing a brief conclusion of evidences towards the emergence of long-range dependency, as a result of other kind of interactions beyond explicit long memory behavior of individuals and how long memory filters can be applied to filter relevant information.

3.3 A MULTI-AGENT PREDICTION MARKET BASED ON BOOLEAN NETWORKS

The main idea of this first model is to simulate the dynamics behind the interactions between individuals, an external source of information and a market clearer, which aggregates the individual beliefs, in order to compose public opinion (or probability of outcome) of events. Hence, as Jumadinova, Matache and Dasgupta (2011) point out:

> Prediction markets have been shown to be a useful tool in forecasting the outcome of future events by aggregating public opinion about the events' outcome. Previous research on prediction markets has mostly analyzed the prediction markets by building complex analytical models. In this paper, we posit that simpler yet powerful Boolean rules can be used to adequately describe the operations of a prediction market.

The basic structure of this model is composed of individual agents, which update their respective beliefs according to a Boolean based rule, where they assume a binary belief state: 1 when they believe that a specific event will happen; 0 when they believe that a specific event will not happen. The factors that are weighted in order to assume one or another state are:

- The individual past belief state, given by $S(t-1)$;
- Overall average of the individuals past belief state (condensed into "prices" between 0 and 1 – continuous variable), given by $P(t-1)$;
- External Information, represented by a Bernoulli Random Variable, given by $B(t)$, with a probability q of obtaining a 1 and $(1-q)$ of obtaining a 0.

The overall average of individuals beliefs are condensed into prices (aggregated probabilities) according to the following mathematical expression:

$$P(t) = \sum_{i=1}^{N} \frac{S_i(t)}{N} \tag{38}$$

Furthermore, there is a mathematical function that updates the belief state according to the following expression:

$$\begin{cases} \text{if } [w_1 \cdot S_i(t) + w_2 \cdot B(t) + w_3 \cdot P(t) > z_i] \rightarrow \ S_i(t+1) = 1 \\ \text{else } S_i(t+1) = 0 \end{cases} \tag{39}$$

where

$$\sum_{i=1}^{3} w_i = 1 \tag{40}$$

and $w_i \in [0,1]$. Moreover, z_i is the individual bias, generated randomly for each agent.

In the implementation of this code, for the purpose of this chapter, in order to generate heterogeneity between the individuals, it was imposed that $z_i \sim N(0.5, 0.1)$, fixed at the first simulation step.

For simplicity, in this chapter it was adopted a probability $q = 0.5$, $w_1, w_2 = 0.3$ and $w_3 = 0.4$, in order to avoid any apparent bias in the generated time series and any very strong autocorrelation over individual past states.

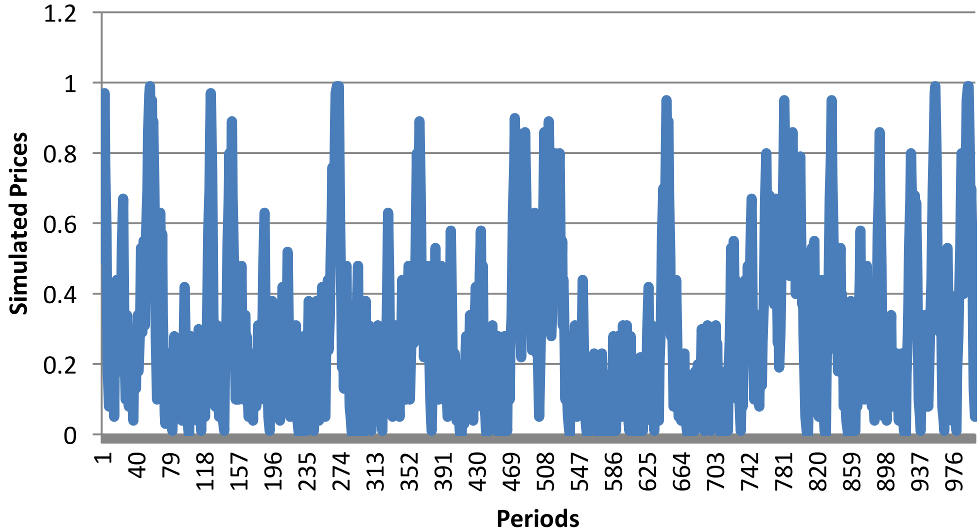
Thus, basically, this set of rules represents a simple Boolean Network, where all agents are interconnected (which simulates a situation of synchronous information and perfect information symmetry), simplified by the existence of an external “Market Maker” agent, which condensates all agents (nodes) beliefs into something that resembles a price. On the other hand, the state of each agent does not depend on any spatial position, since they are all virtually connected and information flows instantaneously, resembling individuals interconnected by a mechanism such as Internet.

Mathematically speaking, it turns out that this network configuration implies in a linear feedback relationship between the agents behavior and their respective collective behavior, which can amplify the system oscillations or stabilize them, depending on the parameter value (w_3).

For the purpose of this chapter, this model was implemented using the software Insight Maker and 100 simulations were carried out, where each one generated price time series that encompassed 2000 ticks. An example of a resulting series is shown in Figure 3.1.

Thus, it was calculated the Künsch (1987) and GPH (GEWEKE; PORTER-HUDAK, 1983) estimates of the fractional difference coefficients over these price series, in order to test the presence of long memory components.

Figure 3.1 – Simulated prices over time.



The average Künsch (1987) estimate for the fractional difference coefficient obtained by the author for the 100 simulations was 0.4869814, while the average GPH (1983) estimate was 0.1457286. If taken into account the fact that the past state is Boolean and the autoregressive part of the function is still weak (less than 0.9), both results provide strong evidences towards the presence of long memory components in this kind of process – they must be statistically different from zero, which can be verified by analyzing the distribution of the estimated parameters. The distribution of the fractional difference estimates (GPH) is described in Figure 3.2.

According to Figure 3.2, it is clear that this process exhibits long-memory properties, avoiding any spurious result from a single simulation, as it relies on a Bernoulli random variable to generate part of the stochastic fluctuations.

In Figure 3.3 it is shown the distribution of the fractional difference estimates according to Künsch (1987).

Figure 3.2 – Histogram of the parameter *d* by the GPH method.

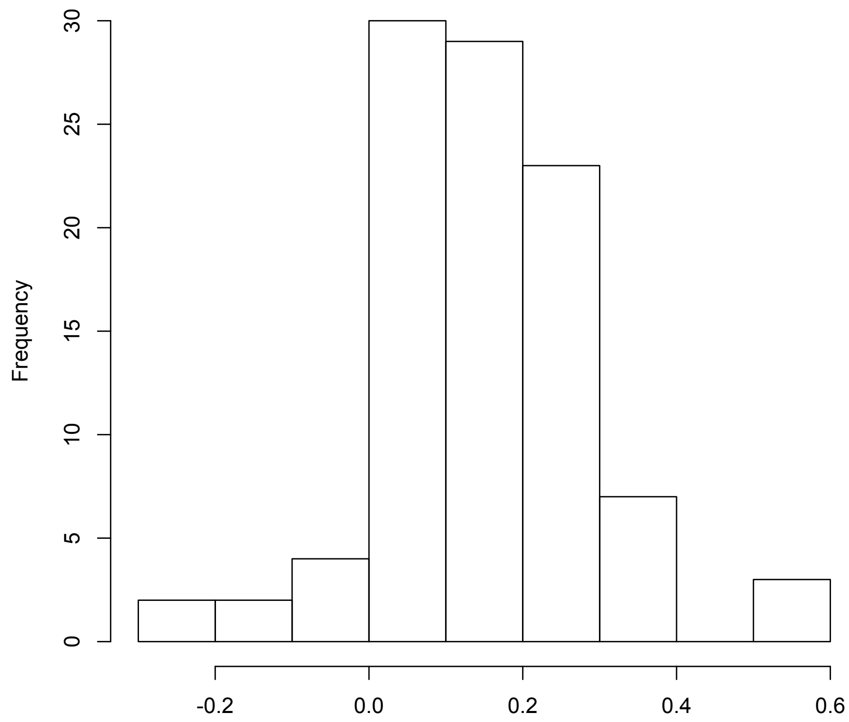
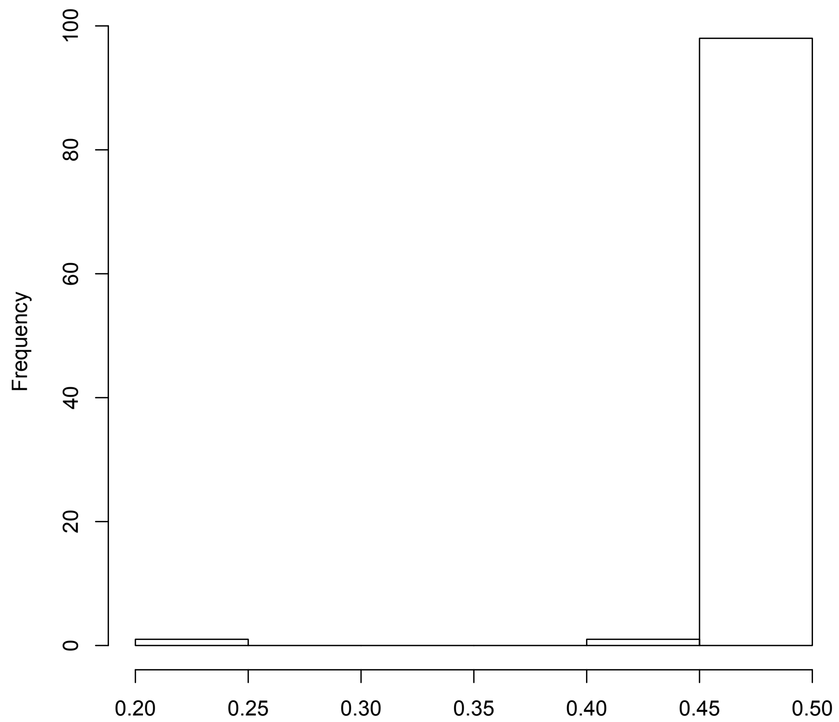


Figure 3.3 – Histogram of the parameter *d* according to Künsch (1987).

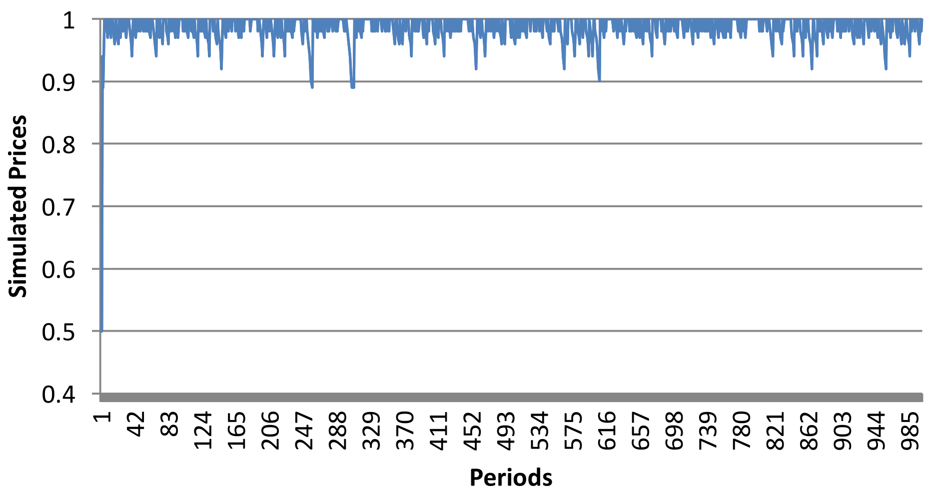


According to Figure 3.3, it is important to notice that the shape of the distribution is completely different from the GPH (1983) estimates. Still, according to these results, it suggests the presence of long range dependency.

Nonetheless, in order to reduce the heterogeneity between the simulated individuals, this experiment was repeated using $z_i \sim \mathrm{N}(0.5, 0.05)$, fixed at the first simulation step.

When the heterogeneity is reduced, the system behaves completely different. It rapidly converges towards a low-level price or a high-level price, resembling a white noise around these levels, as shown in Figure 3.4.

Figure 3.4 – Simulated prices over time.



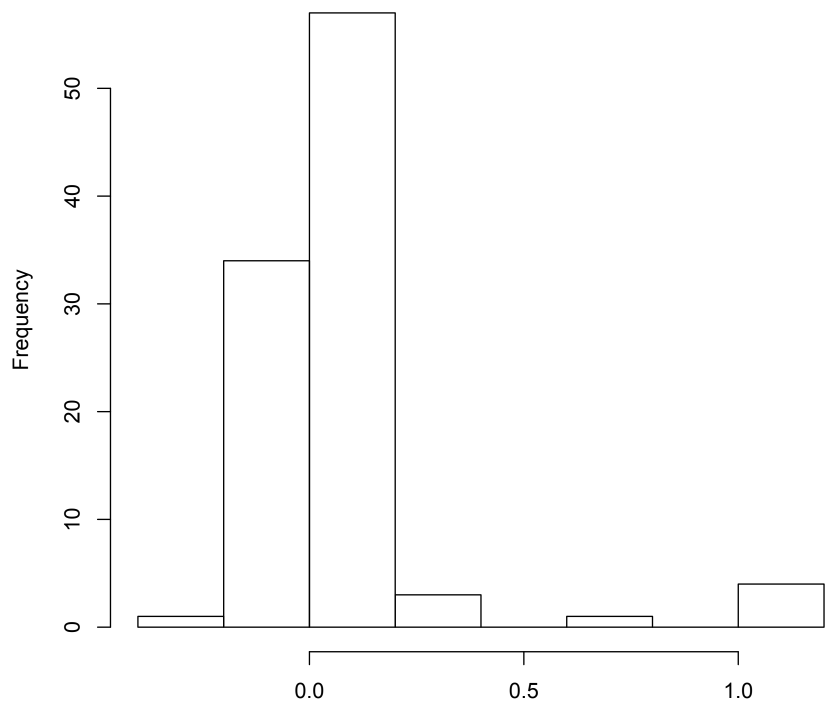
With a probability $q = 0.5$ of having a positive or negative external information, according to the mentioned Bernoulli variable, the system converges rapidly towards a high-level or a low-level price. In the case of Figure 3.4, it converged towards a high-level price.

Having estimated the fractional difference coefficients, it was obtained an average Künsch (1987) estimate of 0.3211652, while the average GPH (Geweke; Porter-Hudak, 1983) estimate was 0.06440041.

Again, if taken into account the fact that the past state is Boolean and the autoregressive part of the function is still weak (less than 0.9), the GPH results provide weak evidences towards the presence of long memory components in this kind of process – in this case, it suggests a White Noise Process. On the other hand, the Künsch (1987) estimates provide evidences towards the presence of these long range dependencies.

The distribution of the fractional difference estimates (GPH) is described in Figure 3.5.

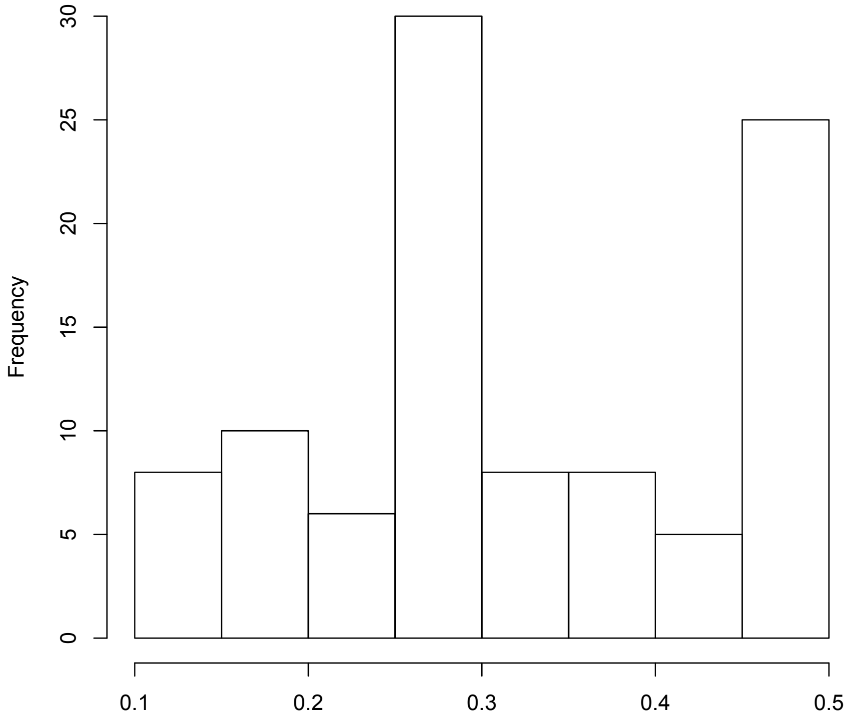
Figure 3.5 – Histogram of the parameter *d*.



In contrast to the previous distribution, in Figure 3.5 is clear that this process does not exhibit long memory properties. In Figure 3.6 it is shown the distribution of the fractional difference coefficient estimates according to Künsch (1987).

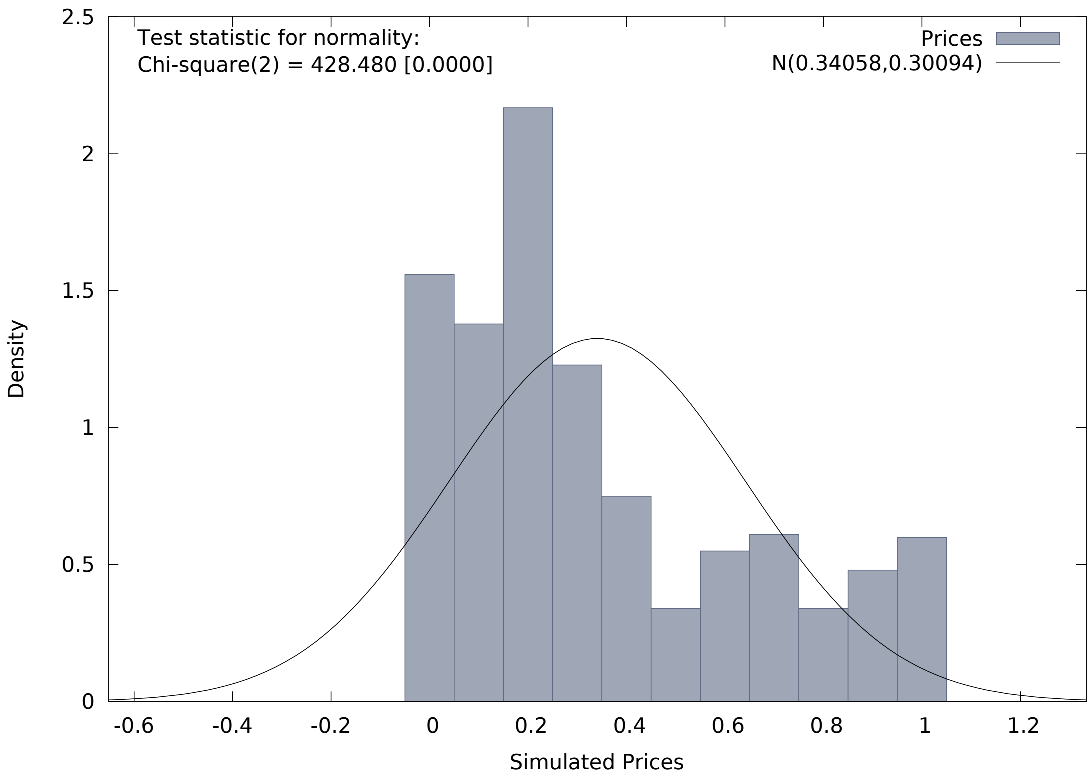
When analyzing Figure 3.6 it is important to notice that its shape is completely different from those obtained in the previous cases. The distribution suggests the presence of a long range dependency in the analyzed stochastic process.

Figure 3.6 – Histogram of the parameter *d*.



Another interesting fact is that the shape of the simulated price distribution is completely different in both cases, as seen in Figure 3.7.

Figure 3.7 – Histogram of simulated prices.



In Figure 3.7, that synthesizes the distribution of the prices in the first case, it can be seen that the prices are very far from a normal distribution, with considerable high *fat fails*, asymmetry and so on, resembling a power law-like distribution.

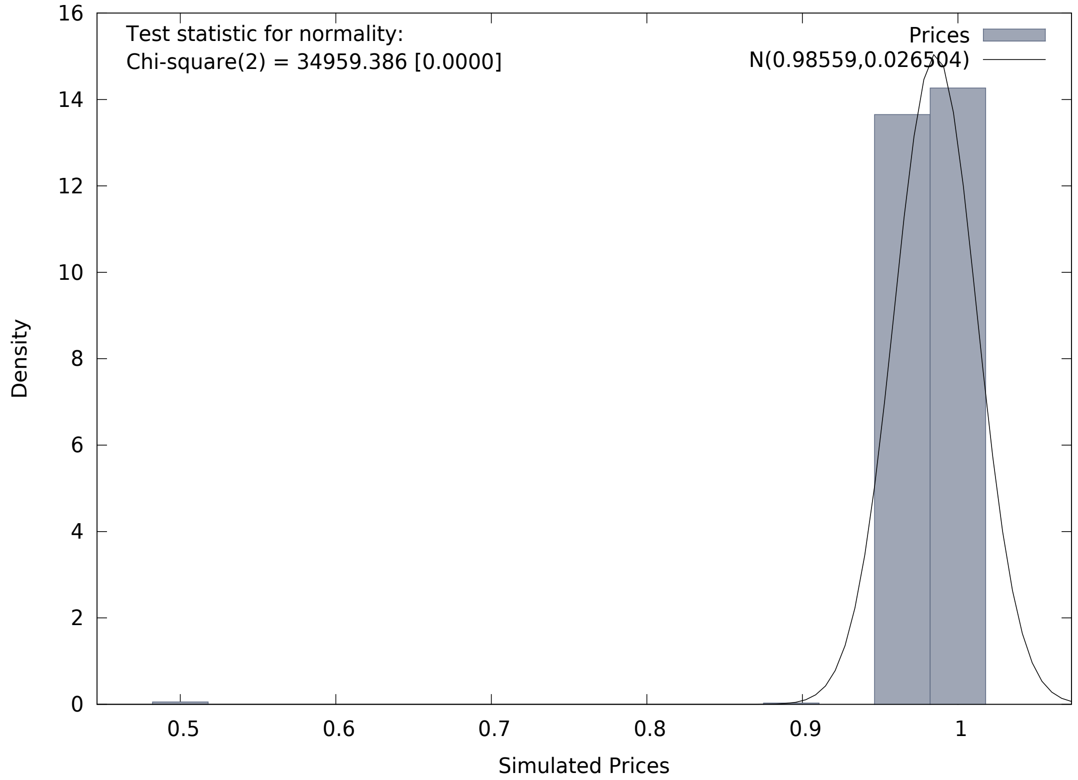
Except for the percentile between 0.2 and 0.3, which has a very strong peak, there is a typical exponential decay, which is one of the most important features in a power law distribution.

Furthermore, the characterization of such behavior is very important, due to the fact that this is one of the most notorious aspects present in self-similar processes, which are naturally one of the emergent properties of complex systems.

Thus, as it is widely known that Power Laws are an important characteristic of self-similar processes, and on the other hand, knowing that long-range dependency can arise from such processes, this is one more interesting finding towards the obtainment of empirical evidences of the presence of such property.

This same procedure was applied to other set of experiments made in the complementary case, as shown in Figure 3.8.

Figure 3.8 – Histogram of simulated prices.



In Figure 3.8, which synthesizes the distribution of the prices in the second case, it can be seen that the prices are not far away from a normal distribution, despite the fact that the statistical test for normality rejects a normal distribution.

This distribution is far more symmetric than the previously presented and does not exhibit huge fat tails, not suggesting a Power Law like distribution, which is an interesting evidence in terms of the absence of long range dependency.

In terms of the Modified R/S Statistic Test and V/S Statistic Test, with 5% of significance, the results show that 50% and 60% (respectively) of the simulated samples have long-range dependency in the heterogeneous samples, while 95% and 82% (respectively) of the simulated samples do not have long-range dependency in the homogenous samples.

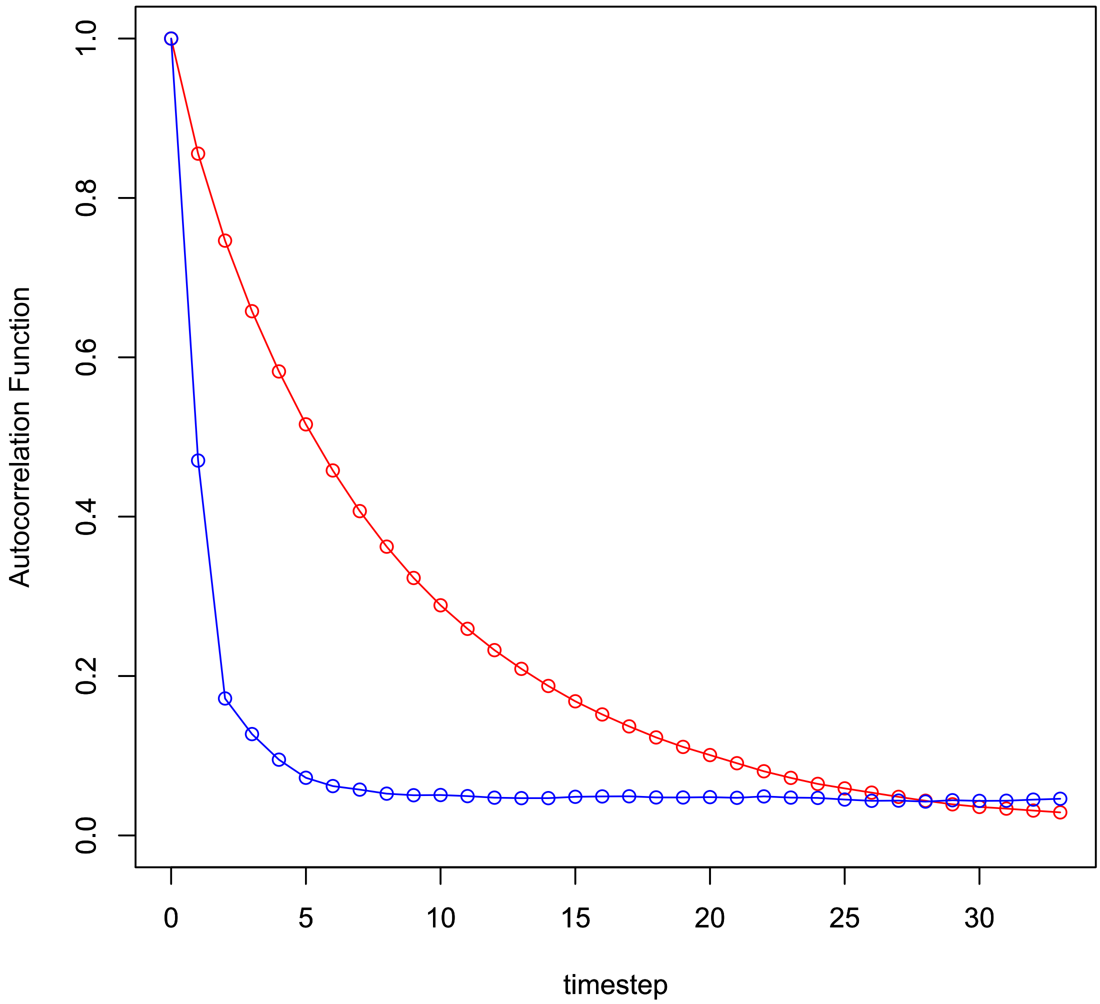
Finally, as these two statistical tests are very sensitive to the bandwidth choice of q, as discussed in Chapter 2, it was calculated the mean of autocorrelation functions of all simulations, in order to analyze the presence of slow hyperbolical decay (long memory) versus fast exponential decay (short memory).

In Figure 3.9 it is possible to see that there is a slow hyperbolic decay in the red line and a fast exponential decay in the blue line, which represents respectively the heterogeneous agents process and the homogeneous agents process, evidencing the rise of long memory, despite possible doubts according the Modified R/S and V/S Statistic Tests.

Hence, what can be said in terms of agent-based models is that the introduction of heterogeneity among the agents is an important factor in terms of the existence of long memory properties in any process. When heterogeneity is considerably reduced, it seems these properties vanish, at least in this model.

Furthermore, it is worth mentioning that the implied process in the agents discussed in the present section is rather different from those discussed in the traditional literature, as in Granger (1980) and Zaffaroni (2004).

Figure 3.9 – Plot of autocorrelation functions.



The idea of heterogeneity present in the sum of cross-sectional processes in these original works which results into long memory properties, are a consequence of the heterogeneity present in the autoregressive part of each individual agent. These works discuss the rise of long memory properties in the aggregation of stochastic processes of the following form:

$$x_{i_t} = \alpha_i \cdot x_{i_{t-1}} + \varepsilon_{i_t} \tag{41}$$

where α_i is the autoregressive parameter distributed according to a specific distribution (so this is the heterogeneity source in the implied process), ε_{i_t} are normally distributed disturbances with $\varepsilon_{i_t} \perp \varepsilon_{j_t}$, and x_{i_t} is the *i*-th individual real-valued stochastic variable that may be aggregated. In an adaptive learning framework, this can be thought as the way each individual makes their respective forecast about the system next state.

On the other hand, in the underlying process here discussed, x_{i_t} is a binary stochastic process (which has a completely different nature); all autoregressive parts are kept constant across the different agents and a bias factor is introduced (each agent with its own bias). Nonetheless, ε_{i_t} are distributed according to a Bernoulli distribution, and there is also a linear feedback relationship between the collective behavior of the agents and its individual behavior, which is fed into their individual process according to the same parameter weights.

When all these factors are put together and compared against the traditional literature, an important conclusion can be made. If all agents forecast exactly in the same way (which is an explicit assumption in this model), the way they perceive these same forecasts generate such long memory behavior, which is something not yet discussed in the existing literature.

In addition to the mentioned points above, it is worth mentioning that Jumadinova, Matache and Dasgupta (2011) have chosen a Bernoulli distributed variable for the external information, given the fact that it is very easy to obtain a map within the range $[0; 1]$ that models the behavior of the mean field of the proposed Boolean Network.

It is straightforward to derive a mean-field $p_r(t)$ that denotes the mean probability that a specific event r will occur at instant t:

$$\begin{aligned} p_r(t+1) &= P(r(t+1)=1) \\ &= P(r(t+1)=1 \mid r(t)=0)\cdot\big(1-p_r(t)\big) \\ &\quad + P(r(t+1)=1 \mid r(t)=1)\cdot p_r(t) \end{aligned} \tag{42}$$

and according to the rules implied in each agent (node), it follows that (for a generic probability distribution of the variable B):

$$P(r(t+1)=1 \mid r(t)=0) = P(w_1\cdot p_r(t) + w_3\cdot B > z) \tag{43}$$

$$P(r(t+1)=1 \mid r(t)=1) = P(w_1\cdot p_r(t) + w_3\cdot B + w_2 > z) \tag{44}$$

As can be seen in the original work, it is possible to derive the following map, when B follows a Bernoulli process:

$$p_r(t+1) = \begin{cases} 1, \text{if } p_r(t) > \dfrac{z}{w_1} \\ q \cdot \left(1 - p_r(t)\right) + p_r(t), \text{if } \max\left\{\dfrac{z - w_2}{w_1}; \dfrac{z - w_3}{w_1}\right\} < p_r(t) \leq \dfrac{z}{w_1} \\ p_r(t), \text{if } \dfrac{z - w_2}{w_1} < p_r(t) \leq \dfrac{z - w_3}{w_1} \\ q, \text{if } \dfrac{z - w_3}{w_1} < p_r(t) \leq \dfrac{z - w_2}{w_1} \\ q \cdot p_r(t), \text{if } \dfrac{z - w_2 - w_3}{w_1} < p_r(t) \leq \min\left\{\dfrac{z - w_2}{w_1}; \dfrac{z - w_3}{w_1}\right\} \\ 0, \text{if } p_r(t) \leq \dfrac{z - w_2 - w_3}{w_1} \end{cases} \tag{45}$$

According to Jumadinova, Matache and Dasgupta (2011), this map does not show any complex behavior, which is possible to see when analyzing its fixed points – that are obtained by evaluating the conditions where $p_r(t+1) = p_r(t)$. However, if z is different for every agent (in the case where heterogeneity is introduced), the derived map is not valid anymore.

Hence, rather than trying to derive (if possible) a map for such condition, it is already possible to check in the previous simulations that complexity emerges, but still, depending on the degree of heterogeneity introduced in the system, the system floats around the original fixed points zero and one present in the derived map above. When a large heterogeneity is introduced over the parameter z, such behavior completely disappears, giving rise to a full complex behavior over $p_r(t)$.

Consequently, it is possible to show that, depending on the degree of heterogeneity introduced, the system can behave deterministically (as in the derived map), probabilistically, around two main fixed points, and a very complex behavior with long memory components, as in the third case. Despite the other traditional generalizations, as can be seen in Granger (1980) and Zaffaroni (2004), additional simulations must be conducted for further generalization of these conclusions, which seems to be a new interesting path.

3.4 CELLULAR AUTOMATA AS DYNAMIC STOCHASTIC SYSTEMS

Cellular Automata are one of the simplest dynamic systems capable of showing complex behavior, which, based on a reduced set of rules, can be investigated using the standard analytical framework.

Moreover, Cellular Automata have been used to model a wide range of different problems, where it is not trivial to derive mean-field equations based on the establishment of (partial) differential equations or when it is hard to model local and spatial-dependent interactions. A few examples of what kind of modeling this technique encompasses are:

- spatial-dependent predator-prey interactions as in Vilcarromero, Jafelice and Barros (2010);
- epidemic dynamics as in White, Rey and Sánchez (2007) and Pfeifer et al. (2008);
- urban and populational growth as in Almeida and Gleriani (2011) and Mavroudi (2007);
- social change (NOWAK; LEWENSTEIN, 1991);
- consumer behavior (ROUHAUD, 2000);
- city traffic (ROSENBLUETH; GERSHENSON, 2011).

In order to advance in this subject, it is worth explaining what is a Cellular Automaton first. According to Wolfram (2002):

> A cellular automaton is a collection of "colored" cells on a grid of specified shape that evolves through a number of discrete time steps according to a set of rules based on the states of neighboring cells. The rules are then applied iteratively for as many time steps as desired.

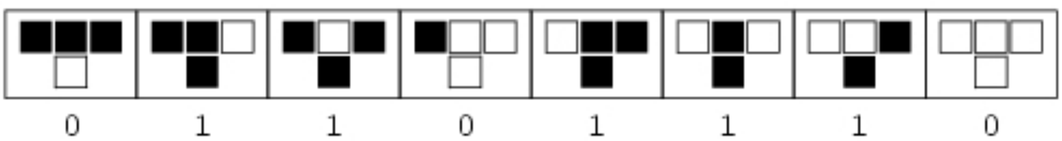
Nonetheless, according to the same author, it is also important to notice that:

> Cellular automata come in a variety of shapes and varieties. One of the most fundamental properties of a cellular automaton is the type of grid on which it is computed. The simplest such "grid" is a one-dimensional line. In two dimensions, square, triangular, and hexagonal grids may be considered.
>
> […]
>
> The number of colors (or distinct states) *k* a cellular automaton may assume must also be specified. This number is typically an integer, with *k* = 2 (binary) being the simplest choice. For a binary automaton, color 0 is commonly called "white" and color 1 is commonly called "black". However, cellular automata having a continuous range of possible values may also be considered.

Thus, if in the previous model all agents behave heterogeneously and globally – since their respective state do not depend on any specific neighborhood range, but rely on all agents states (global interaction), in cellular automata occurs the opposite situation. All agents behave locally and homogenously, where each cell represents an agent, so the focus in the study of the long range dependency arise is the underlying state transition rule in each automaton.

In this section of the work, only Elementary Cellular Automata are studied. Despite the word "elementary", they are still capable of displaying behaviors far from "elementary". In fact, they can exhibit even chaotic behavior. For example, the Rule 110 is one of the most intriguing and beautiful pieces of computer software ever written. With a set of 8 simple bitwise rules, one is able to, in principle, compute any calculation or emulate any computer program, as conjectured by Salem and Wolfram (1985) and proved by Cook (2004). (Note that rule numbers are encoded according to the 8-bit sequence of combinations, as in Figure 3.10.)

Figure 3.10 – Encoding of rule 110.



Hence, with this set of simple rules, one is able to generate very complex patterns that seem to evolve chaotically, depending on the initial conditions of the system. According to Wolfram (2002), this kind of system exhibits a behavior that is neither completely stable nor completely chaotic, generating localized structures that appear and interact in various complicated-looking ways, which characterizes what is called Class 4 behavior.

Keeping that in mind, the main idea of this second experiment is to show that there is another interesting possible source of complexity in terms of the rise of long memory processes, treating cellular automata as stochastic dynamic systems, where its individual transition rules may exhibit an equilibrium density over time (which are not intuitive). Hence, whenever large deviations from its respective equilibrium points occur, long-range dependency may arise. Furthermore, as it is shown later, there are some rules that, despite large deviations induced by the configuration of initial conditions, each system may also have convergence rates with different speeds, affecting the rise of long-range dependency.

Instead of working with regime switches explicitly, which is an already known source of long-range dependency, the goal of this present exercise is to show that large deviations caused by external shocks (such as a policy change) may lead to a behavior that mimics smooth regime switches, pointing back to the real need of building models with micro foundations of agents, as widely discussed in the economic literature.

So, in order to build such exercise, the first step is to aggregate all the black cells (represented by ones) in each time step, given the fact that they are randomly started according to a binomial probability function in two distinct cases:

- $p_0 = p^*$;
- and $p_0 \neq p^*$

where p^* denotes a stable fixed point (that is going to be introduced later). Consequently, if $p_0 = p^*$, no large deviations are expected; if $p_0 \neq p^*$, large deviations should be expected, but, also, as the time passes p_t should converge to p^*. Consequently, p_t also denotes the expected value of the density of black cells at each time step.

Moreover, in order to establish this experiment as a tool to investigate possible sources of long memory properties, if one considers the following expression for the density[1] of black cells:

$$\rho(t) = \frac{S(t)}{N} = \sum_{i=1}^{N} \frac{x_{i,t}}{N} \tag{46}$$

where $x_{i,t}$ represents the value of the cell at position i at time step t, $S(t)$ represents the amount of black cells present at each time-step and, over this number, long-range dependency statistics are calculated.

The reason to use such measurement is the fact that the Central Limit Theorem ensures that the variable $S(t)$ is normally distributed for a sufficiently large N if $x_{i,t}$ is binomially distributed. One should also notice that if p_t is a stable fixed point, in theory, $S(t)$ is also stable normal distributed over time. Hence, it is expected that the sequence of i.i.d. normal variables will compose a stable white-noise process, i.e. a short memory process.

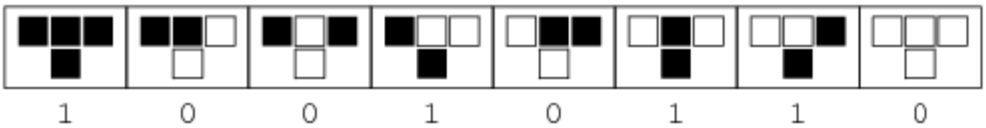
On the other hand, if p_t is not a stable fixed point, following the assumptions made, a sequence of regime switches happens until the system reaches its equilibrium point. Initializing such discrete system far away from its stable fixed point, would lead the system to a convergence to its respective equilibrium point, simulating (in a simplified manner) an external shock over dynamic stochastic system. Also, as it is known that regime switches may cause long-range dependency (STOCK; WATSON, 1996) – and this convergence behaves like a sequence of smooth regime switches, it is expected that this deviation from the equilibrium point may also generate a long memory behavior.

In order to derive a mathematical framework to establish the probabilities p^* for each rule, the approach here used assumes that the previous state of the system is a Markov process, i.e. a random system that changes its states according to a transition rule that only depends on the current state. Thus, the transition rule can in fact describe the probability that the next cell will be 1 or 0, according to a Binomial Probability, as described in Wolfram (1983).

[1] It is worth mentioning that the density of black cells is equal to the probability of picking a black cell at random.

So, for example, if rule 150 is considered (Figure 3.11), and established that $x_{i,t}$ represents the value of the cell at position i at time step t, it is possible to write the probability $P[x_{i,t+1} = 1]$. To accomplish that, it is necessary to sum up the individual probabilities of the cases where $x_{i,t+1} = 1$, according to a Binomial Probability.

Figure 3.11 – Encoding of rule 150.



For notation simplicity, writing $P\left[x_{i,t} = 1\right] = p_t$, $P\left[x_{i,t} = 0\right] = (1 - p_t)$ and $P\left[x_{i,t+1} = 1\right] = p_{t+1}$, one is able to develop the following equation:

$$p_{t+1} = p_t^3 + 2 \cdot p_t \cdot (1 - p_t)^2 + p_t^2 \cdot (1 - p_t) = f(p_t) \tag{47}$$

which is a discrete nonlinear system that describes the underlying probability in the Markov Process. This Markovian assumption is made in order to reduce the complexity of the mathematical treatment of the system.

If one solves this equation by establishing $p_{t+1} = p_t = p^*$, a fixed point is found, and consequently, its respective properties may be studied, like stability, in order to check if there is an attractor present in this system.

In this example, the system is characterized by the fixed points $p^* \in \{0, 0.5, 1\}$ and, according to the Lyapunov stability criteria, only 0.5 is a stable fixed point. The other fixed points are unstable. This can be verified by calculating:

$$\left.\frac{\partial f(p_t)}{\partial p_t}\right|_{p^*} = \lambda \tag{48}$$

and checking if $|\lambda| < 1$.

So, from the assumption of a Markov Process, if the initial conditions of the system are set using $p_0 \neq 0.5$, the system will converge towards $p_t = 0.5$.

That said, an additional noise is added to the system in order to have a true stochastic system and then, study its long memory properties. To accomplish that, it is introduced a probability based rule, where each cell has a probability p_N of having its state randomly modified, with equal probabilities of assuming 0 (white) or 1 (black), and $(1 - p_N)$ of having its state unmodified.

Having proposed the methodology, the rules 18, 106, 110, 150 and 182 are investigated.

These rules, according to the methodology discussed before, have the following fixed points between 0 and 1 (where valid probabilities can reside), noticing that "S" denotes a stable fixed point and "U" denotes an unstable fixed point.

Table 3.1 - Definition of rules.

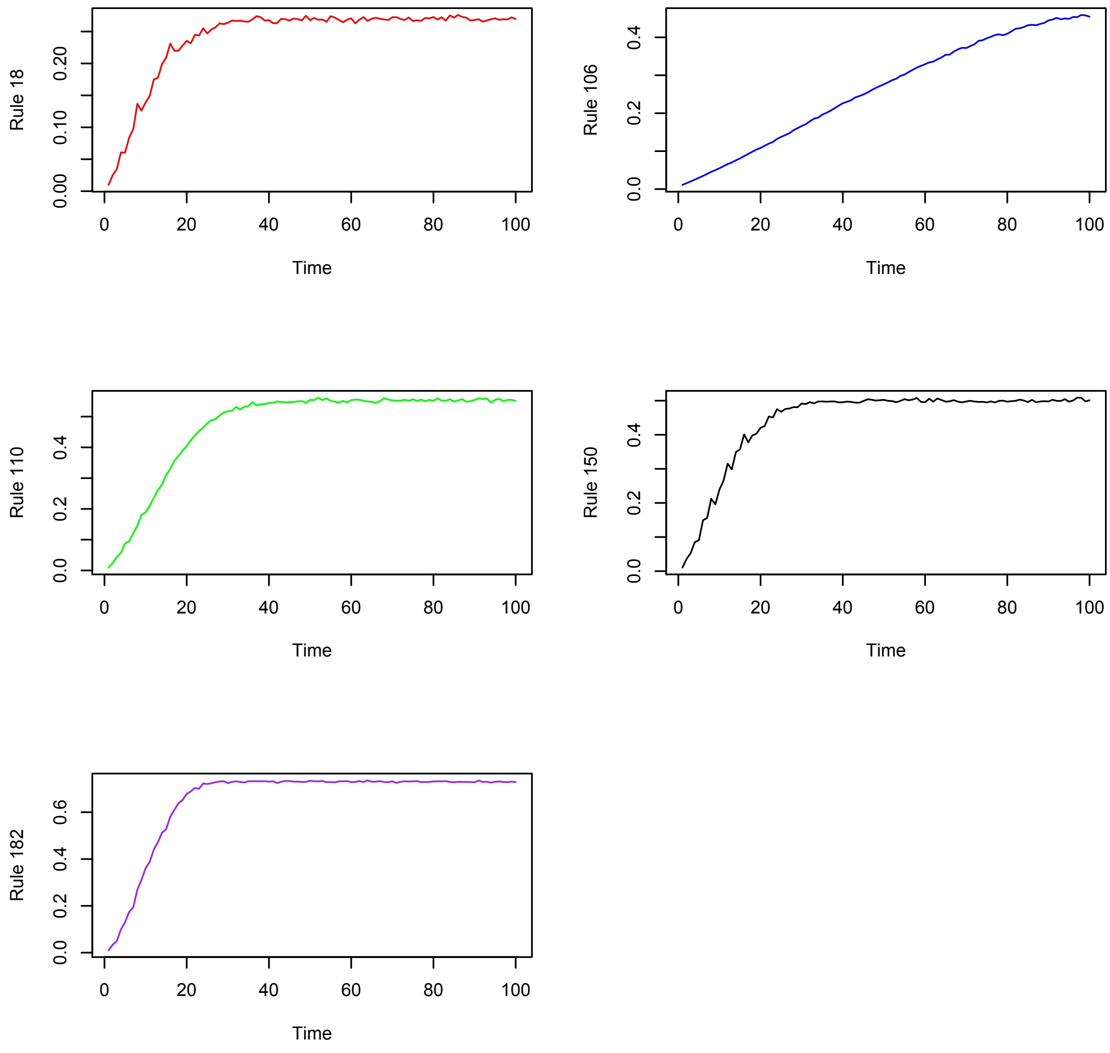
Rule	Fixed Point 1	Fixed Point 2	Fixed Point 3
Rule 18	0 (U)	$1 - \frac{\sqrt{2}}{2}$ (S)	-
Rule 106	0 (U)	0.5 (S)	-
Rule 110	0 (U)	$\frac{\sqrt{5}}{2} - \frac{1}{2}$ (S)	-
Rule 150	0 (U)	0.5 (S)	1 (U)
Rule 182	0 (U)	$\frac{2}{3}$ (S)	1 (U)

Hence, all studied rules presented here have, in theory, one valid attractor. In order to confirm that in practice, it was performed 100 simulations of each rule, with an initial configuration of $p_0 = 0.01$ and $p_N = 0.01$. In other words, the systems were initialized away from their respective equilibrium points, with a low amount of noise. After that, a mean evolution of each one of these processes were calculated and plotted in Figure 3.12.

First, it is worth noticing that the aggregated series seem to behave like having an implicit regime switch, when analyzed globally. In the first part of the evolution of the system, there is a first order behavior in the system, followed by saturation within a range where the steady state resides. This is what was exactly expected in terms of possibility to generate long-range dependency, as can be seen in Diebold and Inoue (2001).

Nonetheless, according to Figure 3.12, it is possible to see that not all rules converge properly to the stable fixed points inferred before. Wolfram (1983) states that the main cause of such behavior is the violation of the assumptions made (i.e. Markov Processes) and the presence of serial correlation / nonlinear behavior, which can be verified by the convergence towards to limiting probabilities – the systems do not converge instantaneously.

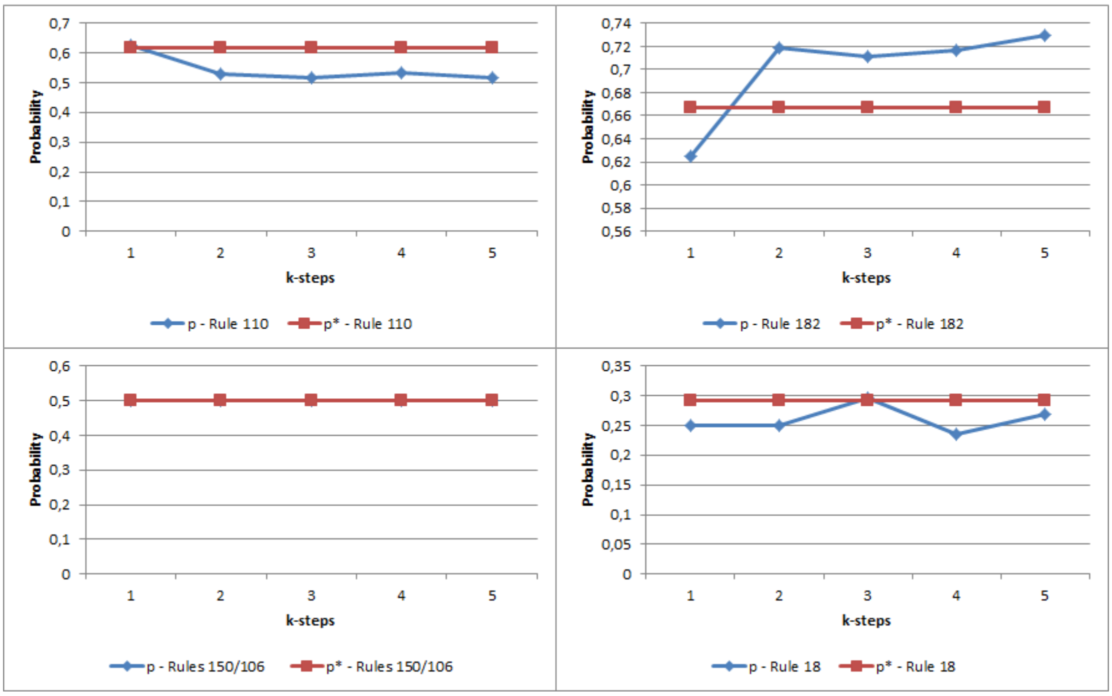
Figure 3.12 – Evolution of mean process.



Still, if a similar inference procedure is carried out, but assuming that only *k*-steps before the current realization the system is initialized by a random process, more consistent results can be obtained. For this, the recurrence relation is rebuilt in order to analyze the probabilities of $P\left[x_{i,t+1} = 1\right]$ in terms of

$\phi\left(x_{(i-2k,t-k)},\ldots,x_{(i,t-k)},\ldots,x_{(i+2k,t-k)}\right)$, assuming an uniform probability for every initial configuration, i.e. $\frac{1}{2^{(2\cdot k+1)}}$ for each case. Hence, Figure 3.13 is obtained.

Figure 3.13 – Limiting probabilities across different methodologies.

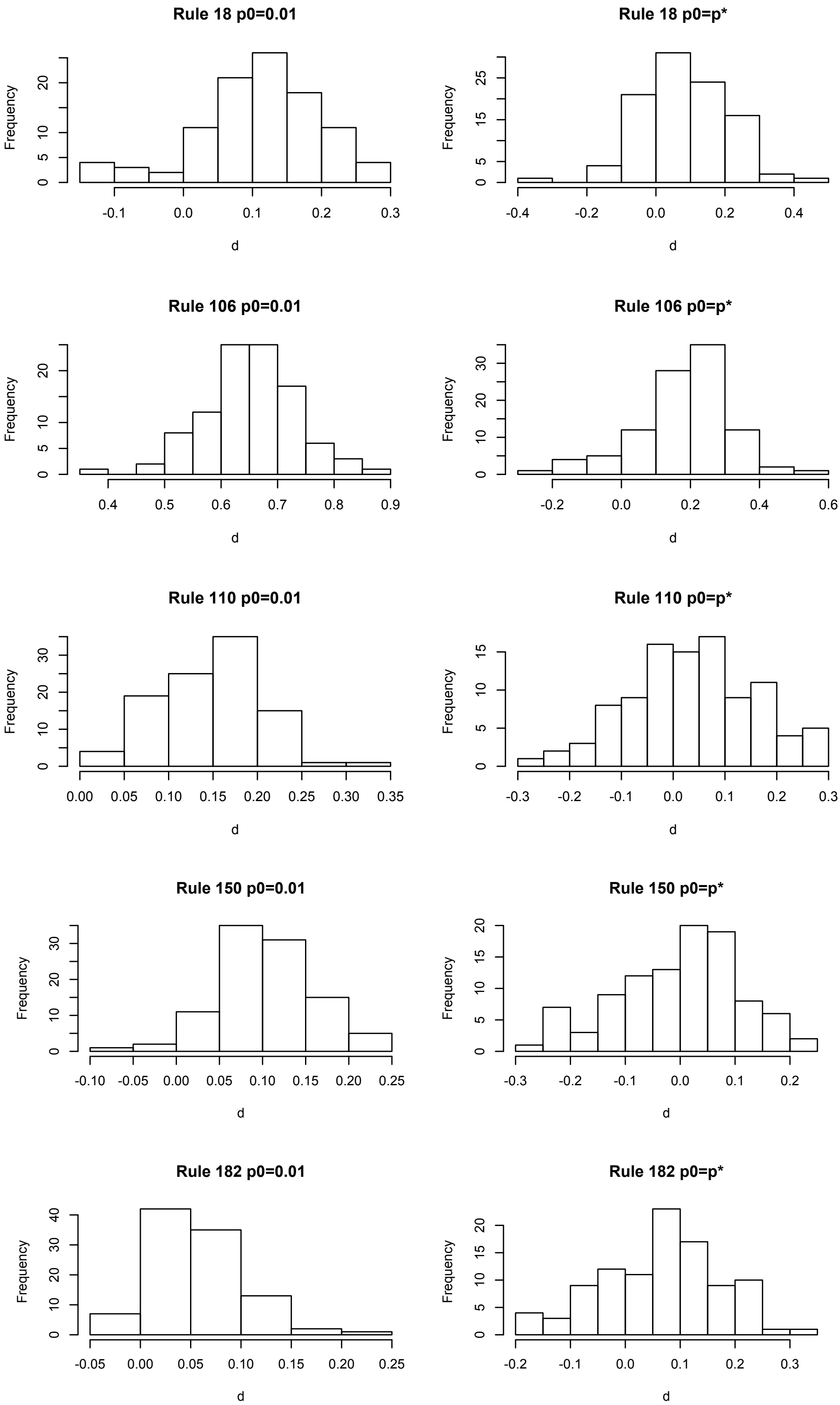


Consequently, it is possible to verify that, despite the fact that there are some deviations when a more robust calculation of the limit is made, the initial framework is relatively robust enough to point out where the limiting probabilities do not reside.

So, given such confirmation, a direct comparison can be carried out between the processes initialized with $p_0 = 0.01$ and $p_N = 0.01$; and $p_0 = p^*$ and $p_N = 0.01$ in order to verify the possible rise of long-range dependency, in terms of statistical tests, autocorrelation functions plot and distributions of the *d* fractional parameter.

In Figure 3.14, the rise of long-range dependency is shown, according to the GPH test, when the systems are initialized away from the limiting density, except for the rule 182, which seems to be insensitive to the fact that the initial state is initialized away from the limiting density; and rule 106, which seems to display long-range dependency independently from the initial state.

Figure 3.14 – Histogram of the GPH procedure for estimating *d*.



Such affirmation can be verified by inspecting the histograms and checking out that their mean are different from zero, which suggest the presence of long memory behavior. The same procedure was also carried out by calculating the Künsch (1987) procedure, according to Figure 3.15.

According to these results, it seems that the Künsch (1987) procedure is more sensitive than the GPH (1983), in terms of detecting long-range dependency. This procedure was able to detect a possible long-range dependency in all cases, as consequence of initializing the cellular automata away from their limiting densities, except for rule 106. So, basically, according to the Künsch (1987) procedure there is long-range dependency in rule 182 (when initializing it far away from its limiting probabilities), whereas it was not found according to the GPH (1983) procedure.

Hence, in order to complement these tests, it was generated the mean of the autocorrelation functions of all these systems, according to Figure 3.16.

According to Figure 3.16, it is possible to see that whenever the system is initialized away from its steady state (in red), it seems to generate some sort of slow decay in the autocorrelation function, except for the rule 106, which seems to always exhibit a slow decay, independently from the initial state; and rule 182, which displays a behavior more similar to an exponential decay (but still, being hard to distinguish from a hyperbolical decay). Moreover, a direct comparison can be made between Figure 3.16 and the evolution of the mean process. While rules 18, 110 and 150 display a similar speed of convergence towards to an equilibrium point, rule 106 exhibits a very slow speed of convergence, and rule 182 displays a very fast convergence towards the steady state of the system.

Figure 3.15 – Histogram of the Künsch (1987) procedure for estimating *d*.

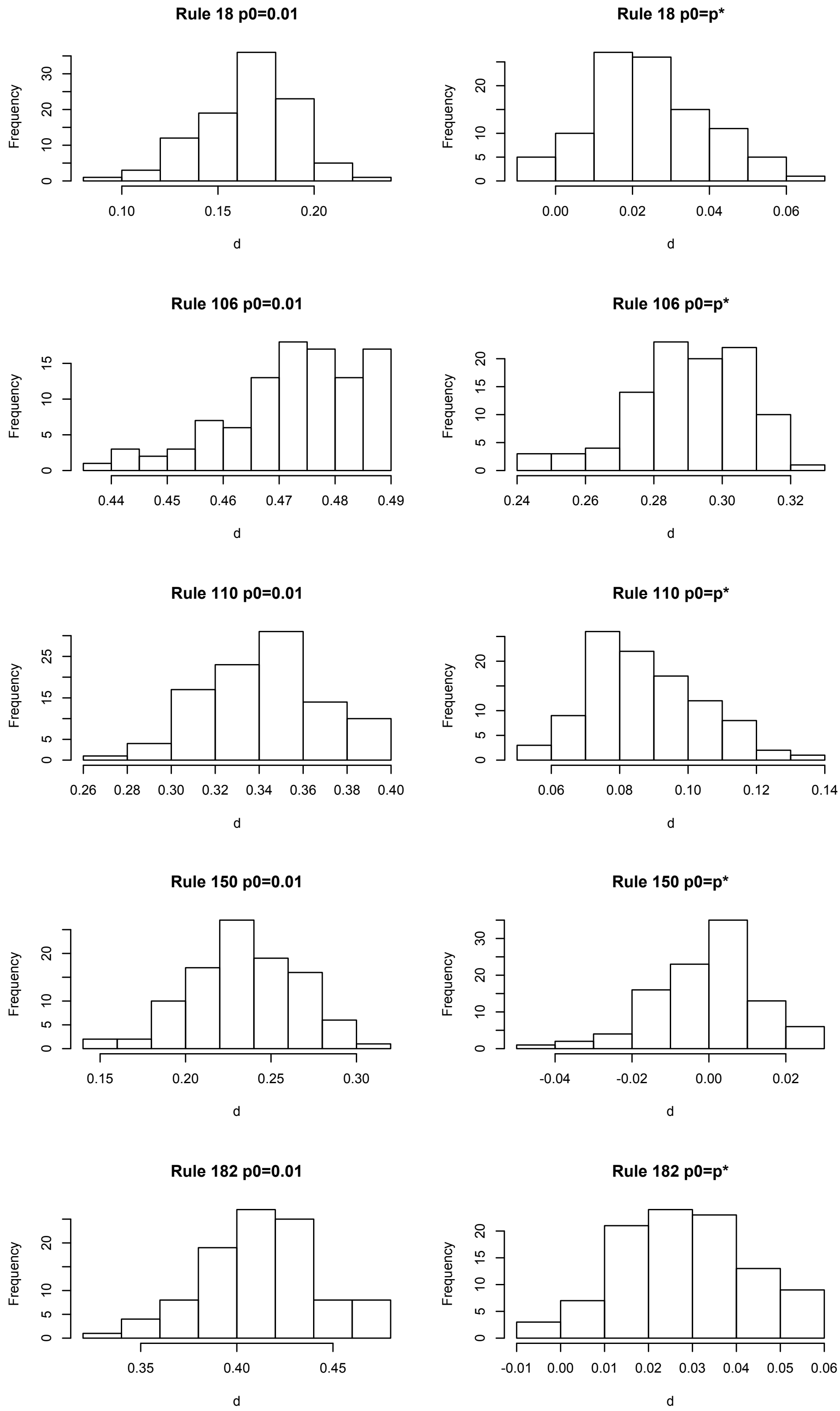
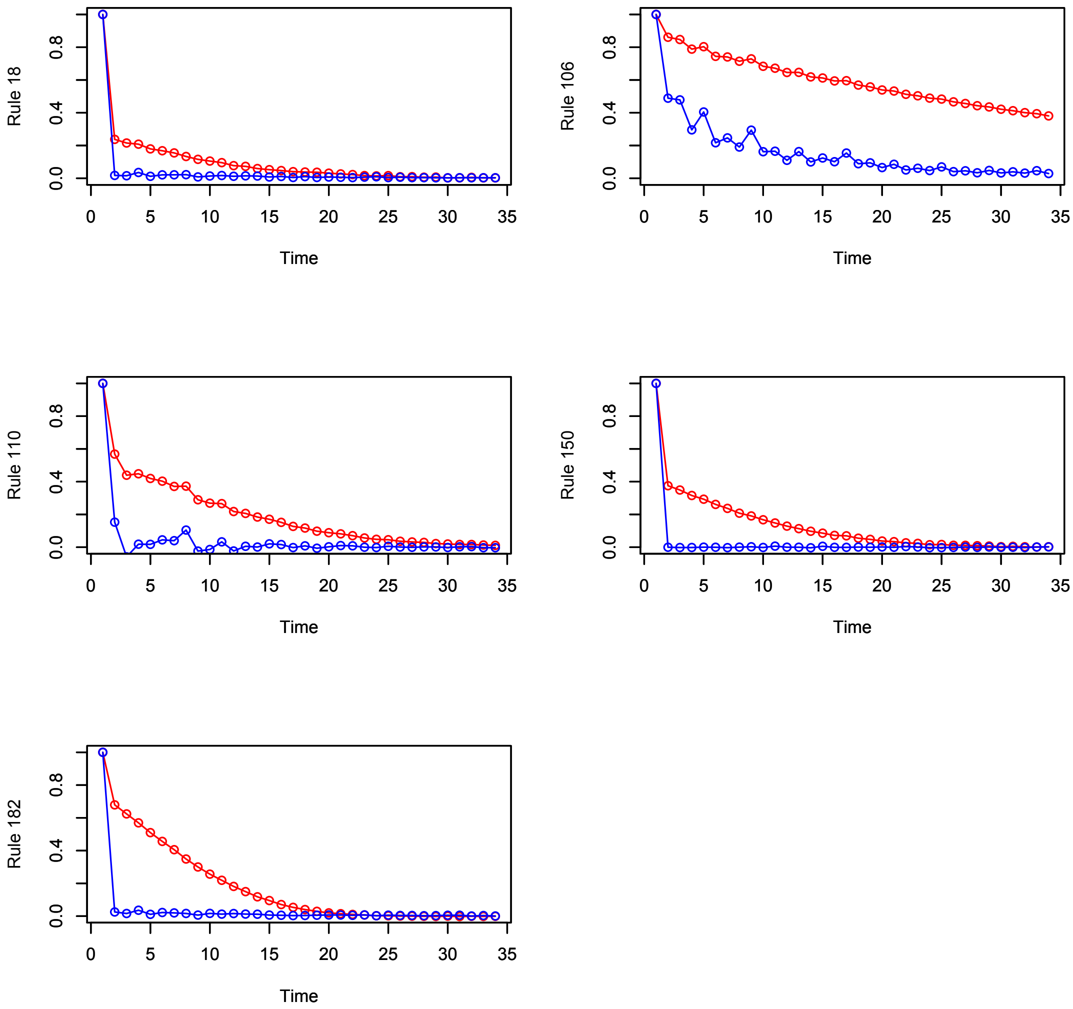


Figure 3.16 – Autocorrelation function plots.



Moreover, it was carried out the V/S and R/S Statistic Tests following the same methodology, as in the previous exercise, in order to detect the presence of long-range dependency. In Table 3.2, it follows the results for the systems initialized with $p_0 = 0.01$. The same was done to obtain the results for $p_0 = p^*$ shown in Table 3.3.

As can be seen, in Tables 3.2 and 3.2, these tests basically summarize the hypothesis of long memory behavior discussed in the other tests and on the evaluation of the autocorrelation function plots. Hence, basically, all rules exhibit the rise of long-range dependency, when initialized far away from their steady state densities, except for rule 106.

Table 3.2 – Acceptance ratio of statistic tests for $p_0 = 0.01$.

Rule	V/S Test Acceptance for Long-Memory at 95%	R/S Test Acceptance for Long Memory at 95%
Rule 18	92%	88%
Rule 106	100%	100%
Rule 110	99%	91%
Rule 150	99%	90%
Rule 182	84%	3%

Table 3.3 – Acceptance ratio of statistic tests for $p_0 = p^*$.

Rule	V/S Test Acceptance for Long-Memory at 95%	R/S Test Acceptance for Long Memory at 95%
Rule 18	27%	24%
Rule 106	84%	85%
Rule 110	30%	22%
Rule 150	8%	11%
Rule 182	35%	24%

Furthermore, there is no consensus whether there is a rise of long memory behavior in rule 182, given the fact that it converges very rapidly towards the equilibrium density. Thus, an interesting discussion can be made in terms of these two pathological cases.

First, if rule 106 is considered, it is worth mentioning that the transition rule for each cell can be defined by the following expression[2]:

$$x_{i,t+1} = x_{i-1,t} \cdot x_{i,t} \cdot \overline{x_{i+1,t}} + x_{i-1,t} \cdot \overline{x_{i,t}} \cdot x_{i+1,t} + \overline{x_{i-1,t}} \cdot x_{i,t} \cdot x_{i+1,t} + \overline{x_{i-1,t}} \cdot \overline{x_{i,t}} \cdot x_{i+1,t} \quad (49)$$

where $\overline{x_{i,t}}$ denotes the "not" operation over the cell at the *i*-th position, in the instant *t*. So, if $x_{i,t} = 1$, then $\overline{x_{i,t}} = 0$, and vice-versa.

[2] Notice that, in this context, the multiplication equals to the logical "and" the sum to the logical "or", given that $x_{i,t}$ can only assume values 0 and 1.

Additionally, if one considers the density of the system given by

$$\rho(t) = \sum_{\mathrm{i}=1}^{\mathrm{N}} \frac{\mathrm{x}_{\mathrm{i,t}}}{N} \tag{50}$$

it is possible to realize that the system is stationary when

$$\rho(t) = \rho(t+1) \tag{51}$$

So, by plugging the initial definition of the transition rule into the density expression, one obtains:

$$\sum_{\mathrm{i}=1}^{\mathrm{N}} \mathrm{x}_{\mathrm{i,t}} = \sum_{\mathrm{i}=1}^{\mathrm{N}} \begin{bmatrix} \mathrm{x}_{\mathrm{i-1,t}} \cdot \mathrm{x}_{\mathrm{i,t}} \cdot \overline{\mathrm{x}_{\mathrm{i+1,t}}} + \mathrm{x}_{\mathrm{i-1,t}} \cdot \overline{\mathrm{x}_{\mathrm{i,t}}} \cdot \mathrm{x}_{\mathrm{i+1,t}} + \overline{\mathrm{x}_{\mathrm{i-1,t}}} \cdot \mathrm{x}_{\mathrm{i,t}} \cdot \mathrm{x}_{\mathrm{i+1,t}} \\ + \overline{\mathrm{x}_{\mathrm{i-1,t}}} \cdot \overline{\mathrm{x}_{\mathrm{i,t}}} \cdot \mathrm{x}_{\mathrm{i+1,t}} \end{bmatrix} \tag{52}$$

Consequently, by factorizing:

$$\sum_{\mathrm{i}=1}^{\mathrm{N}} \mathrm{x}_{\mathrm{i,t}} = \sum_{\mathrm{i}=1}^{\mathrm{N}} \begin{bmatrix} \mathrm{x}_{\mathrm{i,t}} \cdot \left(\mathrm{x}_{\mathrm{i-1,t}} \cdot \overline{\mathrm{x}_{\mathrm{i+1,t}}} + \overline{\mathrm{x}_{\mathrm{i-1,t}}} \cdot \mathrm{x}_{\mathrm{i+1,t}}\right) \\ + \overline{\mathrm{x}_{\mathrm{i,t}}} \cdot \left(\overline{\mathrm{x}_{\mathrm{i-1,t}}} \cdot \mathrm{x}_{\mathrm{i+1,t}} + \mathrm{x}_{\mathrm{i-1,t}} \cdot \mathrm{x}_{\mathrm{i+1,t}}\right) \end{bmatrix} \tag{53}$$

Thus, in order to study the conditions where the system is stationary, it is useful to rewrite the previous expression as a difference equation in terms of *i*, as it follows:

$$\mathrm{x}_{\mathrm{i}} = \mathrm{x}_{\mathrm{i}} \cdot (\mathrm{x}_{\mathrm{i-1}} \cdot \overline{\mathrm{x}_{\mathrm{i+1}}} + \overline{\mathrm{x}_{\mathrm{i-1}}} \cdot \mathrm{x}_{\mathrm{i+1}}) + \overline{\mathrm{x}_{\mathrm{i}}} \cdot (\overline{\mathrm{x}_{\mathrm{i-1}}} \cdot \mathrm{x}_{\mathrm{i+1}} + \mathrm{x}_{\mathrm{i-1}} \cdot \mathrm{x}_{\mathrm{i+1}}) \tag{54}$$

which has the following system as solution:

$$\begin{cases} (\mathrm{x}_{\mathrm{i-1}} \cdot \overline{\mathrm{x}_{\mathrm{i+1}}} + \overline{\mathrm{x}_{\mathrm{i-1}}} \cdot \mathrm{x}_{\mathrm{i+1}}) = 1 \\ (\overline{\mathrm{x}_{\mathrm{i-1}}} \cdot \mathrm{x}_{\mathrm{i+1}} + \mathrm{x}_{\mathrm{i-1}} \cdot \mathrm{x}_{\mathrm{i+1}}) = 0 \end{cases} \tag{55}$$

As there are no possible negative values, it follows that:

$$\overline{x_{i-1}} \cdot x_{i+1} = 0 \tag{56}$$

Solving the following simplified system, one obtains:

$$\begin{aligned} x_{i-1} \cdot \overline{x_{i+1}} &= 1 \\ x_{i-1} \cdot x_{i+1} &= 0 \end{aligned} \tag{57}$$

Hence:

$$\begin{cases} x_{i-1} = 1 \\ x_{i+1} = 0 \end{cases} \tag{58}$$

So, for example, if $x_{i,0} = 0$, without any noise and, with only one $x_{k,0} = 1 \mid k \in \{0, N\}$, the system will remain at its stationary state. The same happens if all the cells are initialized with 0. And, moreover, it is possible to verify that all stationary solutions are those where the cells are initialized as triplets of the form “1_0”.

From the stationary state with $S(t) = 1$ (in other words, with only one black cell), when a very low-level of noise is added, the initial state of the system is preserved with a slow convergence towards the limit density, as verified in Figure 3.13. If there is no noise, nothing happens to the system.

However, it is also interesting to notice that if all cells are initialized as one, except for only one cell, being equal to 0, the long memory behavior becomes more evident, when no noise is present in the system.

On the other hand, when analyzing rule 182, it is possible to derive the following transition rule:

$$\begin{aligned} x_{i,t+1} = {} & x_{i-1,t} \cdot x_{i,t} \cdot x_{i+1,t} + x_{i-1,t} \cdot \overline{x_{i,t}} \cdot x_{i+1,t} + x_{i-1,t} \cdot \overline{x_{i,t}} \cdot \overline{x_{i+1,t}} + \overline{x_{i-1,t}} \cdot x_{i,t} \cdot \overline{x_{i+1,t}} \\ & + \overline{x_{i-1,t}} \cdot \overline{x_{i,t}} \cdot x_{i+1,t} \end{aligned} \tag{59}$$

Repeating the same previous steps, as done for rule 106, it is possible to notice that:

$$x_{i,t+1} = x_{i,t} \cdot \left(x_{i-1,t} \cdot x_{i+1,t} + \overline{x_{i-1,t}} \cdot \overline{x_{i+1,t}}\right) + \overline{x_{i,t}} \cdot \left(x_{i-1,t} \cdot x_{i+1,t} + \overline{x_{i-1,t}} \cdot x_{i+1,t} + x_{i-1,t} \cdot \overline{x_{i+1,t}}\right) \tag{60}$$

Consequently, it is important to verify that the first term of the previous equation given by:

$$\overline{x_{i,t}} \cdot \left(x_{i-1,t} \cdot x_{i+1,t} + \overline{x_{i-1,t}} \cdot x_{i+1,t} + x_{i-1,t} \cdot \overline{x_{i+1,t}}\right) \tag{61}$$

drives the system towards a reversibility trend, differently from rule 106, which keeps a strong inertia (long-range dependency cause).

If a simple initial state is taken, with all $x_{i,t} = 0$, except for one $x_{k,t} = 1$, the system converges very rapidly towards the limit density. Consequently, the autocorrelation function tends to have an exponential decay, even with a significant distance from the limit density, precisely because of the inherent rapid convergence.

When compared to a non-pathological case, like rule 110, again, following the same steps, it is possible to verify that:

$$x_{i,t+1} = x_{i-1,t} \cdot x_{i,t} \cdot \overline{x_{i+1,t}} + x_{i-1,t} \cdot \overline{x_{i,t}} \cdot x_{i+1,t} + \overline{x_{i-1,t}} \cdot x_{i,t} \cdot x_{i+1,t} + \overline{x_{i-1,t}} \cdot x_{i,t} \cdot \overline{x_{i+1,t}} + \overline{x_{i-1,t}} \cdot \overline{x_{i,t}} \cdot x_{i+1,t} \tag{62}$$

Factoring the previous expression:

$$x_{i,t+1} = x_{i,t} \cdot \left(x_{i-1,t} \cdot \overline{x_{i+1,t}} + \overline{x_{i-1,t}} \cdot x_{i+1,t} + \overline{x_{i-1,t}} \cdot \overline{x_{i+1,t}}\right) + \overline{x_{i,t}} \cdot \left(x_{i-1,t} \cdot x_{i+1,t} + \overline{x_{i-1,t}} \cdot x_{i+1,t}\right) \tag{63}$$

one can realize that the term $\overline{x_{i,t}} \cdot \left(x_{i-1,t} \cdot x_{i+1,t} + \overline{x_{i-1,t}} \cdot x_{i+1,t}\right)$ drives the system towards a reversibility much more slower than that one observed in rule 182. That is why regime switches in the non-pathological cases cause long-range dependency.

Therefore, an important conclusion can be drawn from this exercise. Rules that drive local interactions between homogenous agents and the presence of large deviations from the equilibrium points, jointly, can generate long-range dependency.

Additionally, it is important to state that the straightforward analyses carried out in the present exercise is possible due to the mathematical simplicity of the cellular automata here analyzed. If more complex systems were analyzed, it becomes more difficult to draw possible reasons to the rise of long memory behavior.

Thus, this is one of the main reasons why long-range dependency should be studied as an emergent property of the system.

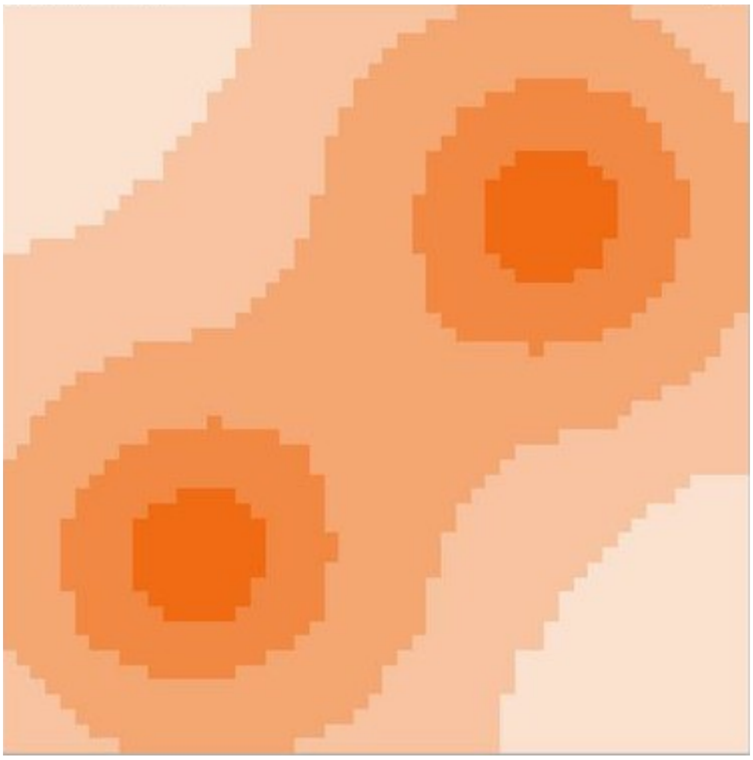
3.5 SUGARSCAPE – A SIMULATION OF AN ARTIFICIAL SOCIETY BASED ON A MULTI-AGENT SYSTEM

The Sugarscape is a large scale agent-based model composed of (of course) agents, the environment (a two-dimensional grid) and a set of rules which governs the interactions between the agents and the environment. This model was originally presented in Epstein and Axtell (1996).

Each cell within the grid can contain different amounts of sugar and sugar capacity, where initially, the amount of sugar is equal to the sugar capacity. Whenever a patch is exploited, the amount of sugar is decreased, but it has a grow back constant that allows restoring part of its sugar capacity.

These grid cells are randomly initialized in order to introduce spatial complexity in the simulation, as presented in Figure 3.17.

Figure 3.17 – Sugar spatial distribution.



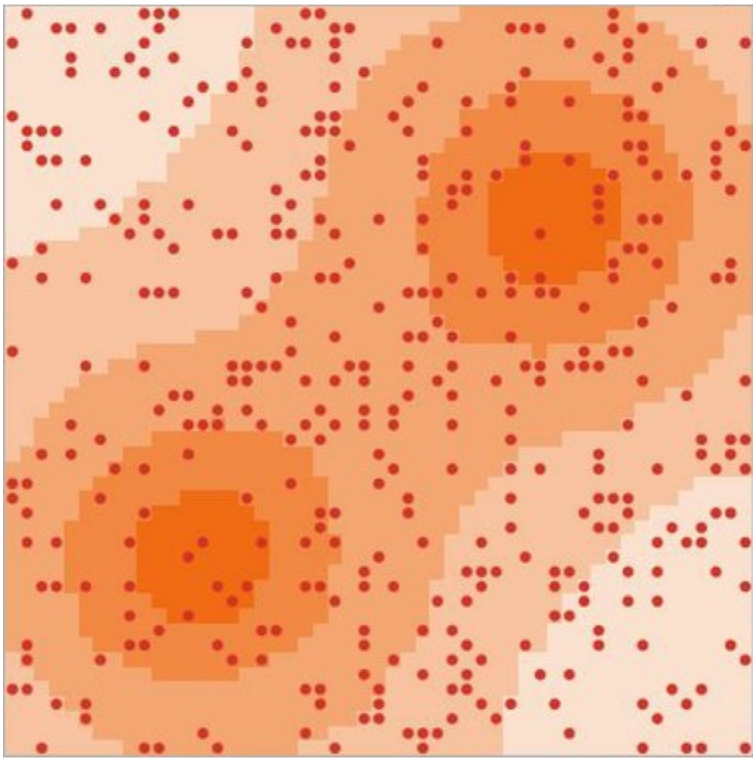
Hence, as can be seen in Figure 3.17, the darker cells represent patches with higher sugar values, and the lighter ones patches with lower sugar values.

Moreover, this grid is populated by individual agents that have different states initialized randomly according to a uniform distribution:

- Amount of stocked sugar (defined by a range of Minimum and Maximum Initial Sugar Stocks);
- Metabolism (defined by a range of Minimum and Maximum Metabolism rates);
- Vision (defined by a range of Minimum and Maximum Vision capability);
- Life Expectancy (defined by a range of Minimum and Maximum Life Expectancy).

Hence, having these variables randomly initialized, the agent actions turn to be heterogeneous among themselves. Moreover, they are placed randomly within this grid, as is shown in Figure 3.18.

Figure 3.18 – Spatial distribution of artificial agents.



Given that, agents can search and gather sugar from cells within their vision range, they consume sugar from their stock according to their metabolism and they die if they run out of stock or if they achieve their life expectancy – when they die, they are replaced by other agents with random initial states.

Furthermore, they can explore only one cell grid at each tick and they select the cell grid according to the highest sugar value. If several patches exist with the same value, the agent chooses the closest one.

So, individual agents act with bounded rationality while exploiting the patches, given the fact that the choice of the patches is made according to the distance and the highest sugar value, but they do not coordinate their actions between themselves, which would be establishing who will explore a specific site, leading to suboptimal choices.

Finally, the execution order of this model is such that agents perform their respective operations according to a pre-specified set of rules, and then, all operations within each cell grid are carried out.

It is worth mentioning that there are also several other versions of this model that include more complex iterations between the agents and more spatial complexity, such as the introduction of another commodity (spice), combat rules between agents, sex and reproduction, genetic inheritance and so on, but for the sake of this work, this basic set of rules exhibit the necessary features to study the role of spatial complexity in terms of rising long range dependency properties.

To achieve that, the present author modified the original model, in order to remove agent heterogeneity, by setting maximum and minimum life expectancy to the number of simulation ticks (avoiding the generation of new random agents by death cause), and the same for vision, initial quantity of sugar and metabolism. Hence, all agents behave the same way, and they do not get replaced by another agent (as in the original model) with a random set of characteristics which may introduce heterogeneity in the system.

Aiming at comparisons between a complex environment and a simple environment, the present author also modified the code to allow the removal of spatial complexity, by setting all patches identically and configuring the sugar restoration parameter to be larger than the agents' metabolism. Moreover, all heterogeneity between the patches is removed by imposing the same sugar capacity values for all of them – all patches having maximum capacity.

Consequently, in a first simulation, the system behaves like a stable intransient deterministic system.

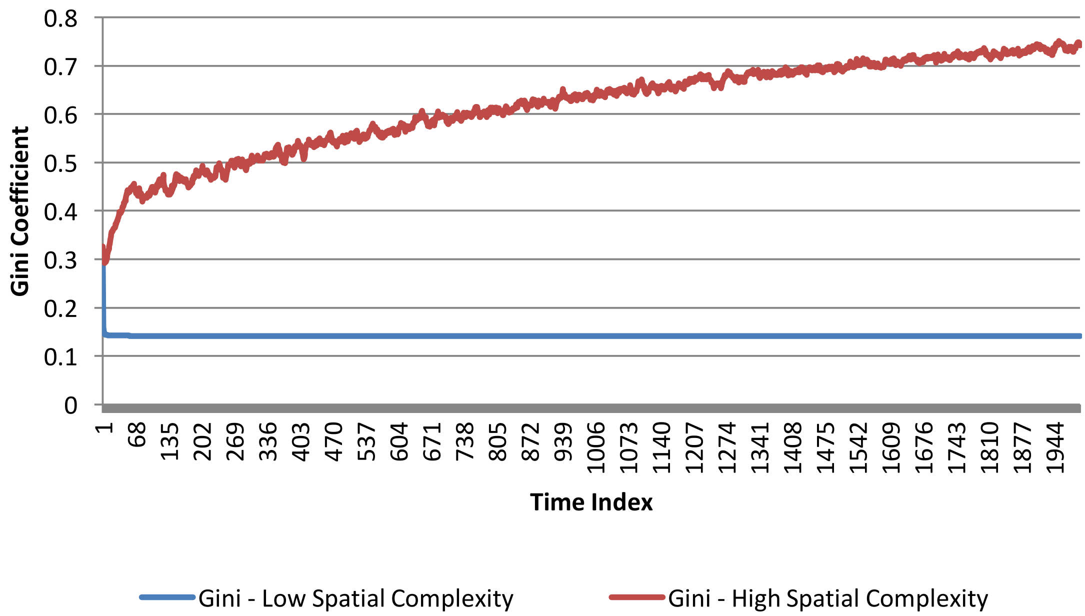
After that, in a second simulation, the model is again modified, in order to generate spatial complexity. To achieve that, the sugar restoration and sugar capacity parameters are set to default values (identical to the original model), keeping all agents initially homogenous between themselves. Furthermore, heterogeneity is imposed over the patches by setting different and random sugar capacities, where only 5% of the patches have the maximum capacity.

Hence, this second configuration produces a result much similar to a stochastic process.

In order to check such results, a Gini coefficient time series was calculated over the food quantity that each agent has – in this case sugar is the Wealth in this simple artificial economy, simulated over 2000 periods.

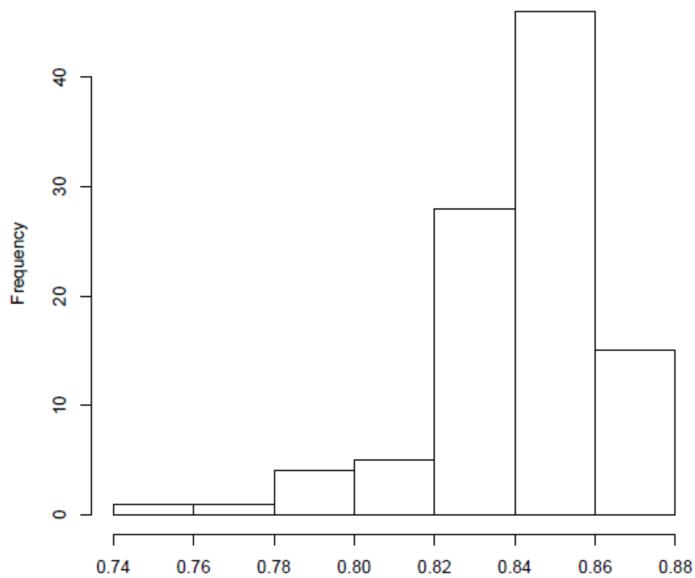
Then, as can be seen in Figure 3.19, in the first discussed modification the system rapidly converges towards a fixed point. However, in the second configuration, the system behaves like a stochastic process, producing long memory properties that are going to be discussed later.

Figure 3.19 – Evolution of simulated Gini indexes.



Keeping in mind such stochastic properties observed in the second configuration, the GPH procedure was performed in order to analyze the long memory properties of such process and it was obtained the distribution presented in Figure 3.20, over 100 different experiments.

Figure 3.20 – Histogram of the parameter *d*.



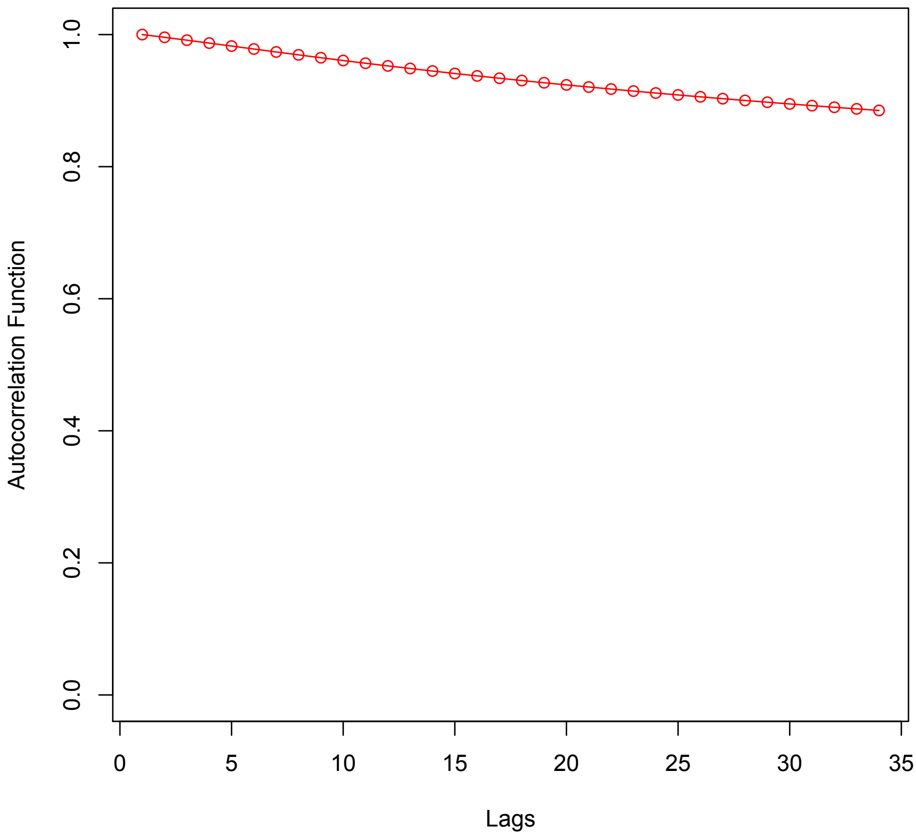
As can be seen in Figure 3.20, there is a strong evidence of the presence of long-range dependency in the analyzed stochastic process. In this specific case, the mean of this distribution is 0.8418075 and its standard deviation is 0.02211333. It is also important to mention that the Künsch (1987) procedure failed in the R package and consequently did not produce any result that could be shown in this chapter.

Additionally, it is also important to notice that the autocorrelation function decays very slowly over the time, as can be seen in Figure 3.21.

Furthermore, both V/S and R/S Statistic Tests point out that all of the 100 generated samples have long-range dependency, at a 95% confidence level. Hence, there is no doubt that long memory behavior emerges in this case.

Therefore, from the cases analyzed along this section, it is possible to show that spatial complexity is also important in the rise of long memory properties, when setting all agents behavior homogeneously and avoiding direct local interactions between themselves – in order to try to avoid all factors discussed previously.

Figure 3.21 – Autocorrelation function plot.



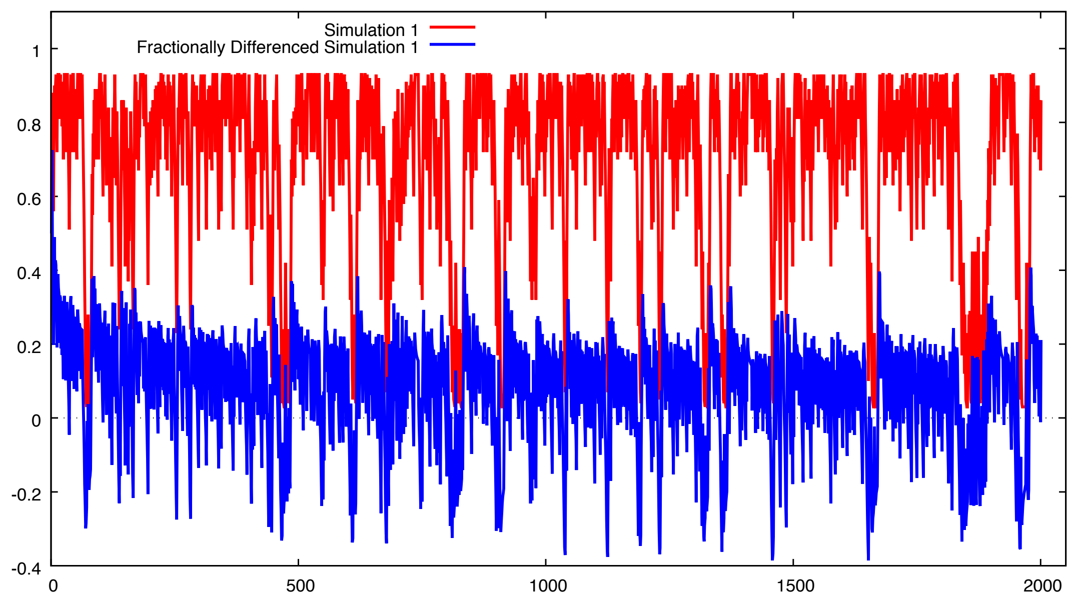
3.6 IDENTIFYING SHOCKS AND DISTURBANCES WITH LONG-MEMORY FILTERS

Having presented possible sources of long-range dependency, it makes sense starting to discuss whether it is possible to use the reversal approach to identify stylized facts, such as large disturbances from the equilibrium points, in order to advance to possible practical applications of long memory filters and its statistical theory.

The main point to be discussed in this section is the usage of long memory filters and how they can be used to identify anomalies in the temporal dimension.

If one departs from the artificially generated time series from the first analyzed model (Prediction Market example) and calculates the fractionally differenced series, as in Figure 3.22, it is possible to notice that the fractionally differenced simulation contains all significant short-term fluctuations that occurs in the original simulation, except for the fluctuations caused in lower frequencies. The large deviations from the equilibrium points can be seen right in the beginning of the fractionally differenced series and at the end of it, where the largest disturbances occur.

Figure 3.22 – Plots of the original simulation and fractionally differenced series over time – simulated prices.



The same conclusions can be obtained while observing Figures 3.23 and 3.24. The shift in the level caused by a large deviation from the equilibrium in the Rule 110 (Figure 3.23) can be better seen in the fractionally differenced series, which is caused by the elimination of the effects of autocorrelation. Hence, a good part of the observed fluctuations are caused by the deviation from the density equilibrium, and they occur until the system reaches stability.

After reaching stability, only standard fluctuations occur, as a consequence of nonlinearity and stochastic fluctuations introduced in the system.

The same pattern can be observed in Figure 3.24. When spatial complexity is introduced, it generates inequality among the simulated agents, as a consequence of competition between themselves for resources (in this case, sugar). This inequality is captured by the Gini Coefficient over time.

The initial shock in the inequality caused by competition (which is a consequence of the resource distribution in the nature) is very slowly dissipated. Consequently, the fractionally differenced series can be used as a tool to evaluate how changes in resources availability lead to increases (decreases) in inequality.

Figure 3.23 – Plots of the original simulation and fractionally differenced series over time – rule 110.

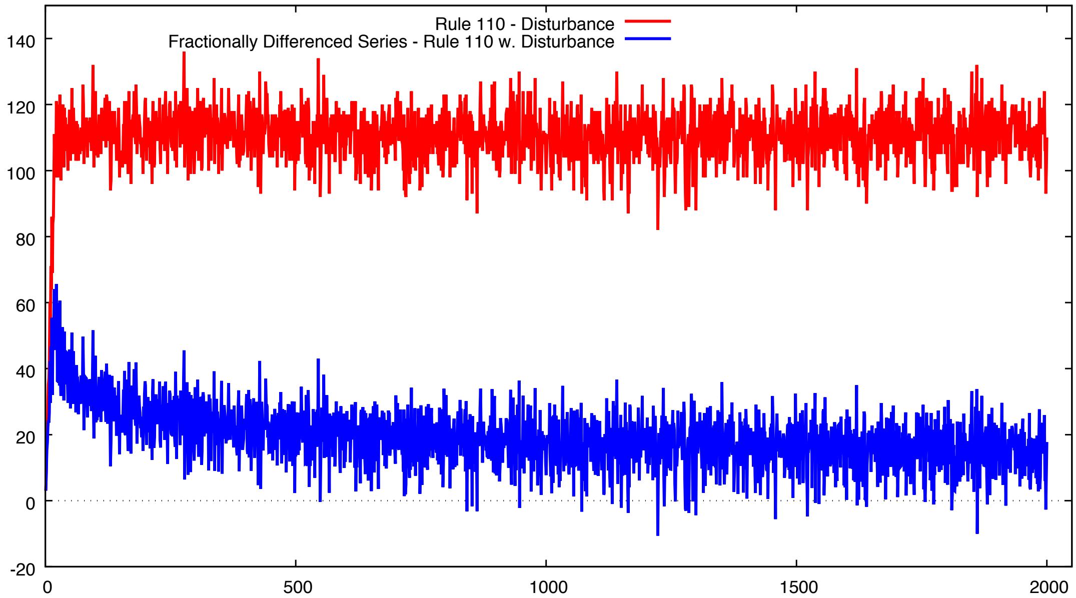
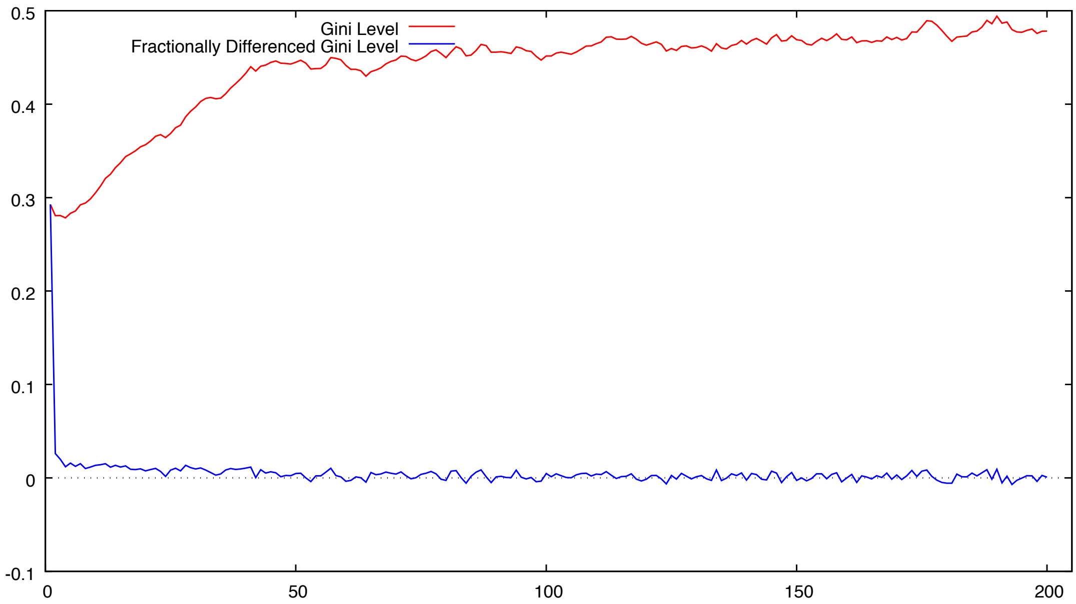


Figure 3.24 – Plots of the original simulation and fractionally differenced series over time – sugarscape models.



Mathematically speaking, when considering a time series given by an ARFIMA(p,d,q) process:

$$Y_t \cdot (1-L)^d = \frac{(1-\theta(q)\cdot L)}{(1-\phi(p)\cdot L)} \cdot \varepsilon_t \tag{64}$$

and, also, if one considers

$$X_t = Y_t \cdot (1-L)^d \tag{65}$$

then, one obtains a simple ARMA(p,q) process:

$$X_t = \frac{(1-\theta(q)\cdot L)}{(1-\phi(p)\cdot L)} \cdot \varepsilon_t \tag{66}$$

So, basically, what is being exhibited in Figures 3.22, 3.23 and 3.24 is Y_t and X_t, this last being a short memory process with only short-term fluctuations. The fact that X_t is free from long range dependency becomes very attractive when stylized facts must be identified in the time frequency (i.e. time frames of specific events) in different economies and within specific economies.

This is very important, because an important parallel can be made between the computational agent based examples simulated in this chapter and real economies:

- Real economies are also subject to large deviations from their equilibrium, which depending on their specific motion laws may generate different responses;
- Real markets have heterogeneous agents;
- Real economies are subject to spatial changes that generates spatial complexity (such as weather and climate changes), with agents competing all the time for resources.

Consequently, it is expected that aggregated time series observed in real economies are biased by the long term autocorrelation, which is a result of specific market idiosyncrasies present in them.

So if one is able to filter out this long range dependency, there are several potential practical applications, including the identification of shocks across different economies – making it easier to observe phenomena such as risk contagion (this is developed in Chapter 4); and simulation of counter-factual scenarios in the presence of long-range dependency (discussed in Chapter 5).

4 USING LONG MEMORY FILTERS TO IDENTIFY RISK CONTAGION ACROSS DIFFERENT ECONOMIES

4.1 PROLOGUE OF THE CHAPTER

Risk analysis plays an important role in a wide extent to economics practitioners in several areas, such as: investment decisions, portfolio allocations, establishment of preventing costs as consequence of their respective materialization and calculation of potential gains/losses (HULL, 2015). In addition to that, it is a consensus that most of the existing tools were not able to predict the domino effect that characterized the risk contagion across the world.

Hence, the main idea of this chapter is to develop and provide a useful tool, which works using the long memory filtering concept, that separates what are the true short term fluctuations that occurs in the risk aversion of economic agents, and how these true short term fluctuations (the true risk shocks) spread into different economies.

This treatment seems to be very important given that, as discussed in Chapter 3, economic agents are subject to market idiosyncrasies that play an important role in the evolution of the risk over time – risk aversion contains long memory effects. So filtering out possible market specific over-reactions may turn to eliminate spurious effects that contaminates the correlation between risk aversion in pairs of countries – keeping only the desired stylized facts, enabling one to obtain a good risk contagion index.

4.2 THE IMPORTANCE OF RISK ANALYSIS AND RISK CONTAGION

Globalization, deregulation and technological advances (such as real-time trading) have significantly increased the transmission of information, causing a greater integration between markets, as discussed in Bergmann et al. (2014). Still according to these authors, these factors contributed to the intensification of the phenomenon of financial risk contagion, which, consequently, increases the complexity in the risk analysis, because external factors must be taken into account when calculating the risk exposure.

To support this statement, it is worth mentioning that the world is witnessing “one of the largest and most complex financial crisis to date”, which started in the U.S. housing market, spread to the American financial market and, finally, to the rest of the world, according to Bianconi, Yoshino and Sousa (2013).

Hence, crucial and complex decisions must be made by portfolio managers, central bankers and regulatory authorities in order to prevent more harmful consequences of this crisis. And to make such decisions, questions “concerning the benefits of diversification, the robustness of financial institutions, and the extent of the domino effect” must be answered, according to Bergmann et al. (2014).

Other potential questions of interest can also be answered when assessing the risk contagion across different economies, such as whether BRIC countries can be major players in sustaining the world economic growth, or if specific economies can be considered insulated from the financial stress of the U.S., according to Bianconi, Yoshino and Sousa (2013).

Thus, quantitative analysis of risk contagion becomes a very important tool, given the fact that understanding how volatility is spread from a specific market to another plays an essential role in the forecast of local risk levels, and consequently, the respective asset price.

Having shown the big picture of the need of measurement instruments of financial shocks, this chapter focuses in developing an indicator of risk contagion, based on the mathematical framework developed in the previous chapters of this thesis.

To develop this task, a brief discussion of risk contagion concept is made. After that, a review of the most important methodologies for the calculation of different assets/markets correlation is made. Then, based on the previous points discussed in this thesis (about the nature of fractional integration in real market data), a new risk contagion correlation based index is developed, by applying a fractional difference filter over conditional variance data and by combining this to the traditional rolling correlation index.

Afterwards, comparisons to a benchmark model – Dynamic Conditional Correlation (DCC) developed by Engle (2002) – are made, showing that this present tool has a set of interesting features that makes it very attractive to economics and finance practitioners.

4.3 MEASUREMENT OF RISK CONTAGION

There is no common agreement on what constitutes risk contagion or how to measure it, as can be seen in Buchholz and Tonzer (2013).

Yet, for the sake of this work, following the same authors, a contagion episode is characterized by a significant increase in volatility-adjusted correlations. Several other works use the same definition (FORBES; RIBOGON, 2002; BOYER ET AL., 2006; CAPORIN ET AL., 2013). But still, there is no uniform mathematical framework to do such calculations.

For example, Bianconi, Yoshino and Sousa (2013) analyze the risk contagion across volatility analysis using simple unconditional volatility measures, vector autoregressions, cointegration and conditional volatility and correlations among stock and bond returns. Among these different analyses, a “heat map” was built in order to show the evolution of the financial stress by an index, and when plotted in parallel, it is possible to check how the risk (“heat”) was spread across different economy sectors and different economies over time.

Bergmann et al. (2014) use a multivariate copula approach to test risk contagion effects over equities in different economies, subject to the choice of the copula method.

Buchholz and Tonzer (2013) use a DCC model – following Engle (2002) – and test whether there is an increase in the dynamic conditional correlation, by testing dummy variables and verifying their respective statistical significance, in a bivariate framework (using country pairs).

Chiang *et al.* (2007) test if correlations behave statistically differently over time using a DCC model to nine Asian daily stock-return data, adjusting the returns with an AR(1) coefficient and a global factor (returns over US Equities) and search for increases in the dynamic correlations during different periods, introducing arbitrary divisions in the sample, whether that subsample belongs to a crisis period or not.

Bratis, Laopodis and Kouretas (2015) evaluate and compare contagion/interdependence cross-country and cross-market using the DCC methodology, focusing the risk contagion across European Monetary Union, pointing out a difference between interdependence and contagion. To test for contagion, they employ one-sided tests for mean differences among subsamples, in order to see whether means of conditional correlations are different between tranquil and turbulent periods.

4.4 BRIEF REVIEW ON METHODOLOGIES FOR DYNAMIC CORRELATIONS

According to Engle (2002), several simple methods such as rolling historical correlations and exponential smoothing techniques are widely used, but more complex methods such as varieties of multivariate GARCH and Stochastic Volatility models, despite they have been extensively investigated by researchers, are used by a few sophisticated practitioners.

The simplest technique is the rolling correlation – and by far, the most popular, according to Engle (2002), which can be calculated as:

$$\rho_{1,2,t} = \frac{\sum_{s=t-n-1}^{t-1} r_{1,s} \cdot r_{2,s}}{\sqrt{\left(\sum_{s=t-n-1}^{t-1} r_{1,s}^2\right) \cdot \left(\sum_{s=t-n-1}^{t-1} r_{2,s}^2\right)}} \tag{67}$$

There is also the exponential smoother used by RiskMetrics, based on a parameter λ, which establishes more weight to the most recent data, given by:

$$\rho_{1,2,t} = \frac{\sum_{s=1}^{t-1} \lambda^{t-s-1} \cdot r_{1,s} \cdot r_{2,s}}{\sqrt{\left(\sum_{s=t-n-1}^{t-1} \lambda^{t-s-1} \cdot r_{1,s}^2\right) \cdot \left(\sum_{s=t-n-1}^{t-1} \lambda^{t-s-1} \cdot r_{2,s}^2\right)}} \tag{68}$$

According to Engle (2002), it is important to notice that there is no guidance from the data on how to choose λ. Hence, λ is imposed to be 0.94 for all assets / indices used, in order to ensure a positive definite correlation matrix.

In addition to these models, it is possible to estimate multivariate GARCH models, such as VECH, BEKK and DCC, which are going to be explained below.

In order to introduce them, it is worth explaining first the concept of scalar GARCH approach, which consists of a stochastic process of the form

$$x_t = \ln(P_t) - \ln(P_{t-1}) \tag{69}$$

$$\mu_t = E[x_t|\mathcal{F}_{t-1}], h_t = Var(x_t|\mathcal{F}_{t-1}) \tag{70}$$

where $\mathcal{F}_{t-1}$ denotes all available information until $(t-1)$. Hence, if one considers by hypothesis $\mu_t = 0$ and $h_t = (x_t^2|\mathcal{F}_{t-1})$, it is intuitive to build a (GARCH) model of the form:

$$x_t = \sqrt{h_t}\epsilon_t \tag{71}$$

$$h_t = \alpha_0 + \sum_{i=1}^{r} \alpha_i \cdot x_{t-i}^2 + \sum_{j=1}^{s} \beta_j \cdot h_{t-j} \tag{72}$$

Additionally, if one considers

$$\upsilon_t = x_t^2 - h_t \tag{73}$$

and rewrite the previous equation for h_t as:

$$x_t^2 = \alpha_0 + \sum_{i=1}^{q} (\alpha_i + \beta_i) \cdot x_{t-i}^2 + \upsilon_t + \sum_{j=1}^{s} \beta_j \cdot \upsilon_{t-j} \tag{74}$$

one obtains an ARMA(*q*,*s*) model for x_t^2, where υ_t is a martingale difference, as can be seen in Morettin and Tolói (2006).

Having presented how a simple scalar GARCH model works, it is possible to extend it to a multivariate context, where

$$X_t = \varepsilon_t = H_t^{-\frac{1}{2}} z_t \tag{75}$$

where X_t is a (*N* x 1) vector and z_t an i.i.d. random vector with the following characteristics:

$$E[z_t] = 0 \tag{76}$$

$$E[z_t z_t'] = I_N \tag{77}$$

$$z_t \sim G(0, I_N) \tag{78}$$

with G being a continuous density function and E denoting the expected value operator. Furthermore, let:

$$E_{t-1}[\varepsilon_t] = 0 \tag{79}$$

$$E_{t-1}[\varepsilon_t {\varepsilon_t}'] = H_t \tag{80}$$

$$E[\varepsilon_t {\varepsilon_t}'] = \Sigma \tag{81}$$

where E_{t-1} denotes the conditional expected value using all available information until $(t-1)$. Hence, it is possible to define a dynamic correlation matrix, where

$$R_t = Corr_{t-1}(\varepsilon_t) = D_t^{-\frac{1}{2}} H_t D_t^{-\frac{1}{2}} \tag{82}$$

with

$$D_t = diag\left(h_{11,t}, \dots, h_{NN,t}\right) \tag{83}$$

There are two alternative approaches, according to Rossi (2010):

- Models of H_t;
- Models of D_t and R_t.

The simplest model for a multivariate GARCH model is a vector ARMA process in squares and cross-products of the disturbances, following the same logic of the scalar GARCH process, which is basically a model of H_t, using the whole information set $\mathcal{F}_{t-1}$.

According to Engle and Kroner (1995), if one denotes the vector-halt operator by **vech**, which stacks the lower triangular elements of a *N* x *N* square matrix as an $\left[N \cdot \frac{(N+1)}{2}\right]$ x 1 vector, such as:

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \tag{84}$$

then

$$vech(A) = \begin{bmatrix} a_{11} \\ a_{21} \\ a_{22} \end{bmatrix} \tag{85}$$

So, if the conditional covariance matrix H_t is symmetric, then $vech(H_t)$ contains all elements in H_t.

Consequently, it is natural to write a multivariate extension of the GARCH model as:

$$vech(H_t) = W + \sum_{i=1}^{q} A_i \; vech(\varepsilon_{t-i}\varepsilon'_{t-i}) + \sum_{j=1}^{p} B_j \; vech\big(H_{t-j}\big) \tag{86}$$

As can be seen in Rossi (2010), the major issue with this kind of model is the number of parameters to be estimated (order $O(N^4)$). So, for example, if one considers $N = 5$, $p = q = 1$, the model would contain 465 parameters. So, restrictions must be imposed in this representation in order to decrease the number of parameters to be estimated, such as orthogonality assumptions and other kind of restrictions.

One of the most popular is the positive definiteness restriction, also shown in Engle and Kroner (1995). This restriction consists of considering the following model:

$$H_t = CC' + \sum_{k=1}^{K}\sum_{i=1}^{q} A_{ik}\varepsilon_{t-i}\varepsilon'_{t-i}A'_{ik} + \sum_{k=1}^{K}\sum_{i=1}^{p} B_{ik}H_{t-i}B'_{ik} \tag{87}$$

where the intercept matrix W is decomposed into CC', where C is a lower triangular matrix, being a positive semidefinite matrix.

If, for the sake of simplicity of this exposition, one assumes $K = 1$ and a GARCH(1,1), then:

$$H_t = CC' + A_1\varepsilon_{t-1}\varepsilon'_{t-1}A'_1 + B_1H_{t-1}B'_1 \tag{88}$$

Consequently, if the diagonal elements in C are restricted to be positive and a_{11} and b_{11} are also restricted to be positive, there is no other C, A_1 and B_1 with an equivalent representation. So, the restrictions eliminate all other observationally equivalent structures.

On the opposite way, as pointed before, there is the DCC methodology, which is basically a model of D_t and R_t. The main idea of the DCC model is to estimate in a two-step approach the conditional correlation based on the univariate GARCH models for each single variable, and a model of R_t, which consists of a filter derived according to Engle (2002).

So, basically, the author derives a filter where:

$$Q_t = \Sigma\,(1 - \alpha - \beta) + \alpha\, H_t + \beta\, Q_{t-1} \tag{89}$$

with Σ being the unconditional covariance matrix (i.e. long-run covariance matrix). And, hence, if Q_t is a filtered covariance matrix, with parameters α and β estimated in a second step and H_t calculated using the standardized residuals from the univariate GARCH models estimated in a first step, then it makes sense normalizing the Q_t in terms of the conditional filtered variances in the main diagonal, in order to obtain a dynamic conditional correlation, such as:

$$R_t = diag(Q_t)^{-1} \cdot Q_t \cdot diag(Q_t)^{-1} \tag{90}$$

calculated at each time step.

Each model has its particular advantages and weaknesses. BEKK and VECH are harder to estimate but do not require any arbitral choice of parameters – they are estimated by maximum likelihood procedures. The exponentially adjusted model (also known as RiskMetrics procedure) and the rolling correlation model are easier to estimate, but tend to generate biases as consequence of the timeframes and smoothing parameter choices. The weakest point in the DCC methodology, as pointed out by Caporin and McAleer (2013), is the absence of any derivation of the filter and its mathematical properties; and the lack of any demonstration of the asymptotic properties of the estimated parameters as well.

But still, despite this lack of theoretical foundations for DCC, among these different existing methods for multivariate conditional covariance, Engle (2002) points out that the DCC model performs better than all other traditional techniques in terms of mean absolute error of correlation estimates in a computational exercise, where the data generation process is known.

Hence, for the purpose of this chapter, the DCC model is going to be used as a benchmark, in order to compare the risk contagion based on correlation indices against a new correlation index developed and discussed here.

4.5 A DYNAMIC CORRELATION INDEX BASED ON LONG MEMORY FILTERS

The motivation to build a dynamic correlation index based on long memory filters comes basically from the following major points, discussed so far in Chapter 3 and in the previous section:

- Long Memory processes seem to be generated by idiosyncrasies of real market structures/economic specificities – so a good part of what is being observed is contaminated by the way the agents overreact due to specific market structures (such as laws of motion, spatial complexity and composition of agents). Filtering out these market structures may turn comparable shocks across different economies;
- Long Memory filters seem to be a good tool to identify what are the true innovations (genuine shocks), and what is being a cumulative result of the autocorrelation structure present in each real market/economy – complementing the first point above;

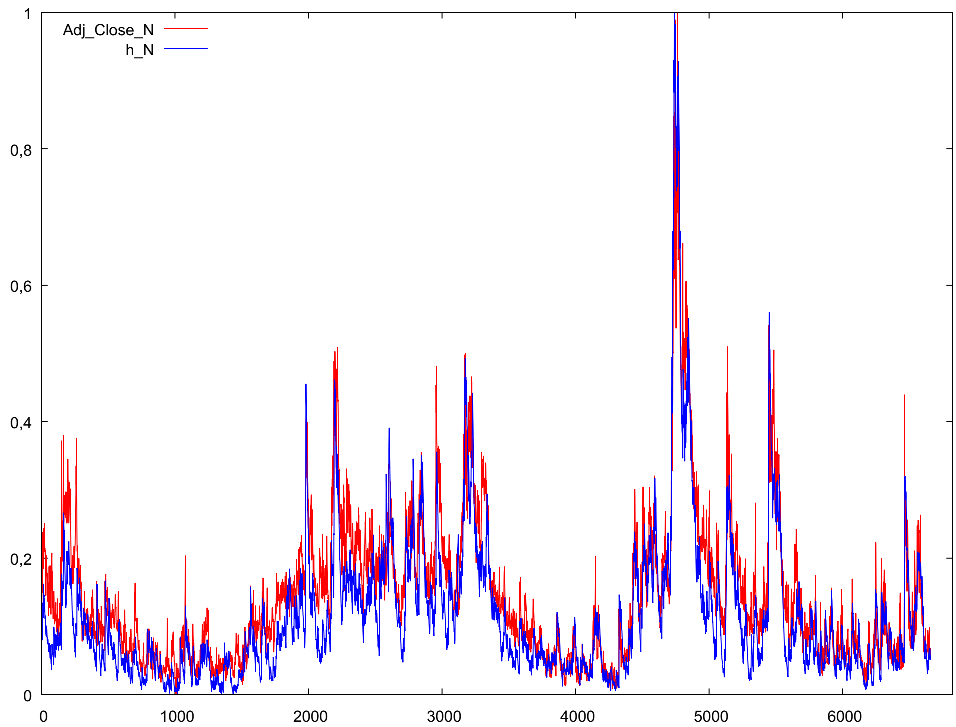
- A set of issues were found while investigating what have been done so far in terms of risk contagion investigation using DCC, which basically requires from the researcher previous *ex-ante* knowledge about what he is willing to identify – so having a filter which does not require any previous knowledge seems to be a very interesting feature to analyze real-time data.

Keeping that in mind, the first important thing is to recognize that univariate GARCH models are attractive estimators of the conditional variance of equity indices. Moreover, they are able to replicate the main stylized facts observed in risk measures such as the VIX (Implied Volatility over S&P500 options), as can be observed in Figure 4.1.

It is possible to see in Figure 4.1 that a significant part of the shocks observed in the VIX original series (in red) can also be seen in the conditional variance extracted by using a univariate GARCH(1,1) model. So, developing a correlation based index which is still based on GARCH models is a good approach.

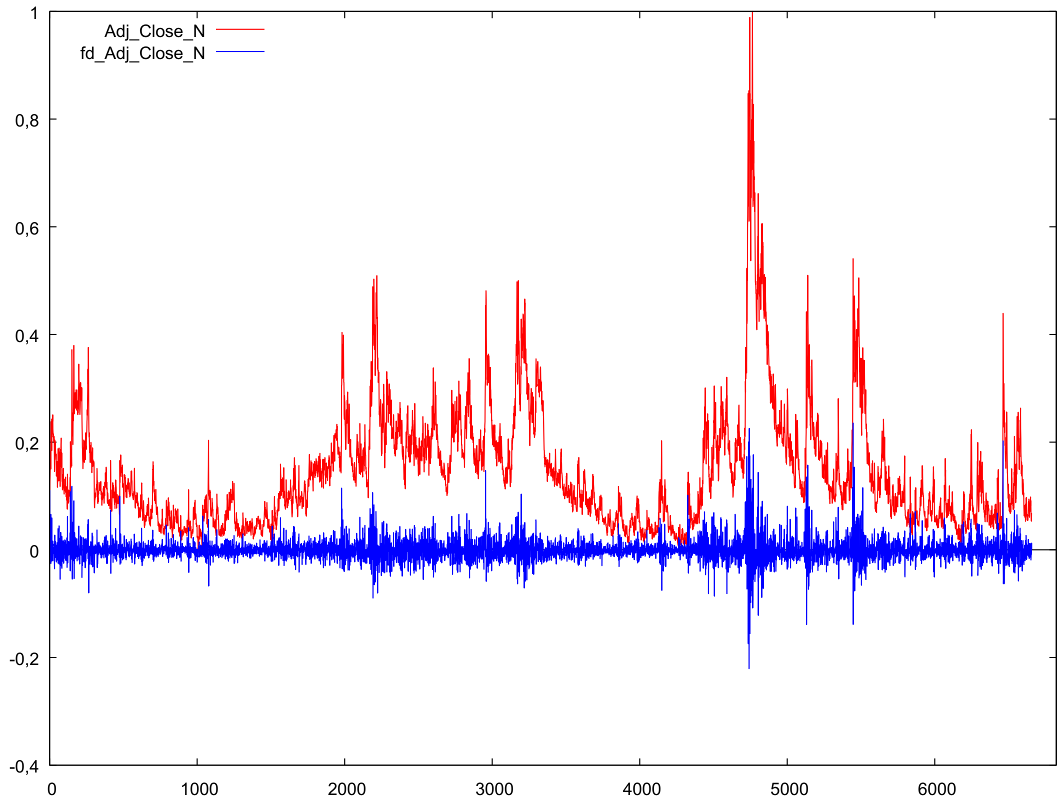
In addition to that, when one applies a long-memory filter to these conditional variances extracted by using GARCH models, it is possible to verify that a yet more interesting feature arises.

Figure 4.1 – VIX *versus* GARCH estimates of equity volatility at trading days.



According to Figure 4.2, whenever there is a peak on the original series (in red), there are also peaks in the filtered series (in blue). Additionally, all level variations that result from the cumulative process inherent to the autocorrelation function (long memory components) are filtered out from the original series (in red). Thus, two major points must be discussed.

Figure 4.2 – Conditional volatility estimate of equity volatility at trading days versus filtered series.



The first important point to be addressed is about all specificities of each economy, which may generate long range dependency, in conjunction with large shocks – as discussed in Chapter 3. So, this autocorrelation issue is filtered out, enabling the researcher to visualize the shocks of interest, without possible overreaction biases.

The second important feature becomes from the fact that, if a filtered series contains only short-term fluctuations, the peaks in the filtered series will naturally lead to the identification of stressing moments, discarding any further requirement of imposing *ex-ante* knowledge of the data due to the presence of long-term serial correlation. Hence, there is no need to impose, *a priori*, dummy variables to check whether there is an increase in the correlation between the volatilities.

When taking into account all these aspects, this approach seems to be attractive because its resulting index should answer two questions: whether shocks in risk aversion of the economic agents occur in a pair of countries; and how strong is this contagion – independently from the overreactions in each economy caused by market idiosyncrasies. Thus, the proposed methodology consists of the following steps:

1. Calculate the univariate conditional standard deviations (variance) according to scalar GARCH models;
2. Calculate the fractional difference parameters for each conditional standard deviation series, according to the estimators discussed in Chapters 2 and 3;
3. Filter out all the long range dependency using the fractional difference operator as a filter for each conditional standard deviation series;
4. Calculate rolling correlations between pairs of countries using timeframes, according to the interest of the researcher.

In order to simplify any reference to this correlation model, this is going to be referred as DCLMF (Dynamic Correlation based on Long Memory Filters), for the purpose of this work.

4.6 ESTIMATION OF THE CORRELATION MODELS

In order to begin with the discussions related to the performance comparison of the proposed correlation index in the previous section, it is worth saying that three different time series were used here, aiming the discussing of risk contagion:

- Shanghai Stock Exchange Index Data;
- Standard & Poor's 500 Index Data;
- Bovespa Stock Exchange Index Data.

All of them ranging from April 27th of 1993 until May 13th of 2016, including only all common trading days, in order to estimate risk contagion. So, basically, the dataset comprises more than five thousand trading days, composing a large amount of available data to estimate long range dependency properties.

Two models were estimated for comparison purposes, as described before: a standard DCC model and a DCLMF model; and two pairs of bivariate analysis were carried out:

- Risk contagion from the U.S. to China;
- Risk contagion from the U.S. to Brazil.

Hence, following the methodology in Engle (2002) and using the rugarch package for R Statistics, two DCC estimates were obtained:

- A first DCC estimate for measuring risk contagion/market integration between U.S. and Brazil (see Table 4.1);
- A second DCC estimate for measuring risk contagion/market integration between U.S. and China (see Table 4.2).

As one can notice in Tables 4.1 and 4.2, the univariate GARCH estimates are identical for U.S. (as expected). In addition to that, it is possible to verify that the weight of the long-run covariance matrix is very low in both models, due to the high values of the sum of α and β coefficients.

Moreover, in both models the values of beta seem to be high, indicating that the filter is smoothing the correlation values, by imposing a high autoregressive coefficient and incorporating a small amount of covariance caused by contemporaneous innovations, when taken into account the formulation of Q_t.

Table 4.1 – DCC estimates for the United States against Brazil.

Distribution: mvnorm	Model: DCC(1,1)
Number of parameters: 15	[VAR GARCH DCC UncQ]: [0+12+2+1]
Number of series: 2	Number of observations: 5461
Log-likelihood: 32422.87	Av. log-likelihood: 5.94
Elapsed time: 9.708487	

(continues)

Table 4.1

(continued)

Optimal parameters

Parameter	**Estimate**	**Std. error**	***t*-value**	**Pr (> |*t*|)**
[ld_SP500].mu	0.000610	0.000139	4.40333	0.000011
[ld_SP500].ar1	0.860183	0.122137	7.024279	0.000000
[ld_SP500].ma1	-0.890330	0.109685	-8.11718	0.000000
[ld_SP500].omega	0.000001	0.000005	0.25852	0.796007
[ld_SP500].alpha1	0.091038	0.083125	1.09519	0.273434
[ld_SP500].beta1	0.901011	0.082574	10.91157	0.000000
[ld_IBOV].mu	0.001089	0.000245	4.45312	0.000008
[ld_IBOV].ar1	-0.503998	0.291468	-1.72917	0.083779
[ld_IBOV].ma1	0.536801	0.284605	1.88613	0.059278
[ld_IBOV].omega	0.000008	0.000007	1.08725	0.276925
[ld_IBOV].alpha1	0.096893	0.015060	6.43394	0.000000
[ld_IBOV].beta1	0.889174	0.014363	61.90622	0.000000
[Joint]dcca1	0.028868	0.004841	5.96365	0.000000
[Joint]dccb1	0.968054	0.005769	167.81081	0.000000

Information criteria

Criterion	**Akaike**	**Bayes**	**Shibata**	**Hannan-Quinn**
Value	-11.869	-11.851	-11.869	-11.863

Table 4.2 – DCC estimates for the United States against China.

Distribution: mvnorm	Model: DCC(1,1)
Number of parameters: 15	[VAR GARCH DCC UncQ]: [0+12+2+1]
Number of series: 2	Number of observations: 5461
Log-likelihood: 32013.49	Av. log-likelihood: 5.86
Elapsed time: 8.771168	

Optimal parameters

Parameter	Estimate	Std. error	*t*-value	Pr (> |*t*|)
[Id_SP500].mu	0.000610	0.000139	4.38886	0.000011
[Id_SP500].ar1	0.860183	0.122158	7.04158	0.000000
[Id_SP500].ma1	-0.890330	0.109700	-8.11606	0.000000
[Id_SP500].omega	0.000001	0.000005	0.25733	0.796927
[Id_SP500].alpha1	0.091038	0.083500	1.09027	0.275595
[Id_SP500].beta1	0.901011	0.082964	10.86030	0.000000
[Id_SSE].mu	0.000314	0.000220	1.42831	0.153202
[Id_SSE].ar1	-0.122529	0.387699	-0.31604	0.751971
[Id_SSE].ma1	0.117182	0.387503	0.30240	0.762344
[Id_SSE].omega	0.000002	0.000004	0.59481	0.551969
[Id_SSE].alpha1	0.071776	0.031185	2.30162	0.021357
[Id_SSE].beta1	0.927224	0.032100	28.88573	0.000000
[Joint]dcca1	0.05314	0.006058	0.87720	0.380378
[Joint]dccb1	0.990693	0.014132	70.10521	0.000000

Information criteria

Criterion	Akaike	Bayes	Shibata	Hannan-Quinn
Value	-11.719	-11.701	-11.719	-11.713

After the estimation of the univariate GARCH processes for each stock index and having estimated the DCC models for each pair of countries, it is also necessary to estimate the fractional difference parameters, in order to apply the long memory filter – for the DCLMF estimate.

Hence, GPH (1983) and Künsch (1987) estimates were calculated, which can be followed in Tables 4.3, 4.4 and 4.5.

Table 4.3 – Fractional difference parameter estimates for conditional variance (Bovespa index).

```
h_IBOV, T = 5461

Local Whittle Estimator (m = 173)
  Estimated degree of integration = 0.504972 (0.0380143)
  test statistic: z = 13.2837, with p-value 0.0000

GPH test (m = 173)
  Estimated degree of integration = 0.530032 (0.0519439)
  test statistic: t(171) = 10.2039, with p-value 0.0000
```

Table 4.4 – Fractional difference parameter estimates for conditional variance (S&P 500 index).

```
h_SP500, T = 5461

Local Whittle Estimator (m = 173)
  Estimated degree of integration = 0.819189 (0.0380143)
  test statistic: z = 21.5495, with p-value 0.0000

GPH test (m = 173)
  Estimated degree of integration = 0.806639 (0.0456596)
  test statistic: t(171) = 17.6664, with p-value 0.0000
```

Table 4.5 – Fractional difference parameter estimates for conditional variance (Shangai Stock Exchange index).

h_SSE, T = 5461
Local Whittle Estimator (m = 173)
Estimated degree of integration = 0.543186 (0.0380143)
test statistic: z = 14.289, with p-value 0.0000
GPH test (m = 173)
Estimated degree of integration = 0.575488 (0.0428547)
test statistic: t(171) = 13.4288, with p-value 0.0000

As there was no significant difference between GPH (1983) and Künsch (1987) estimates for the three time series, averages between the estimates for each series were calculated, which were used to compute the fractionally differenced (filtered) series, as can be seen in Figures 4.3, 4.4 and 4.5.

Figure 4.3 – Conditional variance estimate of equities at trading days versus filtered series (S&P 500).

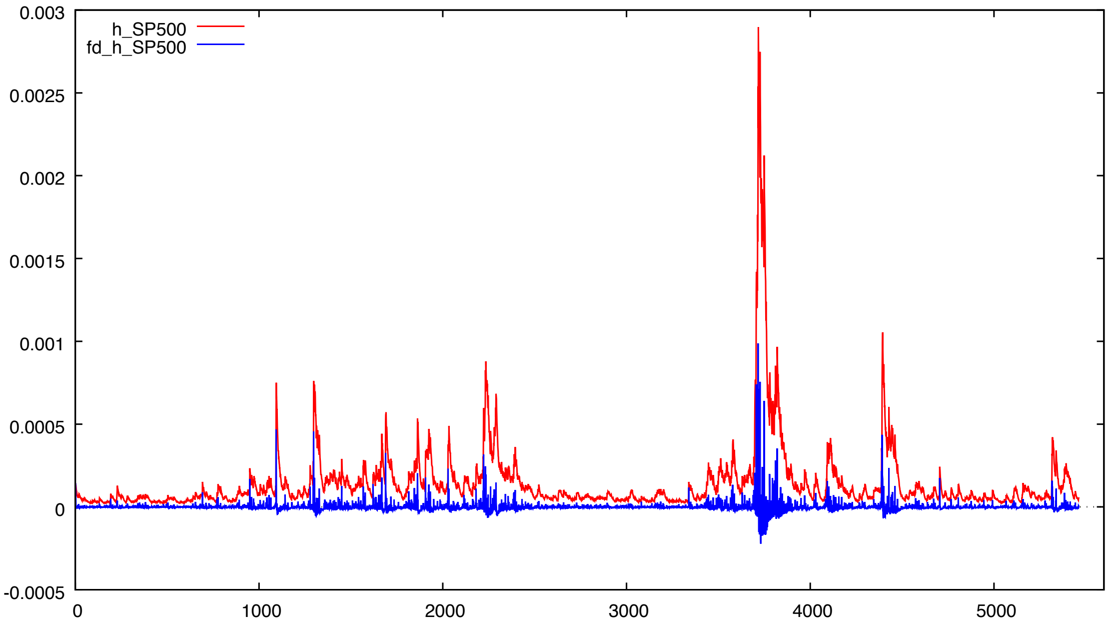


Figure 4.4 – Conditional variance estimate of equities at trading days versus filtered series (Bovespa index).

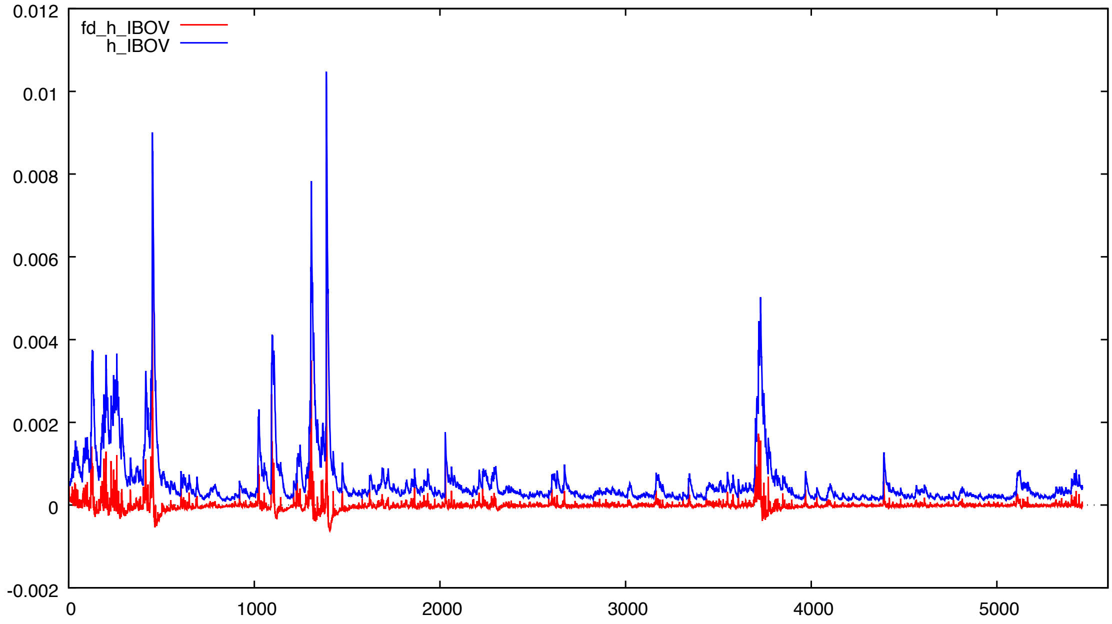
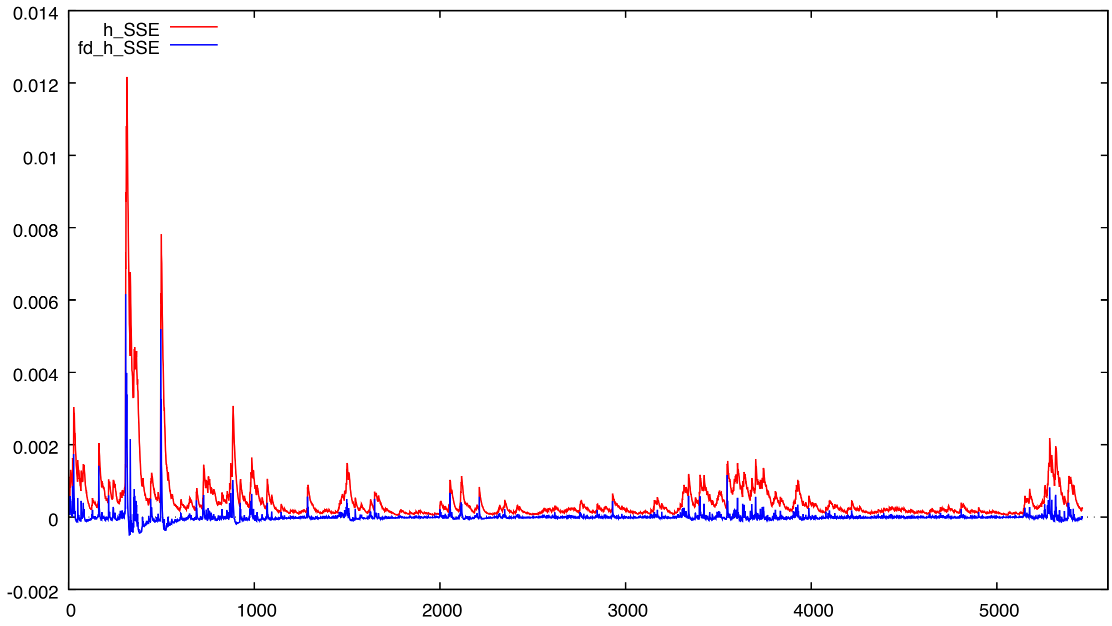


Figure 4.5 – Conditional variance estimate of equities at trading days versus filtered series (Shangai Stock Exchange index).



According to Figures 4.3, 4.4 and 4.5, it is possible to verify that the filtered series are stationary series that contain the peaks whenever there are over-reactions in the original levels. Consequently, these peaks can point out whenever there are stressing moments, without the contagion of the subsequent data points related to the autocorrelation structure inherent to each market/economy. So, when calculating correlation between these filtered series, it is possible to capture how much turbulence is passed from a market to another one, without the interference of the way the agents cumulatively react to them.

After the calculation of these filtered series, rolling correlations are calculated at different time frames and comparisons between this method and the DCC method are carried out, in the next section.

4.7 COMPARING THE DYNAMIC CORRELATION BASED ON LONG MEMORY FILTERS (DCLMF) AND THE DYNAMIC CONDITIONAL CORRELATION (DCC)

The comparison between the dynamic correlations calculated using different methodologies is not trivial, since there is no observed component to serve as a parameter to measure the goodness of the fit.

Choosing the DCC as the standard benchmark model seemed to be a good way, since it became the standard tool in the risk contagion analysis, as discussed in the literature review, and, moreover, it seems to be the best model to capture the dynamic correlation structure, when the data generation process is known – in this case, there is a parameter to measure the goodness of fit, which can be done by means of computational simulation experiments.

Hence, in order to establish which estimative is the best, actually a comparison must be made between the results achieved with these models against what has been found so far in the existing literature.

The first point to be addressed is the timeframe to be chosen in terms of the DCLMF correlation index. In Figures 4.6 and 4.7 it is possible to see how the index behaves according to the timeframe.

It is possible to check in Figures 4.6 and 4.7 that for both estimates – S&P 500 *versus* Shanghai Stock Exchange Index; S&P 500 *versus* Bovespa Index, respectively – that high frequencies turn out to result in noisy correlations. On the other hand, as the frequencies become lower (larger timeframes), the estimates become smoother. Consequently, medium frequencies should be used. In this case, arbitrarily, it was chosen timeframes of 252 trading days.

Having chosen the timeframes, the next exercise consists of comparing the DCC estimates against the DCLMF estimates. To accomplish that, plots of both estimates for both pairs of countries were plotted in Figures 4.8 and 4.9.

According to Figure 4.8, it is possible to see that both estimates (DCC and DCLMF) share common information – there is a correlation of 0.31 between both estimates. Moreover, both estimates provide proofs that, in general, China is insulated from the US, despite the fact that there are evidences of risk transmission between these markets – this finding is in line with Bianconi, Yoshino and Sousa (2013) and Syriopoulos, Makram and Boubaker (2015). The DCLMF points to an average correlation of 0.06 and the DCC has an average of 0.02 (in the whole period) – supporting the previous statement.

Figure 4.6 – DCLMF correlation estimates in different timeframes (Shangai Stock Exchange index *versus* S&P 500).

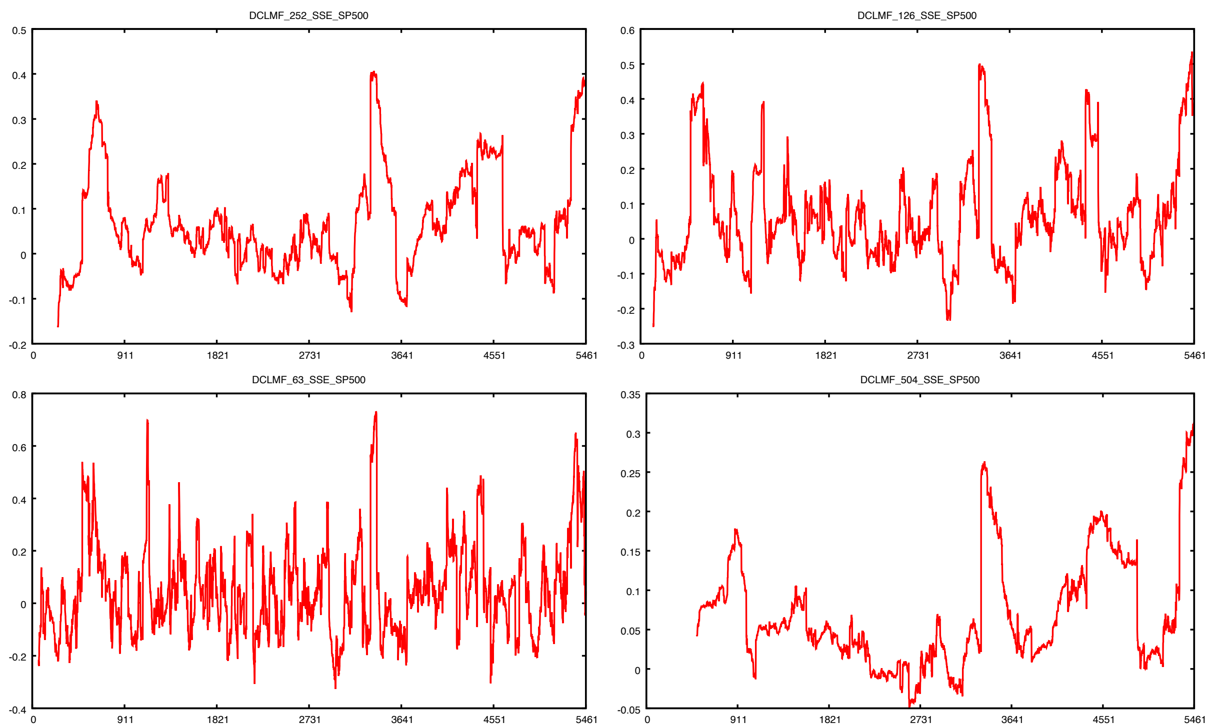


Figure 4.7 – DCLMF correlation estimates in different timeframes (Bovespa index *versus* S&P 500).

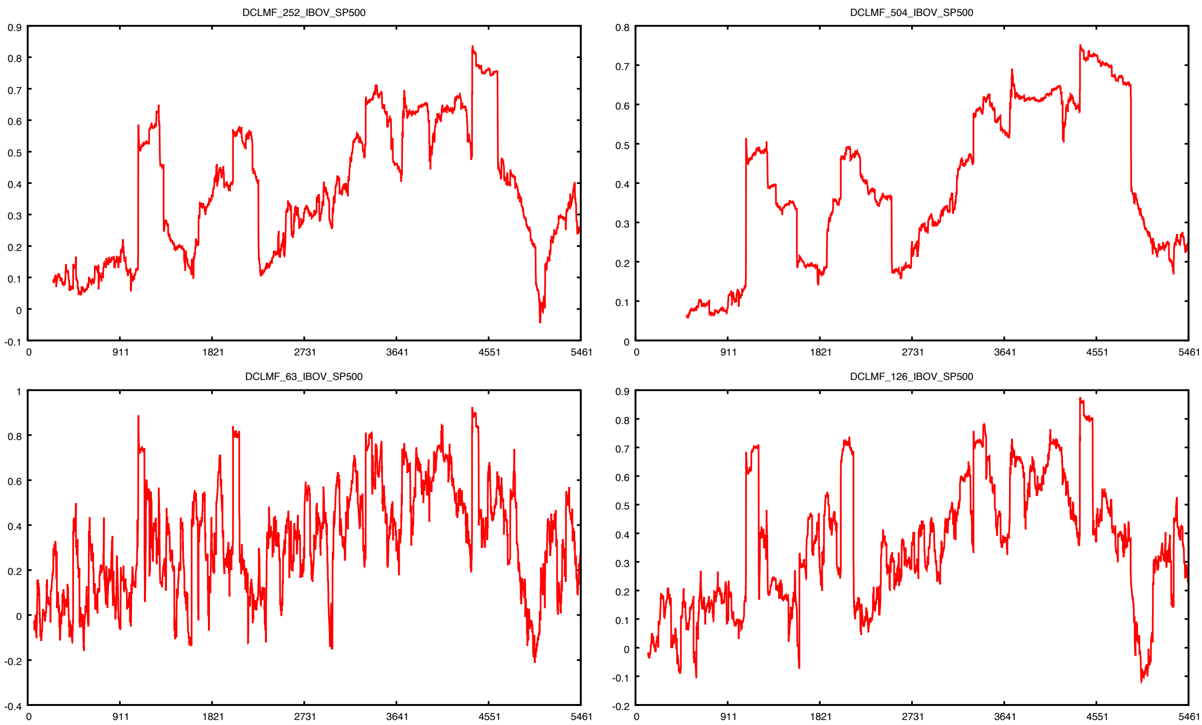


Figure 4.8 – Dynamic correlation estimates (Shangai Stock Exchange index *versus* S&P 500).

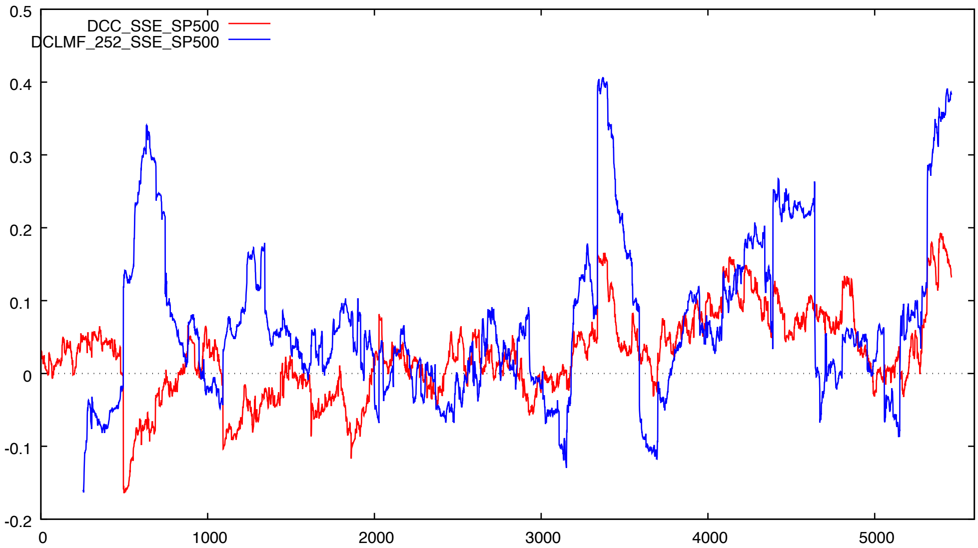
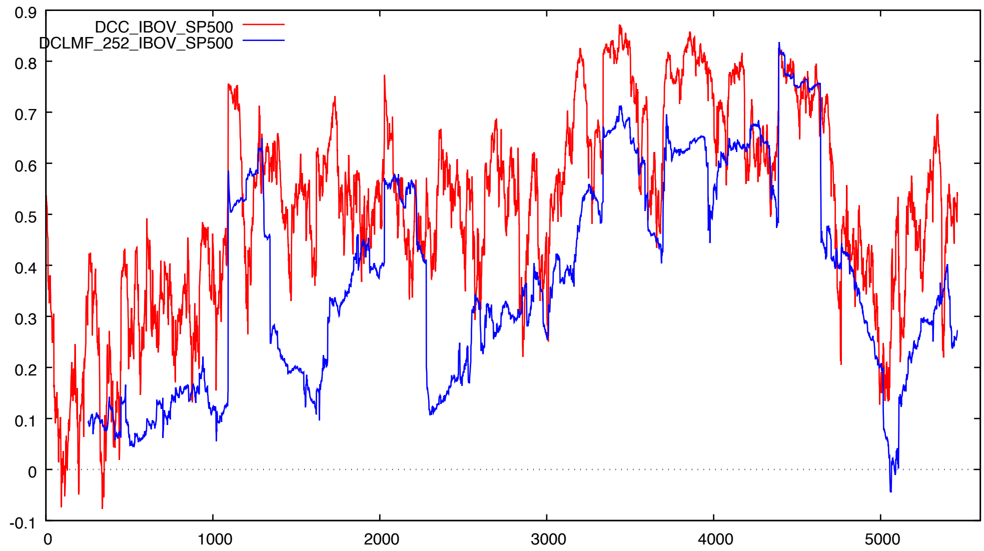


Figure 4.9 – Dynamic correlation estimates (Bovespa index *versus* S&P 500).



But it is also worth mentioning that in several periods, the correlation estimates seem to point to opposite directions. This is a resulting effect from the fact that while DCC recovers the correlation by means of the standardized residuals (which can be positive or negative), the DCLMF uses the fractionally differenced conditional variance series – which are always positive. Consequently, it seems that while DCC results point to economic agents taking opposite positions in the market in some of the stressing moments, the DCLMF results focus only in the risk transmission – which is something more desirable.

According to Figure 4.9, DCC and DCLMF also share common information – there is a 0.76 correlation between both estimates. But, differently from the comparison between China and U.S., Brazilian and U.S. markets seem to be reasonably integrated, with an average correlation of 0.3751 (DCLMF) and 0.5175 (DCC) – again, in line with the findings of Bianconi, Yoshino and Sousa (2013) and Bergmann et al. (2014). In addition to that, DCLMF seems to allow faster level shifts in comparison to the DCC (when economies are integrated). Again, this is a result of the fact that while DCC uses standardized residuals to recover the dynamic correlation, DCLMF uses the fractionally differenced conditional variance series.

Furthermore, it is worth remembering that the DCC correlation estimates are obtained based on an assumed law of motion for the Matrix Covariance, which is then decomposed into the correlation matrix. Given that, it is not expected that fast level shifts should occur in the correlation estimates obtained by the DCC methodology.

Additionally, a more detailed comparison between the volatilities and the dynamic correlation indices is carried out, in Figures 4.10 and 4.11, for China and U.S. markets.

Figure 4.10 – Dynamic correlation estimates and conditional variances (Shangai Stock Exchange index *versus* S&P 500) between 400th and 1200th days.

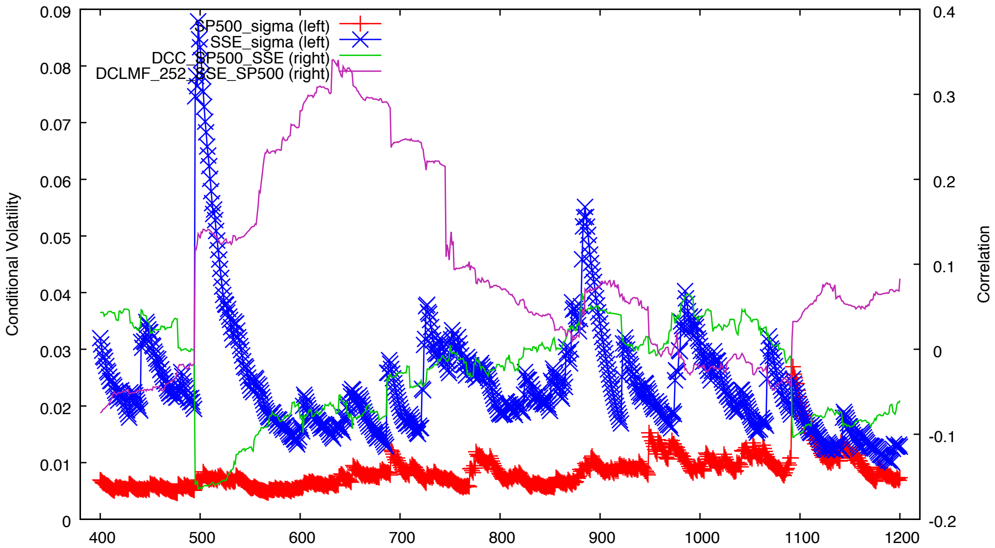
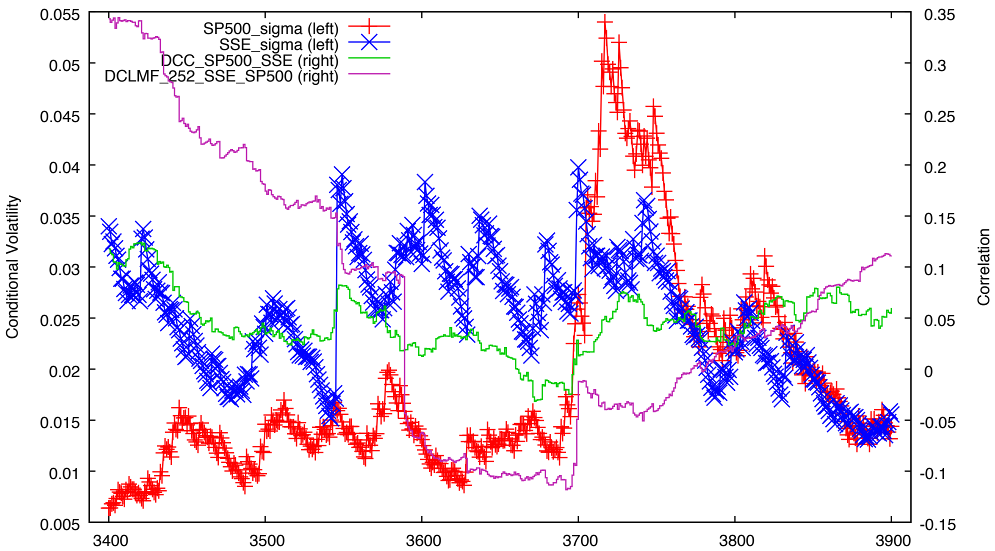


Figure 4.11 – Dynamic correlation estimates and conditional variances (Shangai Stock Exchange index *versus* S&P 500) between 3400th and 3900th days.

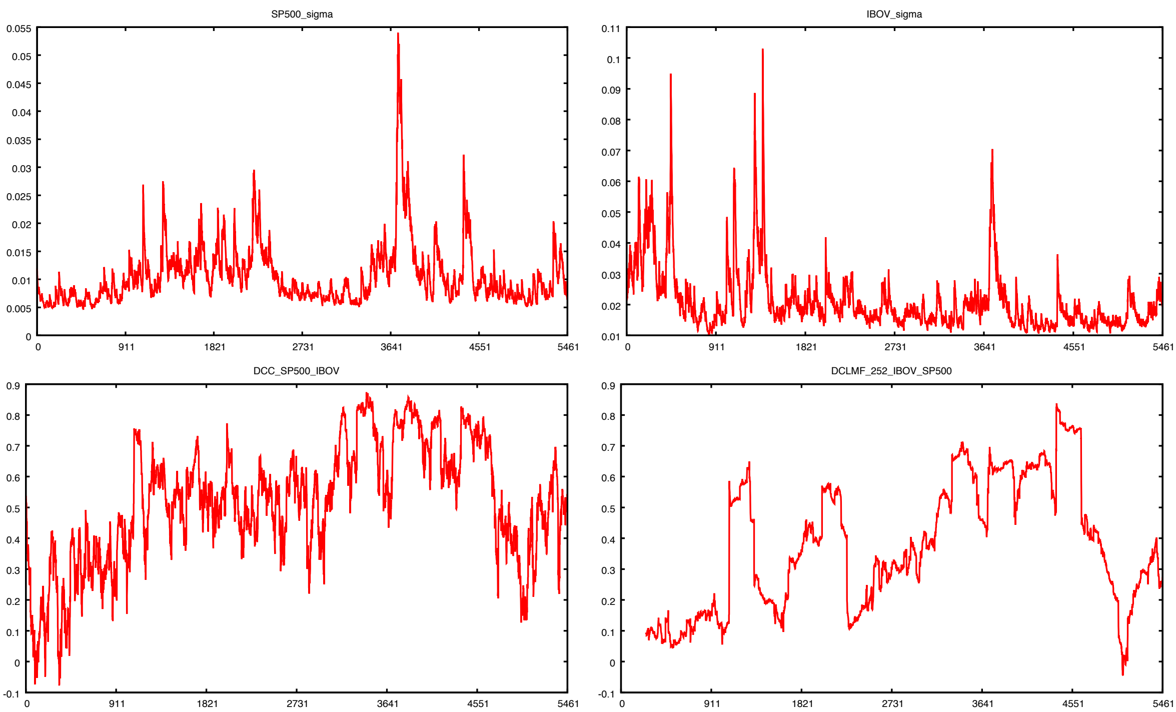


As told before, in general, DCC and DCLMF point out a very low degree of average risk contagion between China and U.S. markets. But, when a deeper look is taken, one can notice that nearby the instant t = 3599, there is an increase in the conditional variance in the Chinese market, while there is a decrease in the S&P 500 conditional variance. Given that, the DCLMF correlation rapidly decreases to a negative correlation while DCC is still positive. The same situation occurs nearby the instant t = 500.

In specific events, it seems that DCLMF better captures the risk contagion while DCC probably is capturing the effects of the standardized residuals – which does not necessarily hold a direct relationship with risk contagion.

The same exercise was repeated for the Brazilian and U.S. Markets in Figure 4.12. From this figure, it is possible to see that DCLMF is less noisy than DCC estimates. Moreover, in specific regions, such as t = 2731, where DCC suggests a stronger market integration (correlation), DCLMF suggests no significant risk contagion – which are substantively different. Additionally, a large part of (if not all) the peaks that occur in the conditional volatilities seem to be captured by the DCLMF index, in this specific case. In this case it seems that both methodologies complement each other.

Figure 4.12 – Dynamic correlation estimates and conditional volatilities (Bovespa index *versus* S&P 500).



Analyzing all the previous empirical points here discussed, it seems that DCLMF seems to better capture stressing moments rather than the standard DCC methodology, which, in average, seems to be a better metric of integration between markets. The composite behavior between a rolling correlation based on true component short memory components (filtered using long memory filters) seems to provide smooth estimates enough to detect risk contagion episodes, but with enough flexibility to capture level shifts, which characterizes sudden risk peaks.

Consequently, if DCLMF seems to be able to capture risk contagion episodes, it is a good opportunity to use it as a practical tool to detect the 2008 crisis effects and the effects related to the Quantitative Easing episode, in order to measure what are the transmission effects of risk aversion in real economy sectors. Such exercise is carried out in the next chapter of this PhD Thesis.

5 WHAT WOULD HAVE HAPPENED TO BRAZIL WITHOUT SUBPRIME CRISIS CONTAGION: AN APPLICATION OF LONG MEMORY FILTERS

5.1 PROLOGUE OF THE CHAPTER

The Subprime Crisis and the Quantitative Easing shocks caused a “Tsunami” in the world’s economy. And this was the major excuse found in Brazil, during the last 5 years, for its huge economic failure. To what extent did these crises contribute for this economic failure? How does long-range dependent phenomena can help one to answer that question? Why answering these questions is so important, given the current Brazilian political scenario?

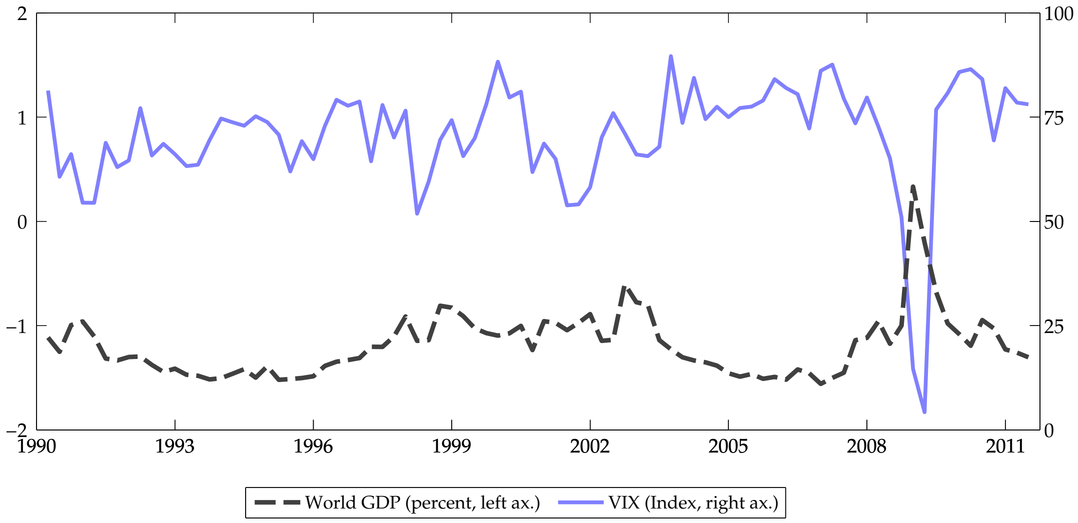
5.2 RISK CONTAGION FROM FINANCIAL SECTORS TO REAL SECTORS: A THEORETICAL BACKGROUND AND THE IMPORTANCE OF THE BRAZILIAN CASE

Understanding volatility goes beyond understanding the impacts of uncertainty over bond markets, capital markets and stock markets. Increasing attention has been paid to this question, due to the effects of the Subprime Crisis in the real sectors of economies around the world, as can be seen in Fornari and Mele (2009). Now, it seems people are starting to realize that Wall Street is not disconnected from the rest of the real sectors, as they thought it would be.

It is clear in Figure 5.1, which reproduces a proxy of the World GDP growth composed by 33 advanced and developing economies (which covers more than 90% of the world GDP, using PPP-GDP weights) versus the VIX index, that an increase in the uncertainty in financial markets (measured by VIX) is able to affect the real sectors of the economy.

Aiming a better understanding of this relationship, first, it is worth mentioning that economic agents are risk averse, at least to some extent. Some of them seek more risk aiming larger returns; others seek less risk for safer returns. Consequently, conditional volatility (which is a risk measure) is fundamental for investment decisions, given that economic agents, in general, want to maximize their respective profits minimizing their risk exposure, as discussed in the previous chapter (HULL, 2015).

Figure 5.1 – Quarterly world GDP growth and VIX index.



Source: Cesa-Bianchi, Pesarani and Rebucci (2014).

Then, it is very helpful examining the relationships between the markets that are directly subject to volatility and uncertainty shocks, the economic agents that react to these shocks and the real sectors. There are three main transmission mechanisms that interconnect them:

From the viewpoint of the firm, irreversible investments are subject to postponement due to uncertainties of demand and funding conditions – as can be seen in Bloom (2009). Reducing investments leads to a fall in economic activity, which may also lead to reductions in labor income and dividends.

From the viewpoint of households, uncertainty in the labor income and dividends – both function of economic activity – induces to precautionary savings, which leads to a consumption reduction, according to Leland (1968) and Kimball (1990). Loans also play a major factor in consumption. Reducing consumption also leads to a fall in the economic activity.

The third transmission channel is *via* Risk-premium shocks, which can be basically thought as a demand for safe and liquid assets, such as short-term US treasury securities, as can be seen in Fisher (2014). To better understand this transmission channel, it is worth remembering that the financial system funds flow from savers to those who have a shortage of funds, either by means of market-based financing (such as debt bonds) or bank-based finance. Consequently, if there are uncertainty shocks over these markets (which may be caused by real sectors, as in the Subprime Crisis), economic agents may not find attractive investing their surplus in risky assets, such as debt bonds or other securities. Hence, they increase their risk premium in order to take this additional risk, which leads to increases in the loan/funding rates. And those who rely on others' money to make new investments or simply consume might not be able to do that anymore. Thus, this is how (in a very simplified manner) uncertainty shocks lead to Risk-premium shocks, which impact directly on the real sectors, as can be seen in Winkler (1998). For a in depth view of such effects, see Smets and Wouters (2007).

That said, it is clear that, undoubtedly, the Subprime Crisis that occurred in the U.S. impacted real sectors worldwide, given the fact that there were significant risk contagion episodes between several financial sectors and U.S. financial sectors, which subsequently induced a worldwide recession.

Nonetheless, this topic has specially received more attention in Brazil. As can be seen in Langevin (2014), the economic recession that is being observed in the country has led to a contraction of 3.8 percent in terms of GDP growth, an inflation surpassing 10 percent and an increasing unemployment, which induced large popular mobilizations, with the singular refrain that the president should be impeached.

During this process, the president has continuously blamed the international crisis as the main cause for the observed recession in Brazil – which can be easily verified in all her public speeches. Hence, it turns out to be very attractive to analyze the extension of the Subprime Crisis effects in the Brazilian GDP. Was the president right? Or wrong?

According to Barbosa and Pessôa (2014), two major causes were identified and related to the observed recession in Brazil: the initial subprime shocks (in 2008) and the set of "Quantitative Easing" policies (which started significantly in August/2010).

The initial Subprime shocks (occurred in 2008) were basically translated into a demand drop in the developed economies, which were a significant driver for a high demand for commodities. And these demands were the major force behind the internal demand in Brazil.

Additionally, the Quantitative Easing was, basically, a set of policies of buying long-term assets, in order to reduce long-term interest rates, aiming a faster economic recovery, given the fact that the short-term interest rates were already near to zero boundary. Thus, lowering rates in the U.S. provokes the appreciation of exchange rates (due to the differential between Brazilian Interest Rates and U.S. Interest Rates, which play a central role in the demand/supply of U.S. dollars in Brazil), which lowers the competitiveness of Brazilian products.

Considering all these aspects, the majority of insights provided by Barbosa and Pessôa (2014) and other works are purely qualitative. The most quantitative approach was shown in Kanczuk (2013, 2014), which uses a DSGE model to quantify various shocks and analyze different policies during the past decade. Unfortunately, the main focus of both of these works is not the comprehension of the Subprime Crisis and Quantitative Easing effects. This is the gap that this present chapter aims to fill, by means of employing a pure time series practical approach, in conjunction with the theoretical insights developed in the previous chapters.

5.3 MEASURING THE CRISIS EFFECTS IN THE BRAZILIAN GDP: METHODOLOGY

In order to measure the discussed effects on the Brazilian GDP, the methodology here developed consists of five steps:

- Assessing the risk contagion episodes by employing the DCLMF correlation discussed in Chapter 4;
- Filtering out the risk contagion by means of an artificial simulated ARMA process over the fractionally differenced volatility series – in order to simulate what would be the “natural volatility levels” of the Brazilian market, without the risk contagion;
- Estimate a real monthly GDP series for the Brazilian Economy, based on the nominal monthly GDP series available from the Brazilian Central Bank;
- Estimate an elasticity between uncertainty (using equity volatility from Ibovespa as proxy) and the real GDP estimates;
- Calculate a counterfactual GDP evolution, based on the volatility free of risk contagion episodes, to assess the costs of the crisis.

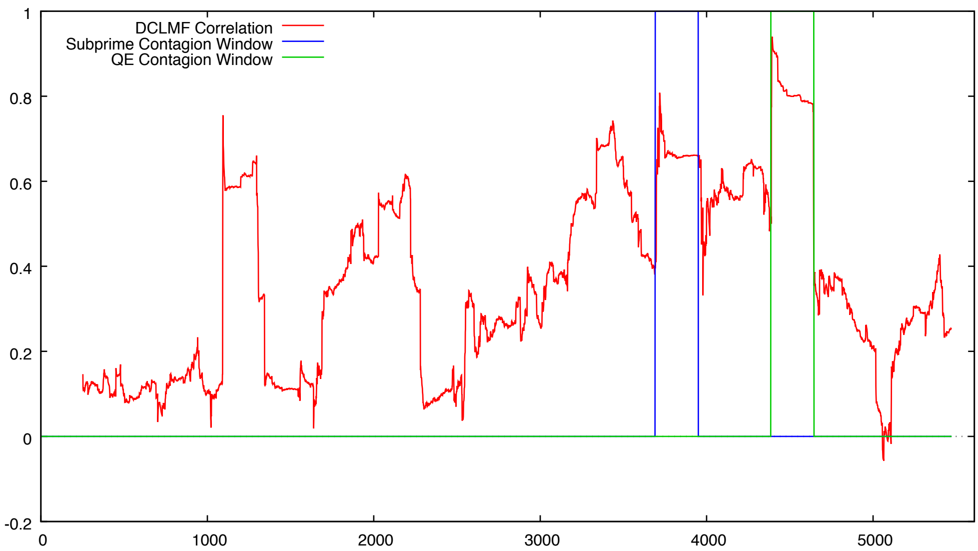
Each of these five steps is explained in details in the next sections, with their respective results. But before a detailed presentation of each of these steps, it is worth mentioning that the present approach assumes that both volatility and real economic activity are affected by the same set of unobserved common factors. This assumption seems to be plausible, given the existing interconnections between financial sectors and real sectors through the transmitting channels discussed in the previous section. Nevertheless, other works such as Cesa-Bianchi, Pesaran and Rebucci (2014) also make use of the same assumptions.

5.3.1 Step 1: Assessing Subprime Crisis Risk Contagion Episodes in Brazil through the DCLMF Correlation

In order to analyze the Subprime Crisis risk contagion from the U.S. to the Brazilian economy, it was used the DCLMF methodology to calculate correlations free from market specific autocorrelation structures, in order to study whether a true short-term innovation affected risk aversion indices in both economies.

To accomplish that, it was chosen a timeframe of 252 trading days for the rolling regression, calculated following the methodology presented in Chapter 4. Then, by means of a graphical inspection, it was defined the risk contagion windows, which are necessary to the application of the Step 2. The results of imposing risk contagion windows versus the calculated correlations are shown in Figure 5.2.

Figure 5.2 – DCLMF correlation and risk contagion windows.



As can be seen in Figure 5.2, it is possible to see that the imposed risk contagion windows fit with the rapid transitions observed in the DCLMF correlation. It is also important to notice that the identified timeframes are compatible with the observations made by Barbosa and Pessôa (2014), who point the original Subprime Crisis peak and the Quantitative Easing episodes as the most relevant events in terms of risk contagion.

The first episode started in Sept/08 and the second episode started in Aug/2010, that are compatible with the imposed risk contagion windows, which were established according to the observed transitions in the DCLMF correlation.

5.3.2 Step 2: What would be the Conditional Variance without the Risk Contagion?

Filtering the conditional variance in Brazil in terms of the risk contagion from the U.S. economy is not a trivial task. For example, Cesa-Bianchi, Pesaran and Rebucci (2014) use a GVAR-VOL approach to estimate the impulse responses of the major economies in conjunction with their respective volatilities, in order to measure the effects of the crisis worldwide. The present approach in this work is completely different. It completely relies on measuring the effects of risk contagion in terms of the fractionally differenced conditional variance and the respective identified risk contagion frames in the previous step. In other words, the present approach is totally based on a time series approach.

In Figure 5.3 it is possible to see that there is a rise in the short-term fluctuations in the fractionally differenced conditional variance in the risk contagion episodes delimited by the contagion windows. Keeping that in mind, the idea of the filtering technique here developed is to determine an ARMA structure compatible with this time series, in order to preserve the autocorrelation structure of the time series – which is an important characteristic of the market structure – and simulate an artificial ARMA process, which follows this specific structure to replace the observed peaks with standard short-term fluctuations, as if there was no stress in the market.

Hence, the first task was to fit a simple ARMA model on the fractionally differenced conditional variance series. Starting with a simple ARMA(1,1) with a constant, a first model was obtained according to Table 5.1.

Figure 5.3 – Fractionally differenced conditional variance (Ibovespa) and risk contagion windows.

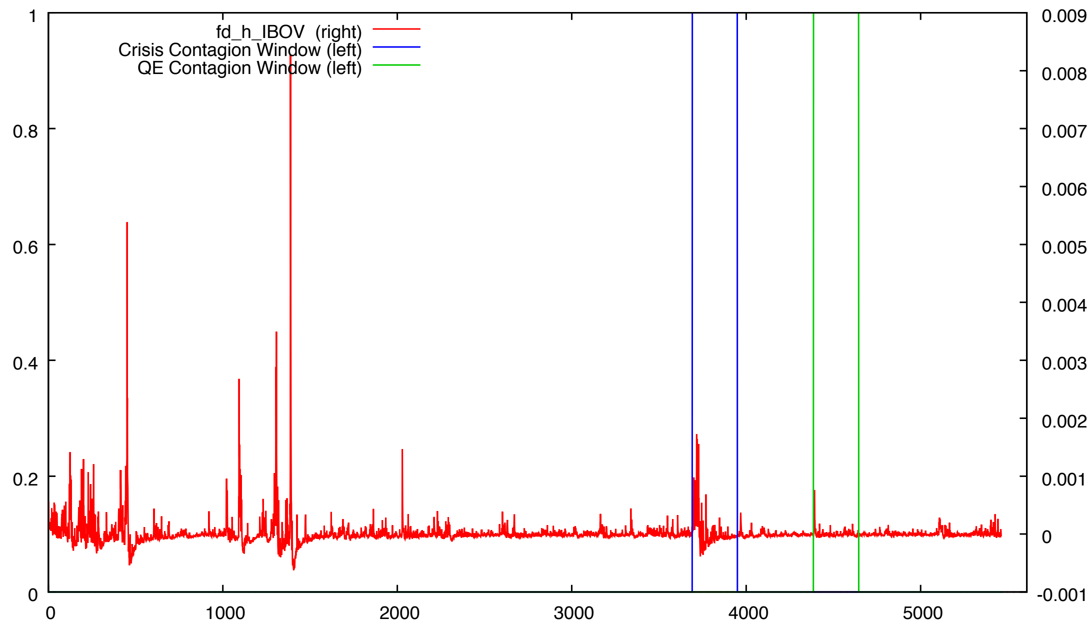


Table 5.1 – Estimation of ARMA process over the fractionally differenced conditional variance series.

Model 3: ARMA, using observations 2-5462 (T = 5461)

Estimated using Kalman filter (exact ML)

Dependent variable: fd_h_IBOV

Standard errors based on Hessian

Variable	Coefficient	Std. error	*z*	*p*-value
Const	6.51683e-8	8.34188e-6	0.7812	0.4347
Phi_1	0.797724	0.0121838	65.47	0.0000
Theta_1	-0.283641	0.0192384	-14.74	3.39e-49

Mean dependent var: 6.24e-6	S.D. dependent var: 0.000229
Mean of innovations: -1.00e-7	S.D. of innovations: 0.000174
Log-likelihood: 39518.18	Akaike criterion: -79028.37
Schwarz criterion: -79001.94	Hannan-Quinn: -79019.15

(continues)

Table 5.1 -
(continued)

	Real	Imaginary	Modulus	Frequency
AR Root 1	1.2536	0.0000	1.2536	0.0000
MA Root 1	3.5256	0.0000	3.5256	0.0000

In Table 5.1 it is possible to verify that the constant is not statistically relevant. Hence, a new ARMA model was estimated in Table 5.2, without constant.

Table 5.2 – Estimation of ARMA process over the fractionally differenced conditional variance series (without constant).

Model 3: ARMA, using observations 2-5462 (T = 5461)
Estimated using Kalman filter (exact ML)
Dependent variable: fd_h_IBOV
Standard errors based on Hessian

Variable	Coefficient	Std. error	*z*	*p*-value
Phi_1	0.798001	0.0121731	65.55	0.0000
Theta_1	-0.283830	0.0192331	-14.76	2.76e-49

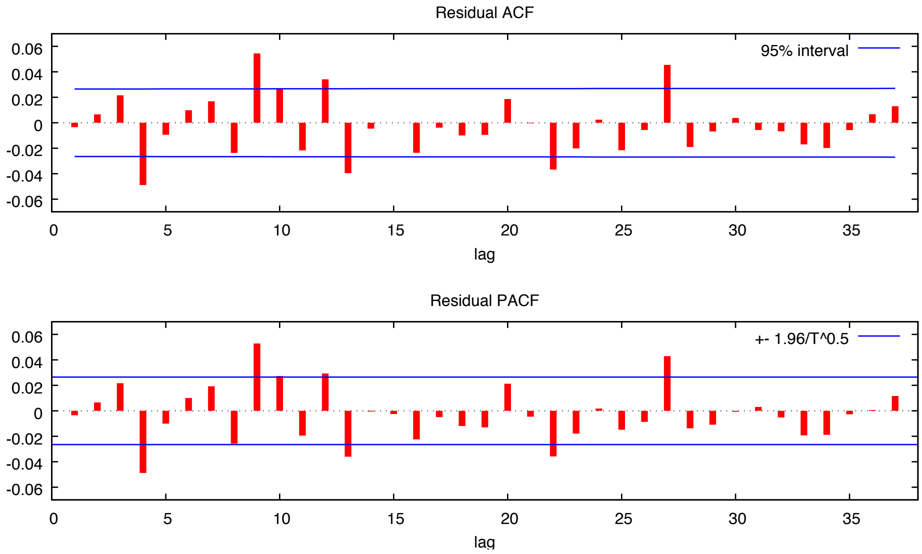
Mean dependent var: 6.24e-6	S.D. dependent var: 0.000229
Mean of innovations: -1.74e-6	S.D. of innovations: 0.000174
Log-likelihood: 39517.88	Akaike criterion: -79029.76
Schwarz criterion: -79009.94	Hannan-Quinn: -79022.84

	Real	Imaginary	Modulus	Frequency
AR Root 1	1.2531	0.0000	1.2531	0.0000
MA Root 1	3.5252	0.0000	3.5252	0.0000

In Table 5.2 it is possible to verify that the constant removal has not affected the estimation of the other parameters. Following the literature on time series analysis (HAMILTON, 1994), to check whether there is a significant remaining

ARMA structure yet to be identified, autocorrelation and partial autocorrelation plots are analyzed in Figure 5.4.

Figure 5.4 – Autocorrelation and partial autocorrelation plots of residuals.



In Figure 5.4 it is possible to see that both autocorrelation and partial autocorrelation plots share the same structure. The direct consequence of that is the fact of existing (approximately) only white noise in the residuals. To better understand such statement, consider a possible ARMA structure remaining over the residuals in the form of:

$$e_t \cdot (1 - \phi B) = a_t \cdot (1 - \theta B) \tag{91}$$

where e_t are the residuals of the initial estimated model, a_t is a white noise process, $(1 - \phi B)$ is the autoregressive polynomial (AR polynomial) and $(1 - \theta B)$ is the moving average polynomial (MA polynomial). If one rewrites the above expression as:

$$e_t = a_t \cdot \frac{(1 - \theta B)}{(1 - \phi B)} \tag{92}$$

one obtains an expression that relate the residuals with white noise in terms of a possible remaining ARMA structure. On the other hand, if one thinks in terms of pure AR and pure MA processes, the autocorrelation function (ACF) of a pure AR process behaves exactly the same way as the partial autocorrelation function (PACF) of a pure MA process (MORETTIN; TOLÓI, 2006).

Consequently, if ACF plots share the same structure as the PACF plots, the AR and MA polynomials approximately cancel out, pointing out that $e_t = a_t$.

Hence, despite the existence of significant lags in terms of ACF and PACF calculated over residuals, it is very likely that they are spurious. To prove such statement, if one considers an extended ARMA model with lags 1, 4, 8 and 9, as suggested by the first lags in the ACF and PACF plots in Figure 5.4, one obtains the model in Table 5.3.

Table 5.3 – Estimation of an extended ARMA process over the fractionally differenced conditional variance series.

Model 3: ARMA, using observations 2-5462 (T = 5461)
Estimated using Kalman filter (exact ML)
Dependent variable: fd_h_IBOV
Standard errors based on Hessian

Variable	**Coefficient**	**Std. error**	***z***	***p*-value**
Phi_1	0.828583	0.0189340	43.76	0.0000
Phi_4	0.0119586	0.0212150	0.5637	0.5730
Phi_8	-0.167756	0.149998	-1.118	0.2634
Phi_9	0.109791	0.125852	0.8724	0.3830
Theta_1	-0.319595	0.0259918	-12.30	9.52e-35
Theta_4	-0.0774525	0.0251454	-3.080	0.0021
Theta_8	0.138934	0.153425	0.9055	0.3625
Theta_9	0.034510	0.0542937	0.6161	0.5378

(continues)

Table 5.3 -
(continued)

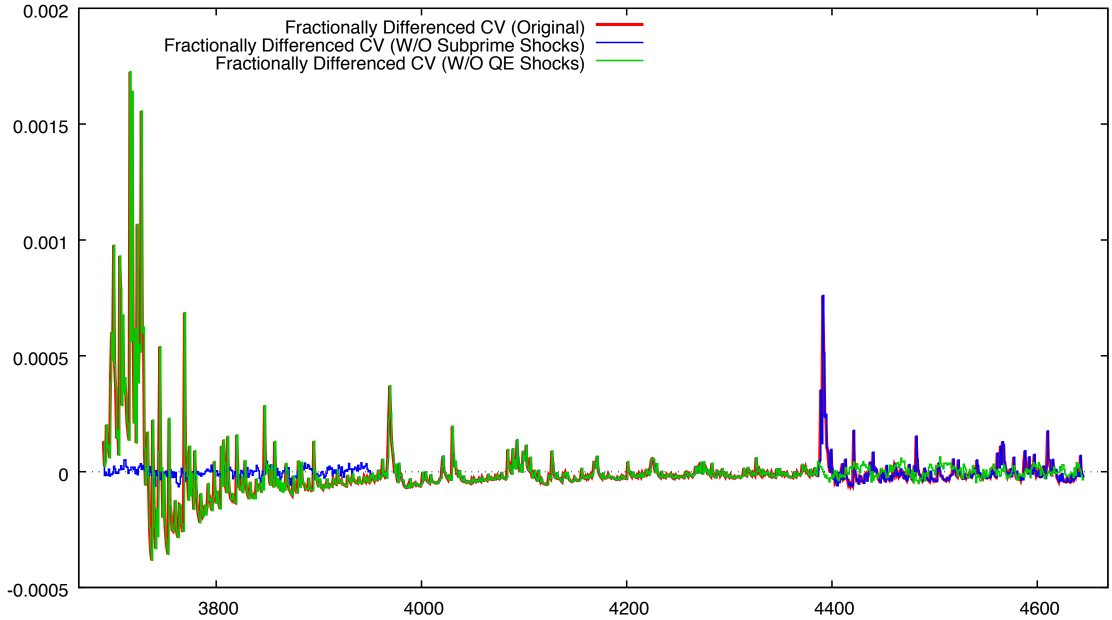
Mean dependent var: 6.24e-6	S.D. dependent var: 0.000229
Mean of innovations: 1.73e-6	S.D. of innovations: 0.000174
Log-likelihood: 39538.22	Akaike criterion: -79058.45
Schwarz criterion: -78999.00	Hannan-Quinn: -79037.70

According to Table 5.3, it is possible to verify that the ARMA(1,1) coefficients barely changed and only one additional parameter (just the fourth lag in the MA polynomial, of all those that the PACF and ACF plots suggested) is statistically significant. And still, despite the value of this additional parameter being close to zero, suggesting a possible spurious relation – does it make sense that the current conditional variance depends on the variance observed four days ago?

Additionally, the inclusion of all these parameters barely changed the information criterion indices, suggesting the choice of the ARMA(1,1) structure. Thus, the simplest model was chosen to serve as a reference to simulate an artificial level of uncertainty, without risk contagion.

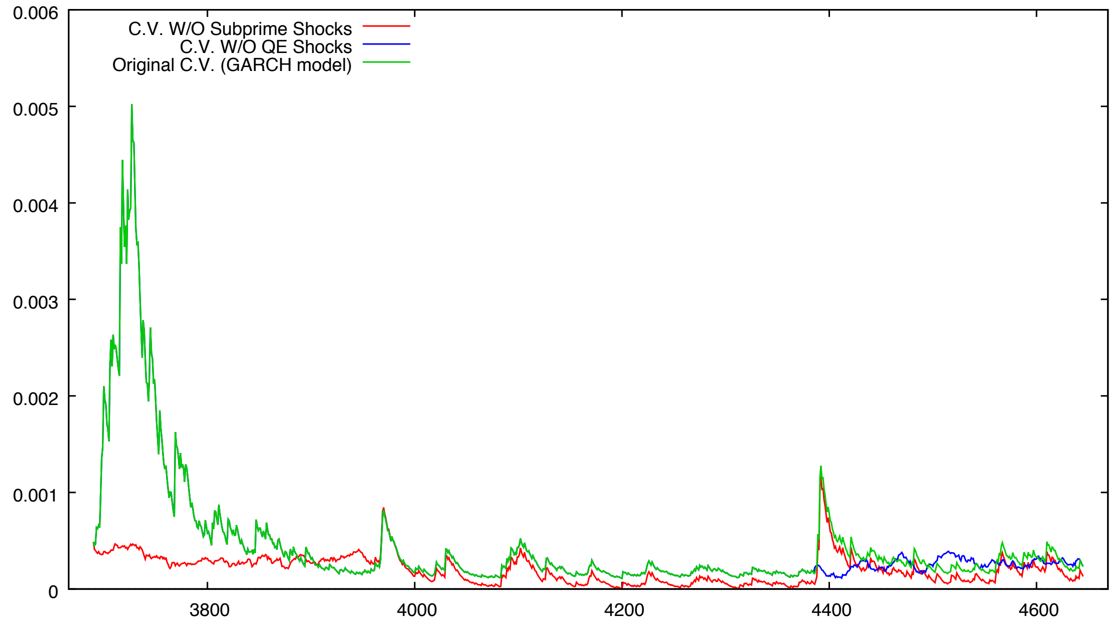
Keeping in mind the estimated ARMA model, one hundred Monte Carlo simulations were carried out, and then averaged, in order to produce an estimate of uncertainty levels without risk contagion episodes. Afterwards, during the contagion episodes – which were explicitly defined by using the DCLMF correlation, the observed fractionally differenced conditional variance (which represents the true short term shocks transmitted to the Brazilian economy) was substituted by the simulated series, as in Figure 5.5.

Figure 5.5 – Fractionally differenced series with and without risk contagion shocks.



In Figure 5.5, during the contagion episodes, as explained before, it was introduced a simulated uncertainty compatible with the autocorrelation structure observed in real data. Then, when a fractional integration is carried out (by reversing the fractional difference operation), estimates of the conditional variance are obtained as in Figure 5.6.

Figure 5.6 – Estimated conditional variance versus counterfactual estimates without risk contagion shocks.



In Figure 5.6 one can observe that volatility estimates without the risk contagion shocks are significantly lower in comparison with the peaks, but not necessarily lower all the time, due to the preservation of the autocorrelation structure, inherent to the use of ARMA models in the simulation process. Furthermore, it is also important to notice that the long memory components play an important role in the slow rate of convergence of the series. This is a significant factor that will be important in the production of the counterfactual scenarios for the GDP discussed later.

5.3.3 Estimates of Monthly Real GDP

Whenever a comparison or a study is made in terms of economic activity, the first thing that should be taken into account is variation of prices. As a consequence of inflation, economic agents are not able to buy the same amount of goods at different periods. Hence, any economic activity index, such as GDP, must be properly adjusted (if the original data is related to the nominal currency), in order to avoid any distortion caused by inflation.

Moreover, in a country like Brazil, the estimation of a monthly real GDP is very important, in order to obtain an estimate of the impact of uncertainty over economic activity. If one uses a quarterly or an annual series in order to estimate such impacts, biased results may arise due to the lack of degrees of freedom; or due to the rapid level shifts that naturally occur in equity volatilities – where significant co-movements can be lost. In addition to that, instead of using industrial production - as in Bekaert and Horoeva (2014) – using GDP seems more adequate, given the importance of primary and tertiary sectors (i.e. agriculture, mineral extraction and services, respectively).

Despite the availability of nominal GDP estimates from the Brazilian Central Bank, obtaining real GDP estimates is a tricky task, due to the fact that GDP deflators are provided only in a yearly frequency. Hence, an alternative methodology is developed in this section.

The basic idea of this methodology consists of combining information from the monthly observed inflation series (IPCA – Brazilian Consumer Price Index, which is the index used for inflation targeting purposes, calculated by Brazilian Statistics Bureau, namely IBGE), real annual GDP data (provided by the Brazilian

Central Bank) and nominal monthly GDP data (also provided by the Brazilian Central Bank). To accomplish that, the following two-step procedure is applied:

- Deflate the nominal GDP using the IPCA series;
- By means of a mathematical programming procedure, apply a geometrical correction factor over the monthly GDP deflated series, which is established by a mathematical programming procedure, where the goal is to minimize to zero the distance between the real observed GDP growth and the corrected deflated GDP growth.

The first step is the traditional deflation procedure, which consists of dividing the nominal GDP by the accumulated observed inflation series. The second step is more complex. Mathematically speaking, the second step consists of finding a set of z_i, which are applied in each year over the months of that respective period such as:

$$\begin{cases} Adj\ GDP_{i,1} = \dfrac{Nom\ GDP_{i,1}}{\pi_{accum(12 \cdot i+1)}} \cdot \dfrac{1}{(1+z_i)} \\ \quad\quad \cdots \\ Adj\ GDP_{i,n} = \dfrac{Nom\ GDP_{i,1}}{\pi_{accum(12 \cdot i+n)}} \cdot \dfrac{1}{(1+z_i)} \\ \quad\quad \cdots \\ Adj\ GDP_{i+1,n} = \dfrac{Nom\ GDP_{i+1,n}}{\pi_{accum(12 \cdot (i+1)+n)}} \cdot \dfrac{1}{(1+z_i)} \end{cases} \tag{93}$$

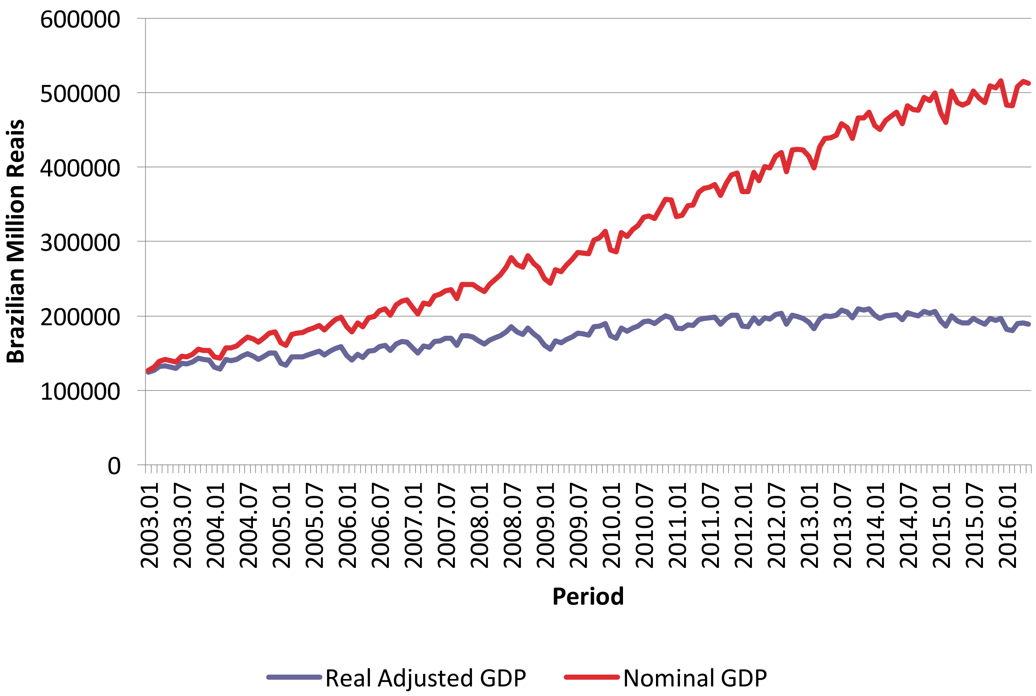
where i denotes the year (the first year in the series is indexed by zero); $1 \le n \le 12$ denotes the month within the i-th year; aiming to minimize the distances:

$$d_i = \left| g_i - \frac{\sum_{k=1}^{12} Adj\ GDP_{i,k}}{\sum_{k=1}^{12} Adj\ GDP_{i-1,k}} \right| \tag{94}$$

with g_i being equal to the observed annual real GDP growth rates.

Solving the stated mathematical programming problem above, one obtains a set of real monthly GDP estimates that are consistent with the observed real GDP growth estimates in yearly basis. A comparison between the obtained real GDP monthly data *versus* the nominal GDP estimates is shown in Figure 5.7.

Figure 5.7 – Real GDP estimates versus nominal GDP estimates (monthly data).



In Figure 5.7, it is possible to verify that a good part of the evolution of the nominal GDP is related to price variations. When one compares the real GDP estimates against the nominal estimates, it is possible to infer that despite the common fluctuations in the series caused by real business cycles, they have totally different trends as consequence of moderate observed inflation. So, it is possible to see the importance of adjusting nominal data to a constant currency.

5.3.4 Calculating the Impacts of Uncertainty in the Real Sectors

The impacts of uncertainty (measured in this case by the volatility in equity prices) over the real sectors were calculated following the methodology discussed in Bekaert and Horoeva (2014). In their work, they calculated the elasticity between the industrial production in the U.S. and the conditional variance (CV). In the present approach, instead of using industrial production, it was used the real GDP estimates in a monthly frequency, as developed in Step 3.

In addition to that, aiming to select the best predictor variable, using a similar approach to Bekaert and Horoeva (2014), which evaluate the best predictor by running all individual regressions, it was built Table 5.4. The underlying model structure used to build Table 5.4 was:

$$\Delta \log(GDP_t) = \beta \cdot \sigma_t^2 \; + c \tag{95}$$

where σ_t^2 is substituted by the appropriate conditional variance, according to the tested variable.

Table 5.4 – Predictive power of different conditional variance estimates.

Param.	**Conditional variance IBOV - Average through time**				
	1 month	**2 months**	**3 months**	**6 months**	**12 months**
Elasticity coeff.	-0.219846	-0.263966	-0.287064	-0.248926	0.0233731
Constant coeff.	0.00376340	0.00403465	0.00416765	0.00388426	0.00210186
***p*-value elasticity**	0.0000480	4.63e-7	0.000127	0.0924	0.9386
***p*-value constant**	0.0000350	0.0000133	0.000013	0.0003	0.2415
R^2	0.129674	0.145762	0.139435	0.063659	0.000299

As shown in Table 5.4, the conditional variance averaged over 2 months seems to be the variable with the best predictive power in terms of explaining economic contractions, as a consequence of uncertainty. It is worth mentioning that, as expected, the obtained elasticity coefficients were negative, meaning that as uncertainty increases, there is a drop in the economic activity.

Finally, it is also important to mention that the values obtained in Table 5.4 are higher than those obtained by Bekaert and Horoeva (2014). The elasticities between conditional variance and economic activity (industrial production) estimated by them range from -0.033 to -0.113. All of them are lower than all statistical significant estimates here obtained – this may be a result of using the whole GDP instead of only using industrial production. On the other hand, the R Squared of the estimated models obtained by Bekaert and Horoeva (2014) are in line with those found here. They range from 0.08 to 0.27. Consequently, it is possible to say that the obtained results for the Brazilian economy (here) are in line with the findings obtained for the American economy, in terms of predictive power of uncertainty *versus* economic activity.

5.3.5 What-If Analysis – What would have happened if Risk Contagion Shocks had not occurred?

After the calculation of the four previous steps, it makes sense developing a methodology to estimate counterfactual predictions of what would have happened to the Brazilian economy if the contagion risk shocks had not occurred and to what extent these shocks were responsible for the current observed economic recession in Brazil.

To accomplish that, the methodology here developed, first, consists of comparing what would have happened to GDP, making use of the calculations made in Step 4 and the counterfactual conditional variances estimated in Step 2.

Then, in order to evaluate the reasonability of the obtained results from the counterfactual analysis, the results are compared against a trend forecast of the potential GDP.

The first part here mentioned is very simple and straightforward. Having estimated the elasticity of real GDP versus uncertainty and knowing the counterfactual estimates of the conditional variance without the risk contagion shocks, it is worth examining the GDP growth in terms of uncertainty from Eq. (95),

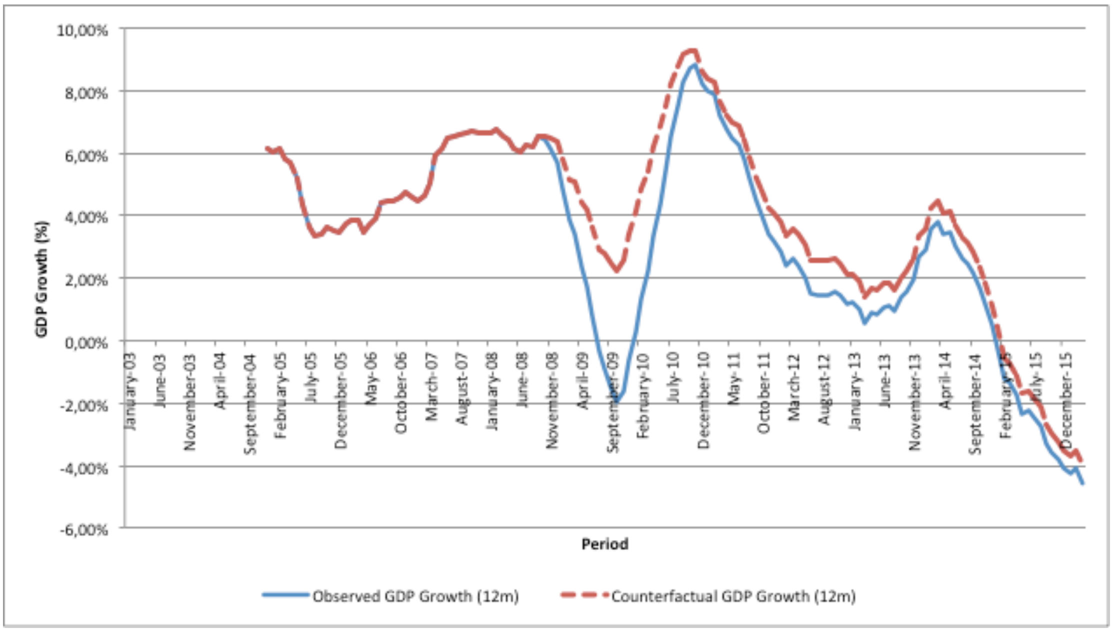
$$\Delta \log(GDP_t) = \beta \cdot \sigma_t^2 + c$$

Hence, if there is an estimate of $\Delta\sigma_t^2$, which is given by the difference between the estimated conditional variance using GARCH models and the counterfactual predictions calculated in Step 2, it is possible to calculate:

$$\Delta^2 \log(GDP_t) = \beta \cdot \Delta\sigma_t^2 \qquad (96)$$

where $\Delta^2 \log(GDP_t)$ denotes the marginal increase in the GDP growth at instant t. Consequently, it is just a matter of integrating the logarithms of the GDP and calculating their respective exponentials, in order to obtain the counterfactual predictions of the series in its original level. Thus, it is possible to analyze Figure 5.8, which contains the counterfactual predictions of the GDP growth against the observed GDP growth.

Figure 5.8 – Observed GDP growth estimates versus counterfactual GDP growth estimates (monthly data).

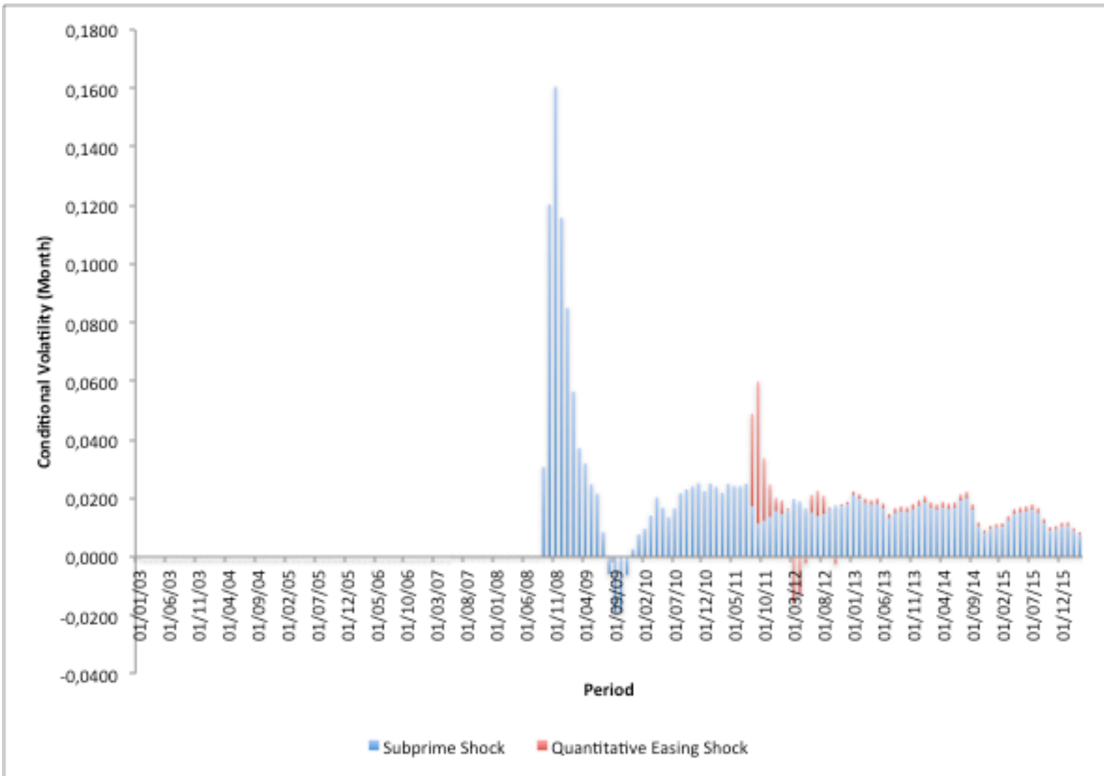


It is possible to verify in Figure 5.8 that, as expected, the counterfactual GDP growth is higher than the observed GDP growth. But more interesting is the fact that the shocks originated from the risk contagion episodes have not been dissipated yet. This is a clear result from the fact that conditional variance exhibit long-range dependency, as discussed through Chapters 3, 4 and this present one.

Presumably, shocks occurred far away in time still have a significant impact in today's uncertainty, as a consequence of the reaction of economic agents in the context of specific market structures. In order to show that, Figure 5.9 was built in order to allow one to visualize the effects of the Subprime shocks and Quantitative

Easing shocks in the uncertainty (conditional volatility) and how they are evolving over time.

Figure 5.9 – Subprime shock and quantitative easing shock participations in equity volatility over time.



In Figure 5.9, as mentioned before, it is possible to see the two major shocks that impacted the Brazilian economy. The Subprime shock was the most significant of them – which is in line with the findings of Barbosa and Pessôa (2014). But the most important finding is the fact that despite the QE shock has been dissipated, the Subprime shock was not fully dissipated yet. This finding not only points a huge impact in the economic activity but in several other economic affairs, such as economic regulation – which relies on asset pricing (and risk) to determine fair returns for public services, investment decisions and monetary policy – as consequence of risk premium shocks.

Taking into account these observations, the second part of the present analysis consists of assessing these obtained results against the Potential GDP (obtained using HP trend filter) forecasts conditional to the information available until the beginning of the crisis. This hypothetical series synthetizes a hypothetical economic activity level, where capital and employment are being fully employed, without inflation pressures, with constant population growth and constant productivity growth resultant of application of new technologies and knowledge.

One might ask if it is reasonable to compare the counterfactual predicted GDP against forecasted Potential GDP trend. There are plenty of reasons to believe that this is reasonable.

First, if a Solow-Swan model is assumed to explain the dynamics of a simple economy, the GDP could be explained as:

$$Y_t = K_t^{\alpha}(A_t\, L_t)^{1-\alpha} \tag{97}$$

$$\Delta L_t = n\, L_{t-1} \tag{98}$$

$$\Delta A_t = g\, A_{t-1} \tag{99}$$

$$\Delta K_t = (s - \delta)\, K_{t-1} \tag{100}$$

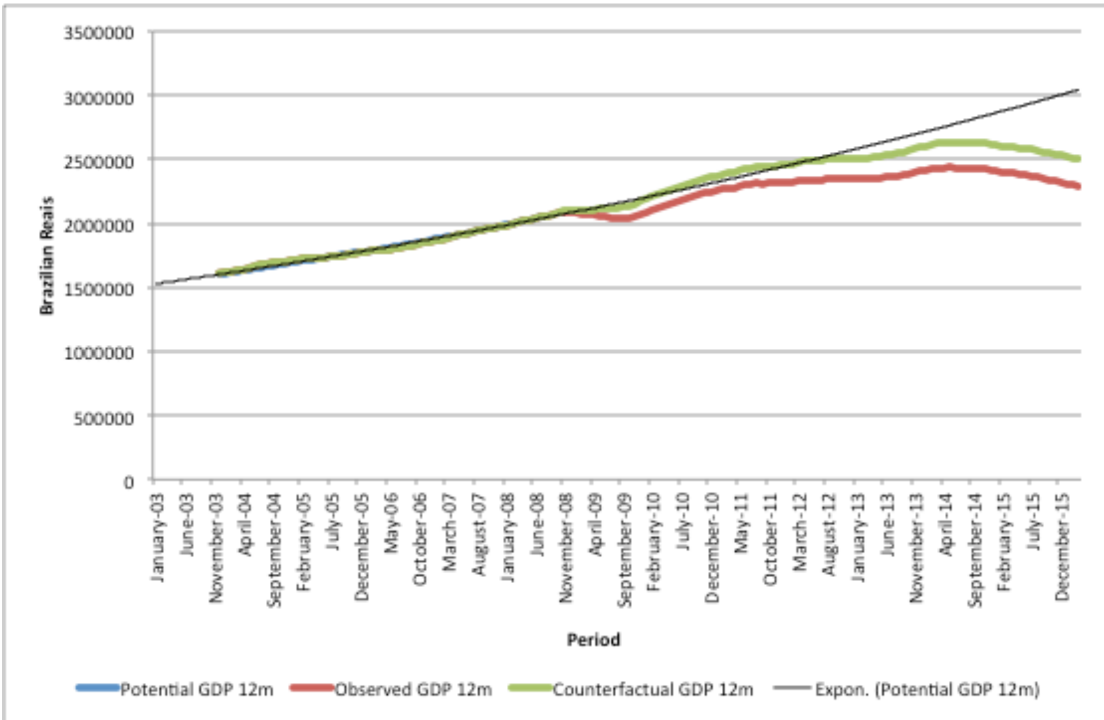
where:

- L_t denotes the labor (i.e. number of workers employed);
- A_t is the productivity level;
- K_t the capital stock used in the production of goods and services;
- Y_t denotes the economic output (GDP), by means of a Cobb-Douglas function, where all inputs are multiplied, ensuring that if no labor is employed, the output is zero; the same for capital and productivity;
- α is an dimensionless parameter that corresponds to the participation of each factor in terms of income obtained in the production of goods and services;
- n is the population growth rate parameter;
- s is an investment rate parameter;
- δ is the depreciation rate of the stock of capital;
- g is the productivity growth rate parameter.

That said, it is widely known that, in Brazil, there were no discontinuities in the period between 2009 and 2016, in terms of population evolution (there were no significant changes in the population structure). But there were structural changes in productivity evolution and in terms of capital investment/use, as seen in Barbosa and Pessôa (2014).

Consequently, if a comparison is carried out between the Potential GDP trend forecast and the counterfactual analysis obtained by the exclusion of risk contagion effects (see Figure 5.10), it is possible to see in what extent the crisis (exogenous factor) contributed to drops in productivity and investments – by means of the three transmission channels discussed in the first part of this chapter, and what is a consequence of internal macroeconomic policies.

Figure 5.10 – Evolution of potential GDP trend, real GDP and counterfactual GDP prediction.



In Figure 5.10 it is possible to see that the counterfactual prediction (in green) floats around the forecasted Potential GDP trend (thin black line) until 2012. On the other hand, Kanczuk (2014) points out that the transmission mechanism of the crisis was more finance than trade, and this is the main reason why, without the uncertainty shocks, the economy would have been able to walk around the Potential GDP trend until 2012 – the weak external demand did not play an important role in the GDP growth. To reinforce such statement, it is worth mentioning that Brazil is a relatively closed economy, as discussed in Kanczuk (2013). So the transmission effect of the crisis did not happen through international trade. Thus, the results obtained in this study are in line with the existing literature and there are also important evidences towards the assumption of the role of risk contagion in the Brazilian GDP growth, which was important but not the most important factor that explains its huge failure.

Coincidentally or not, in 2011, a set of heterodox economic policies entitled “Nova Matriz Econômica” (New Economic Matrix – in free translation) started to be conducted by the Brazilian Government. As discussed in Cunha (2015), during this period, the monetary authorities became less concerned with bringing inflation to the center of the target; there was a closing down of the economy to foreign competition through the enactment of “national content” rules; systematic deployment of public banks for policy objectives and subsidies to targeted industries and huge public investment financed by the private sector.

Complementing the discussions in Cunha (2015) with the findings in Kanczuk (2014), by means of a DSGE model, Brazil has suffered unsuccessful credit shocks (which were an important part of the “New Economic Matrix”) that resulted in a huge deceleration of the GDP growth. Hence, there are solid straight links between the findings of this study, with those currently discussed for the Brazilian economy.

Consequently, calculating the differences between the Potential GDP trend forecast, the counterfactual prediction of the GDP – in a scenario without the exogenous uncertainty shocks, and the estimated real GDP, 70% of the GDP current level was caused by erroneous macroeconomic policies and only 30% was caused by exogenous shocks as consequence of the international crisis. The Brazilian GDP, in comparison with the Potential GDP trend forecast, is 23.63% lower than it should be – if correct macroeconomic policies would have been imposed and if the crisis would have not happened.

Moreover, in U.S. dollars (2009), adjusted by an average nominal exchange rate of R$/USD 2,23 and adjusted by a US-CPI factor of 1.1226, the crisis costs reached approximately 220 billion dollars until Mar/2016.

While the crisis effects are not completely dissipated – due to the long-range dependency inherent to the market structure and to the response of the agents to exogenous shocks; and while proper solutions to the problems caused by the “New Economic Matrix” policies are not found, this terrible situation will become even worse.

Great lessons must be learned from huge mistakes.

6 CONCLUDING REMARKS

During the Introduction of this Thesis, a major provocation was made: several phenomena exhibit long memory properties. But why do they exhibit such feature?

Inspired in the Complex Systems approach, which studies several sorts of phenomena in Physics, Chemistry, Computation and other fields where small perturbations in microscopic scale generates huge disturbances in macroscopic scale, a great motivation for the study of long-range dependency as an emergent feature of complex system arises. If there are interesting properties that emerge with complex behavior, such as power-law, self-similarity, chaotic behavior, etc; is long-range dependency one of them?

Part of these questions started being answered during Chapter 2, where a motivation from the fractional calculus is presented, in the sense that fractional derivatives/fractional differences operators are nothing more than tools to model dynamic systems, where all information available far away in space and time plays an important role in the dynamics, instead of only considering the proximity to the current state in the system.

Also, fractional operators generalize the integer order calculus, providing more flexibility to model or combining different properties in only one model, such as viscoelasticity.

Hence, if complex systems usually are irreducible – which means that it is not possible to predict its future states without passing through all intermediate states, and consequently, the current state depends on all past information, it totally makes sense exploring the possibility of complex phenomena also exhibiting long-range dependency.

After such motivational part, it is introduced the fractional difference parameter and its ability to capture relevant information far away in time, and also, how the most used statistical techniques work, in order to identify these long memory systems.

That said, in Chapter 3, it was shown that it is possible to have long memory properties in variables that measure aggregate behavior of the agents, by analyzing the presence of the fractional difference parameter, as discussed in the traditional literature, without the need of economic/social frictions nor the presence of long memory behavior in the individual agents, due to their respective biological processing unit capabilities (i.e. their brains).

Then, in order to show that, first, traditional estimate calculations were carried out, by estimating the fractional difference parameters, according to the main techniques – GPH (1983) and Künsch (1987). It was also taken into account the perspective of having spurious fractional difference coefficients derived from first order autoregressive structures. To avoid that, the author controlled the parameters of the first of three experiments to avoid any spurious long-range dependency.

Moreover, statistical tests that take into account the presence of short-memory components were also carried out (namely V/S Statistic Test and R/S Statistic Test), in order to reinforce such findings.

So, the author conducted three different computational experiments from previous works, aiming to discuss the rise of long memory properties in the aggregate behavior of the agents, by simulating their individual actions and interactions. These three experiments were taken out in order to remove any possible perspective of conception biases, which may lead to the construction of systems with long memory properties. In other words, the present author chose experiments built by other authors without such analysis perspective, in order to show that such properties may arise according to system features, instead of building models aiming to obtain long-range dependency.

Thus, in the first experiment, adapted from Jumadinova, Matache and Dasgupta (2011), it was shown that long memory properties usually arise from heterogeneity between agents, in a world where space and topology/local interactions do not have any role.

In the second experiment, it was studied a class of individual computational simulations called Cellular Automata, which aims complementing the analysis done in the first experiment. Instead of having heterogeneous agents and a network that does not allow local interactions, it was chosen an experiment where all agents are homogenous (i.e. act the same way, using the same set of rules) and they interact locally, producing a global behavior. Consequently, it was shown that local interactions of individual homogeneous agents may produce long memory patterns observed in the global behavior, as a consequence of the implicit dynamics in the transition rules that drives each agent behavior (laws of motion) – and the presence of equilibrium points in these dynamics. When introducing large deviations from these equilibrium points, there is a convergence to the equilibrium point of each system that can be compared to smooth regime switches, which is known to generate a spectral density similar to a long-range dependent phenomenon.

Furthermore, in the third experiment, in order to complement both previous experiments, it was chosen another model – adapted from Epstein and Axtell (1996), where heterogeneity from the agents was removed and direct interactions between them were also removed. So basically, they only interact with a space that depends on the spatial configuration that the author set up, in order to control the presence of spatial complexity. Thus, such experiment suggests that spatial complexity is another interesting feature that plays an important role in the rise of long memory properties.

Therefore, keeping in mind these three different results, it is possible to show that long-range dependency is a feature that emerges from the system complexity and not necessarily arises from individual long memory capabilities nor economic/social frictions and extends the idea of heterogeneity as a source of complexity, as usually discussed in the traditional economic literature (TEYSSIÈRE; KIRMAN, 2006).

Finally, when these ideas are transposed to a real society, which is composed of heterogeneous agents, bounded rationality, explicit space, local interactions and non-equilibrium dynamics, it makes sense recognizing that macroeconomic and financial markets data would be subject to the presence of long-range dependency, due to the fact that its underlying elements are characterized by such features. And by applying a long memory filter over real data, one would be able to stick only with short-term fluctuations, which may enhance the comprehension of several phenomena in terms of stylized facts.

Nonetheless, it is worth making a final discussion here, about this theme, as part of perspectives for future works and analyses. The traditional literature is focused on building specific models that matches the spectral density of long memory processes, showing possible “confusions” and “misunderstandings” between what is a real “long-range dependent” and a spurious process. This work aims to be provocative in this sense.

It is really hard to establish whether a stochastic process is the true engine behind the data generation process, or just a good approximation of its behavior, especially in the second and third cases here discussed. When systems like these are built and studied, it makes more sense studying them in terms of their respective components that generates such behavior (and study long memory as an emergent property) rather than trying to find a plausible explanation based on the traditional toolset – which is something hard to be made, as seen in the second experiment.

Hence, it is suggested expanding this analysis to other agent based models and cellular automata, setting up other possible sources of complexity within the system, in order to verify if such features appear or not.

Throughout the Chapter 4, a first interesting tool was developed based on the discoveries made in Chapter 3. If part of the autocorrelation structure may be a result of specific market/economic idiosyncrasies, what about applying a long memory filter over real conditional variance data (which is a proxy for the uncertainty of economic agents) and study how risk is spread across different countries?

This is an intriguing and fascinating theme, given the huge impact that the Subprime Crisis caused in the world economy. Moreover, globalization, deregulation and technological advances (such as real-time trading) have significantly increased the transmission of information, causing a greater integration between markets. These factors contributed to the intensification of the phenomenon of financial risk contagion, which, consequently, increases the complexity in the risk analysis, because external factors must be taken into account when calculating the risk exposure.

With the advent of long-memory filters one is able to remove all spurious correlations, which may be consequence of the inherent autocorrelation structure of each market allowing one to deal only with the desired stylized facts. Hence, based on the filtered risk aversion variables (conditional variances obtained by GARCH models), one is able to calculate correlations between pairs of countries, allowing one to obtain a good risk contagion index, which was entitled DCLMF (Dynamic Correlation based on Long Memory Filters).

This methodology was compared to the traditional DCC (Dynamic Conditional Correlation), which is a standard multivariate GARCH approach. In this comparison, it was clear that while DCC is more suitable for understanding the integration between different markets, due to its nature of using the standardized residuals from the GARCH models to recover a conditional correlation; the DCLMF methodology allows rapid shifts of levels (due to the application of long memory filters that remove possible agents over-reactions as a consequence of market structures), acting as a high-pass filter for risk contagion. For the purpose of this comparison, two analyses were carried out: one between U.S. and Brazil; another one between China and U.S. On both, the results obtained seemed to better explain risk contagion episodes, by comparing the conditional volatility series, in conjunction with correlation indices.

Hence, the greatest achievement of this methodology is the fact that it does not require any further analysis, as the traditional DCC does, being a far more straightforward tool.

Given this new tool, it turns out to be interesting to identify the risk contagion episodes from the U.S. economy to the Brazilian economy, in details, as done in Chapter 5.

The Subprime Crisis and the Quantitative Easing shocks, which were originated in the U.S., caused a “Tsunami” in the world economy. And this was the major excuse found in Brazil, during the last 5 years, for its huge economic failure. To what extent did these crises contributed for this economic failure?

Answering this question seems very important, given the current Brazilian economic, social and political scenario. The economic recession that is being observed in the country has led to a contraction of 3.8 percent in terms of GDP growth, an inflation surpassing 10 percent and an increasing unemployment, which induced large popular mobilizations with the singular refrain that the president should be impeached.

Hence, it turns out to be very attractive to analyze the extension of the Subprime Crisis effects in the Brazilian GDP. The huge observed economic failure is her fault or not?

To answer that question, a methodology was developed throughout Chapter 5, which consists of five steps:

- Assessing the risk contagion episodes by employing the DCLMF correlation discussed in Chapter 4;
- Filtering out the risk contagion by means of an artificial simulated ARMA process over the fractionally differenced volatility series – in order to simulate what would be the “natural volatility levels” of the Brazilian market, without the risk contagion;
- Estimate a real monthly GDP series for the Brazilian Economy, based on the nominal monthly GDP series available from the Brazilian Central Bank;
- Estimate an elasticity between uncertainty (using equity volatility from Ibovespa as proxy) and the real GDP estimates;
- Calculate a counterfactual GDP evolution, based on the volatility free of risk contagion episodes, to assess the costs of the crisis.

From this sequence of steps, major conclusions could be drawn:

- The Subprime Crisis, which started in Sept/08 and the Quantitative Easing, which started in Aug/2010, are compatible with the imposed risk contagion windows established according to the observed level shifts in the DCLMF correlation;
- Uncertainty is a relevant explanatory variable in terms of Brazilian GDP growth – in line with the existing literature;
- The uncertainty shock originated from the Subprime Crisis episode has not been dissipated yet. This is a clear result from the fact that conditional variance exhibits long-range dependency. This finding not only points to a huge impact in the economic activity but in several other economic affairs, such as economic regulation – which relies on asset pricing (and risk) to determine fair returns for public services; investment decisions and monetary policy – as consequence of risk premium shocks;
- The comparison carried out between the Potential GDP trend forecast and the counterfactual analysis obtained by the exclusion of risk contagion effects enlightens the extensions of the international crisis (exogenous factor) as a major factor to the drops in productivity and investments in Brazil. But not the

most important factor. According to this methodology, 70% of the GDP drop to the current level was caused by erroneous macroeconomic policies, and only 30% were caused by exogenous shocks as consequence of the international crisis. This observation is in line with the findings of Kanczuk (2014), which points the failure of the "New Economic Matrix" and with the discussions in Cunha (2015);

- The crisis costs have reached approximately 220 billion dollars (2009 USD) until Mar/2016;
- The Brazilian GDP, in comparison with the Potential GDP trend forecast, is 23.63% lower than it should be – if correct macroeconomic policies would have been imposed and if the crisis would have not happened.

REFERENCES

ALMEIDA, C.M.; GLERIANI, J.M. **Cellular Automata and Neural Networks as modeling framework for the simulation of urban land use change**,. In: SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO, XII., Goiânia, 2011. **Anais**... Brasília: INPE, 2011.

ANDREWS, D. Heteroskedasticity and autocorrelation consistent covariance matrix estimation. **Econometrica**, vol. 59, p. 817–858, 1991.

ATANGANA, A.; SECER, A. A Note on Fractional Order Derivatives and Table of Fractional Derivatives of Some Special Functions. **Abstract and Applied Analysis**, vol. 2013, article ID 279681, 8 pages, 2013.

BAILLIE, R. T. Long memory processes and fractional integration in econometrics. **Journal of Econometrics**, v. 73, p. 5-59, 1996.

BARBOSA, F.H.F.; PESSÔA, S.A. International Macroeconomic Impacts on the Brazilian Economy: a note. In: 2014 MEETING OF NATIONAL ECONOMIC ORGANISATIONS (OECD), Paris. **Notes**. Paris: OCDE, 2014. Disponível em: <<https://www.oecd.org/eco/International-Macroeconomic-Impacts-on-the-Brazilian-Economy-paper.pdf >>. Acesso em: 02 jul. 2014.

BEKAERT, G.; HOROEVA, M. **The VIX, the Variance Premium and Stock Market Volatility** - Working Paper Series n. 1675/May 2014. Frankfurt an Main: European Central Bank, 2014.

BERAN, J. **Statistics for Long Memory Processes**. Boca Raton: Chapman-Hall/CRC, 1994.

BERGMAN, D.R.; SECURATO, J.R.; SAVOIA, J.R.F.; CONTANI, E.A.R. U.S. subprime financial crisis contagion on BRIC and European Union stock markets. **Revista de Administração**, São Paulo, v. 50, n. 2, p. 229-240, 2015.

BIANCONI, M.; YOSHINO, J.A.; SOUSA, M.O.M. BRIC and the U.S. Financial Crisis: An Empirical Investigation of Stocks and Bonds Markets. **Emerging Markets Review**, v. 14, p. 76–109, 2013.

BÍBLIA. Livro do Gênesis 41:53. Português. **Bíblia Sagrada: Edição Pastoral-Catequética**. São Paulo: Editora Ave-Maria, 2002.

BLOOM, N. The Impact of Uncertainty Shocks. **Econometrica**, v. 77, n. 3, p. 623-685, 2009.

BOYER, B. H.; KUMAGAI, T.; YUAN, K. How do crises spread? Evidence from accessible and inaccessible stock indices. **Journal of Finance**, v. 61, n. 2, p. 957–1003, 2006.

BRATIS, T.; LAOPODIS, N.T.; KOURETAS, G.P. Systemic risk and financial market contagion: Banks and sovereign credit markets in Eurozone. In: CONFERENCE ON RESEARCH ON ECONOMIC THEORY AND ECONOMETRICS, 14., 2015, Chania. **Working Papers**. Chania: Athens University of Economics and Bussiness, 2015.

BUCHHOLZ, M.; TONZER, L. Sovereign credit risk comovements in the eurozone: Simple interdependence or contagion? **Working Paper series of the UniCredit & Universities Foundation**, vol. 47, May 2013.

CAPORIN, M.; MCALEER, M. Ten Things You Should Know About the Dynamic Conditional Correlation Representation. **Econometrics**, v. 1., n. 1, p. 115-126, 2013.

______; PELIZZON, L.; RAVAZZOLO, F.; RIGOBON, R. **Measuring sovereign contagion in Europe** - NBER Working Paper 18741. Cambridge: National Bureau of Economic Research, Inc, 2013.

CESA-BIANCHI, A.; PESARAN, M.H.; REBUCCI, A. **Uncertainty and Economic Activity: A Global Perspective** - CESifo Working Paper Series 4736. Munich: CESifo Group, 2014.

CHEN, Y.; PETRAS, I.; XUE, D. Fractional Order Control - A Tutorial. In: AMERICAN CONTROL CONFERENCE, 2009, St. Louis. **Proceedings**... St. Louis: American Control Council, 2009. p. 1397-1411.

CHIANG, T. C.; JEON, B. N.; LI, H. Dynamic correlation analysis of financial contagion: Evidence from Asian markets. **Journal of International Money and Finance**, v. 26, n. 7, p. 1206–1228, 2007.

CHOPARD, B.; DROZ, M. **Cellular Automata Modeling of Physical Systems**. Cambridge: Cambridge University Press, 1998.

COOK, M. Universality in Elementary Cellular Automata. **Complex Systems**, v. 15, p. 1-40, 2004.

CUNHA, P.V. **Brazil — Short-term Hurdles and Longer-term Challenges**. Discussion Paper for a seminar in honor of Albert Fishlow, to be held at the Casa das Garças Institute for Economic Policy Studies, Rio de Janeiro, 2015.

DEBNATH, L. Recent Applications Of Fractional Calculus To Science And Engineering. **International Journal of Mathematics and Mathematical Sciences**, v. 2003, n. 54, p. 3413-3442, 2003.

DIEBOLD, F.X.; INOUE, A. Long memory and regime switching, **Journal of Econometrics**, v. 105, p. 131–159, 2001.

ENGLE, R. Dynamic Conditional Correlation: A Simple Class of Multivariate Generalized Autoregressive Conditional Heteroskedasticity Models. **Journal of Business & Economic Statistics**, v. 20, n. 3, p. 339-350, 2002.

______; KRONER, K. Multivariate Simultaneous GARCH. **Econometric Theory**, v. 11, p. 122-150, 1995.

EPSTEIN, J. M. **Generative Social Science: Studies in Agent-based Computational Modeling**. Princeton: Princeton University Press, 2006.

EPSTEIN, J. M.; AXTELL, R. **Growing artificial societies: social science from the bottom up**. Washington D.C.: Brookings Institution Press, 1996.

FERNANDES, F.; GARCIA, C. Long Memory Components in Macroeconomic Series: Are we missing something?. In: AMERICAN CONTROL CONFERENCE, 2016, Boston. **Proceedings.** Boston: American Control Council, 2016.

______. Long Memory Properties and Agent's Behavior: An Interesting Connection. In: Brazilian Behavior Economics and Finance Meeting, 2015, São Paulo. **Proceedings.** São Paulo: Fundação Getúlio Vargas – FGV, 2015.

FISHER, J. D. M. On the Structural Interpretation of the Smets-Wouters "Risk Premium" Shock. **Journal of Money, Credit and Banking**, v. 47, n. 2-3, p. 511-516, 2015.

FORNARI, F.; MELE, A. Financial Volatility and Economic Activity. In: AMERICAN ECONOMIC ASSOCIATION MEETING, 2009, San Francisco. **Annals**... San Francisco: American Economic Association, 2009

FORBES, K. J.; RIGOBON, R. No contagion, only interdependence: Measuring stock market comovements. **The Journal of Finance**, v. 57, n. 5, p. 2223–2261, 2002.

FRISCH, U.; HASSLACHER, B.; POMEAU, Y. Lattice-gas automata for the Navier-Stokes equation. **Physics Review Letters**, v. 56, n. 14, p. 1505-1508, 1986.

GEWEKE, J.; PORTER-HUDAK, S. The estimation and application of long memory time series models. **Journal of Time Series Analysis**, v. 4, p. 221–238, 1983.

GIRAITIS, L.; KOKOSZKA, P.; LEIPUS, R.; TEYSSIÈRE, G. Rescaled variance and related tests for long memory in volatility and levels. **Journal of Econometrics**, v. 112, p. 265–294, 2003.

GRANGER, C.W.J. Long-memory relationships and the aggregation of dynamic models, **Journal of Econometrics**, v. 14,p. 227–238, 1980.

GOLDREICH, O. Invitation to Complexity Theory. **XRDS**, vol. 18, n. 3, p. 18-22, 2012.

GOMEZ-AGUILLAR, J.F.; ROSALES-GARCÍA, J.J.; BERNAL-ALVARADO, J.J; CÓRDOVA-FRAGA, T.; GUSZMÁN-CABRERA, R. Fractional mechanical oscillators. **Revista Mexicana de Física**, v. 58, p. 348–352, 2012.

HERMANN, R. **Fractional Calculus: An Introduction for Physicists, 2nd Edition**. Singapore: World Scientific Publishing Company, 2014.

HON, G. Living Extremely Flat: The Life of an Automaton; John von Neumann's Conception of Error of (in)Animate Systems. In: HON, G.; Schickore, J.; Steinle, F. **Going Amiss in Experimental Research**. Berlin: Springer, 2009. (Boston Studies In the Philosophy of Science, vol. 267), p. 55-71.

HOSKING, J.R.M. Fractional differencing. **Biometrika**, v. 68, p. 165-176, 1981.

HURST, H. Long Term Storage Capacity of Reservoirs. **Transactions of the American Society of Civil Engineers**, v. 116, p. 770-799, 1951.

HULL, J.C. **Risk Management and Financial Institutions**. 4th Edition. New York: Wiley, 2015.

JUMADINOVA, J.; MATACHE, M.T.; DASGUPTA, P. A Multi-Agent Prediction Market Based on Boolean Network Evolution. In: CONFERENCE ON WEB INTELLIGENCE AND INTELLIGENT AGENT TECHNOLOGY, Lyon, 2011. **Proceedings.** Lyon: IEEE/WIC/ACM, 2011.

KANCZUK, F. Um Termômetro para as Macro-Prudenciais. **Brazilian Economic Review**, v. 67, n. 4, p. 497–512, 2013.

______. Brazil through the eyes of CHORINHO. In: DSGE MODELS FOR BRAZIL: SAMBA AND BEYOND, São Paulo, 2014.. **Working Papers** ... São Paulo: FGV-EESP, 2014.

KIMBALL, M. S. Precautionary Saving in the Small and in the Large, **Econometrica**, v. 58, n. 1, p. 53–73, 1990.

KUMAR, D. **Long Memory in the Volatility of Indian Financial Market: An Empirical Analysis Based on Indian Data**. Hamburg: Anchor Academic Publishing, 2014.

KÜNSCH, H. R. Statistical Aspects of Self-similar Processes. In: WORLD CONGRESS OF THE BERNOULLI SOCIETY, 1., Tashkent, 1987. **Proceedings.** Utrecht: VNU Science Press, 1987.

LANGEVIN, M.S. **Brazil's Compounding Crisis**. Washington, D.C.: Council On Hemispheric Affairs, 2016. Disponível em: << http://www.coha.org/wp-content/uploads/2016/03/Brazil's-Compounding-Crisis_M.Langevin.pdf >>. Acesso em: 13 mai. 2016.

LELAND, H. Saving and Uncertainty: The Precautionary Demand for Saving, **The Quarterly Journal of Economics**, v. 82, p. 465–473, 1968

LÉVY, P. Random functions: General theory with special references to Laplacian random functions**. University of California Publications in Statistics**, 1, p. 331–390, 1953.

LIMA, L.R.; XIAO, Z. **Robustness of Stationarity Tests under Long-Memory Alternatives** - Ensaios Econômicos n. 776. São Paulo: Fundação Getúlio Vargas, 2004.

LO, A.W. Long-term memory in stock market prices. **Econometrica**, v. 59, p. 1279–1313, 1991.

LOVERRO, A. **Fractional Calculus: History, Definitions and Applications for the Engineer**. University of Notre Dame, 2004. Lecture Notes.

MACHADO, J.A.T.; SILVA, M.F.; BARBOSA, R.S.; JESUS, I.S.; REIS, C.M.; MARCOS, M.G.; GALHANO, A.F. Some Applications of Fractional Calculus in Engineering. **Mathematical Problems in Engineering**, v. 2010, article ID 639801, 34 pages, 2010.

MANDELBROT, B.B.; WALLIS, J.R. Computer experiments with a fractional Gaussian noises. **Water Resources Research**. v. 5, p. 67-228, 1969.

MAVROUDI, A. **Simulating city growth by using the Cellular Automata Algorithm**. 2007. 54 p. Master Thesis - University College London, London, 2007.

MCLEOD, A.L.; HIPEL, K.W. Preservation of the rescaled adjusted range. A reassessment of the Hurst phenomenon. **Water Resources Research**, v. 14, p. 491-508, 1978.

MORETTIN, P.A.; TOLÓI, C.M.C. **Análise de Séries Temporais**. Second Edition. São Paulo: Edgard Blücher, 2006.

MONTEIRO, L.H.A. **Sistemas Dinâmicos**. Second Edition. São Paulo: Livraria da Física, 2011.

______. **Sistemas Dinâmicos**. Third Edition. São Paulo: Livraria da Física, 2014.

NOWAK, A.; LEWENSTEIN, M. Modeling Social Change with Cellular Automata. In: HEGSELMANN, R.; MUELLER, U.; TROTIZSCH, K. G. **Modelling and Simulation in the Social Sciences from the Philosophy of Science Point of View**. Berlin: Springer, 1996. (Theory and Decision Library, v. 23), p. 249-285.

PFEIFER, B.; KUGLER, K.; TEJADA, M.M.; BAUMGARTNER, C.; SEGER, M.; OSL M.; NETZER, M.; HANDLER, M.; DANDER, A.; WURZ, M.; GRABER, A.; TILG, B. A Cellular Automaton Framework for Infectious Disease Spread Simulation. **Open Medics Informatics Journal**, v. 2, p. 70-81, 2008.

POOSEH, S.; RODRIGUES, H.S.; TORRES, D.F.M. Fractional derivatives in Dengue epidemics. In: NUMERICAL OPTIMIZATION AND APPLICATIONS SYMPOSIUM, Halkididi, 2011. **Proceedings** ... [S.l]. AIP Publishing, 2011.

ROBINSON, P. Semiparametric Analysis of Long-memory Time Series. **Annals of Statistics**, v. 22, p. 515-539, 1995.

ROSENBLUETH, D. A.; GERSHENSON, C. A model of city traffic based on elementary cellular automata. **Complex Systems**, v. 19, n. 4, p. 305-322, 2011.

ROSSI, E. **Multivariate GARCH Models. Financial Econometrics Class Material**. University of Pavia, 2010. Lecture Notes.

ROUHAUD, J. Cellular Automata and Consumer Behavior. **European Journal of Economic and Social Systems**, v. 14, n. 1, p. 37-52, 2000.

RUBINSTEIN, A. **Modeling bounded rationality**. Cambridge: MIT Press, 1998.

SALEM, J. B.; WOLFRAM, S. **Thermodynamics and Hydrodynamics of Cellular Automata**. 1985. Disponível em: <<http://www.stephenwolfram.com/publications/academic/thermodynamics-hydrodynamics-cellular-automata.pdf >>. Acesso em: 29 fev. 2014.

SELLA, R. **Three Essays in Econometrics: Multivariate Long Memory Time Series and Applying Regressions Trees to Longitudinal Data**. 2010. 432 p. Ph.D. Thesis - New York University, New York, 2010.

SMETS, R.; WOUTERS, F. Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach. **American Economic Review**, v. 97, n. 3, p. 586-606, 2007.

STOCK, J.H.; WATSON, M.W. Evidence on structural instability in macroeconomic time series relations. **Journal of Business and Economic Statistics**, v. 14, p. 11–30, 1996.

SYRIOPOULOS, T.; MAKRAM, B.; BOUBAKER, A. Stock market volatility spillovers and portfolio hedging: BRICS and the financial crisis. **International Review of Financial Analysis**, v. 39, issue C, p. 7-18, 2015.

TAQQU, M.S.; WILLINGER, W.; SHERMAN, R. Proof of a fundamental result in self-similartraffic modeling. **Computer Communication Review**, v. 27, p. 5–23, 1997.

TEVEROVERSUSKY, V.; TAQQU, M.S.; WILLINGER, W. A critical look at Lo’s modified R/S statistic. **Journal of Statistical Planning and Inference**, vol. 80, n. 1-2, p. 211-227, 1999.

TEYSSIÈRE, G.; KIRMAN, A.P. **Long Memory in Economics**. Berlin: Springer Science & Business Media, 2006.

VANCE, D. **Fractional Derivatives and Fractional Mechanics**. Washington University, 2014. Lecture Notes.

VILCARROMERO, A.C.S.; JAFELICE, R.S.M.; BARROS, L.C. Autômato Celular no Estudo de um Modelo Presa-Predador. **Biomatemática**, v. 20, p. 147-156, 2010.

WHITE, S.H.; DEL REY, A.M.; SÁNCHEZ, G.R. Modeling epicemics using cellular automata. **Applied Mathematics and Computation**, v. 186, p. 193-202, 2007.

WINKLER, A. **The Dual Role of Financial Markets in Economic Development: Engine of Growth and Source of Instability**, IPC Working Paper, 1998.

WOLFRAM, S. Statistical Mechanics of Cellular Automata. **Review of Modern Physics**, Vol. 55, n. 3, 601-644, 1983.

______. Twenty Problems in the Theory of Cellular Automata. **Physica Scripta**, v. T9, p. 170-183, 1985.

______. **A New Kind of Science**. Champaign: Wolfram Research, 2002.

XIAO, Z. Note on Bandwidth Selection in Testing for Long-Range Dependence. **Economic Letters**, v. 78, n. 1, p. 33-39, 2002.

ZAFFARONI, P. Contemporaneous aggregation of linear dynamic models in large economies. **Journal of Econometrics**, v. 120, n. 1., p. 75-102, 2004.