Università degli Studi di Cagliari

DOTTORATO DI RICERCA

SCIENZE ECONOMICHE ED AZIENDALI

Ciclo XXIX

TITOLO TESI

MARKOV PROCESSES IN FINANCE AND INSURANCE

Settore/i scientifico disciplinari di afferenza

SECS-S/06

Presentata da: ADA LIKA

Coordinatore Dottorato prof. ANDREA MELIS

Tutor prof. STEFANO USAI

Esame finale anno accademico 2015 – 2016

Tesi discussa nella sessione d'esame marzo – aprile 2017

Acknowledgements

First of all I would like to express my sincere gratitude to my advisor Filippo Petroni. You have been a real mentor for me. Your constant support and dedication has made me grow as a research scientist. I would like to thank my official supervisor prof. Stefano Usai and my co-supervisor prof. Marco Micocci for their patience and all their support. I am also very grateful to my two external evaluators for their valuable comments and suggestions. I would like to thank prof. Fabrizio Cesaroni, dott. Giovanni Batista Masala and prof. Micheal John Jones for serving as my committee members as well as the Board of Professors of the PhD Program for their efforts throughout these three years.

A special thanks goes to my family, my mother who has been a life coach for me, my father for always following my studies and showing interest in my research and my sister for always being there for me. You know me better than I know myself and I cannot express in words how fortunate and grateful I feel.

Last but by no mean least, I want to show my gratitude to my sweet half. Thank you for always supporting me in this journey, as in life, even though you don't always completely agree with my decisions.

Contents

Chapter 1

Introduction

Markov chains and, their more general version, semi-Markov processes have been widely used in the econophysics community because they provide a very flexible and efficient tool in modeling problems that cover up a wide range of different disciplines. Recently, there has been an increasing attention in the application of this type of models in finance and the actuarial sciences. This thesis aims to offer one more step in this direction. Two issues, one concerning the actuarial field and the other concerning the financial time series modeling, will be discussed.

Since their first introduction in the fifties, semi-Markov processes have had a rising attention of researchers as they offer a more general approach than the simple Markov chain. In the actuarial field, Markov and semi-Markov processes are very useful and many issues may be described by defining various states and estimating probabilities of transiting from one state to the other. In insurance, for instance, if we consider a simple model of a whole-life insurance, the states of a participant could be defined as $state\ 1 = Alive$ and $state\ 2 = Dead$ where the probability of transiting from one state to the other is the force of mortality.

A new branch based on Markov processes is also developing in the recent literature of financial time series modeling. This branch can be divided in two sub-branches that present two conceptually different approaches: one consists in using the Markov approach to introduce an element of switching between states of the market while in the other, a modified Markov process is constructed in order to consider implicitly the main stylized facts of financial time series.

This thesis, apart from this introduction, is organized in five more chapters. In Chapter 2 a brief introduction of stochastic processes and in particular of Markov chains is offered in order to define, also, the mathematical language and the basic definitions that will be used in the following chapters.

In Chapter 3 an overview of the literature on Markov processes applied to the actuarial sciences and in modeling financial time series is reported. A particular model, called the *Indexed semi-Markov chain* or *ISMC* is delineated. This model, in particular, belongs to the second sub branch mentioned above. The model is used to describe, through a Markov approach, the evolution in time of intra-day price returns of quoted firms. Two other more sophisticated versions of this model are also briefly introduced.

In Chapter 4, a first step in trying to model the salary evolution through a Markov approach is made. In particular, a semi-Markov model with backward recurrence time is proposed to describe salary lines of participants of an Italian Private Pension Scheme of the First Pillar. The semi-Markovian hypothesis is tested by applying a statistic test defined in the recent literature.

In Chapter 5 the ISMC is resumed and a new technique, based on the change point approach is proposed for the discretization of the Index process of the model. The theoretical framework is set and the methodology is

applied to high frequency price return dynamics of a quoted Italian firm.

In the last Chapter, the conclusions of the work are summarized.

Chapter 2

Markov Chains and semi-Markov models

2.1 Introduction to probability

In studying real world phenomenon, very often, the quantities we are interested in do not present a sure predictability for the future. In building a model to describe such phenomenon it is necessary to take into account of the possibility of randomness by attributing a probabilistic nature. This is why, such a model is referred to as a probability model.

2.1.1 Probability and conditional events

The most intuitive way to define probability is through its frequentist interpretation. In this approach, the probability of a random event denotes the relative frequency of occurrence of an experiment's outcome, when repeating the experiment a large number of times.

To better illustrate such concepts, let us suppose we are performing an experiment whose outcome is not predictable in advance as for example, the

roll of a fair dice. Although the outcome is not predictable, we know however the possible outcomes of such experiment.

Definition 1 *The set of all possible outcomes of an experiment is known as the sample space and is usually denoted with* Ω*. Each element of* Ω *is called an elementary event and is usually denoted with* ω*.*

In the experiment of the rolling of a dice we have the state space Ω:

$$\Omega = \{1, 2, 3, 4, 5, 6\}$$

made up of the six possible numbers of the face of the dice that are all the elementary events of the experiment.

Any subset of the sample space Ω is defined as an event and is usually denoted with E. In our example, a generic event might be defined as $E = \{2, 4, 6\}$ denoting the case that an even number appears after the roll of the dice.

Considering any two events E and F of a sample space Ω, we can define:

- $E \cup F$, the union event, that consists of all outcomes that are either in E and/or in F;
- $E \cap F$ or EF, the intersection event, that consists of all outcomes that are both in E and F;
- E^C, the complement event of E, that consists of all outcomes in the sample space that are not in E.

Let us now give a formal definition of probability for a given event E of Ω. First, we need to define a σ-algebra $\mathscr{A}$ that satisfies the following properties:

- $\emptyset \in \mathscr{A}$ - The empty set belongs to $\mathscr{A}$

- $E_1, E_2, \cdots \in \mathscr{A} \Rightarrow \bigcup_{i=1}^{\infty} B_i \in \mathscr{A}$ - $\mathscr{A}$ is closed under numerable union
- $B \in \mathscr{A} \Rightarrow E^C = \Omega | E \in \mathscr{A}$ - $\mathscr{A}$ is closed under complementation

Definition 2 *Given a state space Ω and the relative σ-algebra $\mathscr{A}$, a measure of probability $\mathbb{P}$ is a function that associates to each element of $\mathscr{A}$ a value in the interval $[0, 1]$:*

$$\mathbb{P} : \mathscr{A} \to [0, 1]$$

and satisfies the following properties:

- $0 \leq \mathbb{P}(E) \leq 1, \forall E \in \mathscr{A}$;
- $\mathbb{P}(\Omega) = 1$
- $A_1, A_2, \ldots, \in \mathscr{A}$, *is a sequence of mutually exclusive events,* $\Rightarrow \mathbb{P}(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} A_i$ *and this property holds for all finite subsets of* A_i.

Sometimes the probabilistic evaluation of an event E is subordinated to a specific set of information available H. This is a concept known as *conditional event* and is denoted by $E|H$. Let us suppose that H is an event for which $\mathbb{P}(H) > 0$. Then, the conditional probability of E given H, written $\mathbb{P}(E|H)$, is calculated considering that we already know that the event H is true:

$$\mathbb{P}(E|H) = \frac{\mathbb{P}(E \cap H)}{\mathbb{P}(H)}$$

That can be written as:

$$\mathbb{P}(E \cap H) = \mathbb{P}(E|H)\mathbb{P}(H)$$

or

$$\mathbb{P}(E \cap H) = \mathbb{P}(H|E)\mathbb{P}(E)$$

2.1.2 Random variables

In performing an experiment, often, what we are interested in is some function of the outcome and not the outcome itself. Consider, for example in the tossing of two dices we are often interested in the sum of the dices and not the actual outcome of each. This means that, for each possible realization $\omega \in \Omega$, it is possible to associate a number $X(\omega)$ that represents the numerical value of that particular realization. These quantities, are known as random variable.

In order to give a formal definition of a random variable, it is necessary to introduce the σ-algebra of Borel.

Definition 3 *The Borel σ-algebra of $\mathbb{R}$, usually denoted by $\mathscr{A}(\mathbb{R})$, is the smallest σ-algebra that can be obtained from all opened intervals of the real numbers through unions and intersections.*

Definition 4 *A real valued function X defined on Ω such that $X : \Omega \to \mathbb{R}$ is said to be a random variable if for every Borel set $\mathrm{B} \in \mathscr{B}(\mathbb{R})$, we have that*

$$X^{-1}(B) = (\omega : X(\omega) \in \mathrm{B}) \in \mathscr{A}$$

A random variable X, is called a *discrete* random variable if it can take a finite number of real values $x_1, x_2, \ldots$. In this case, it exists a function $\mathbb{P}_X(\cdot)$ defined as

$$\mathbb{P}_X(x) = \mathbb{P}(X = x)$$

This function is called the *probability mass function* and has the following properties:

- $0 \leq \mathbb{P}(X = x_i) \leq 1$;
- $\sum_{i=1}^{\infty} \mathbb{P}(X = x_i) = 1$

A particular and very important example of a discrete random variable, is the indicator function of a set $E \in \mathscr{A}$:

$$1_E(\omega) = \begin{cases} 1 & if\ \omega \in E \\ 0 & if\ \omega \notin E \end{cases}$$

A random variable X is called an *absolutely continuous* random variable if it can assume all the possible real values of an interval $[a, b]$, which can also be an open interval. In this case it exist a function f_X, known as the *probability density function* that has the following properties:

- $f(x) \geq 0, \forall x \in \mathbb{R}$;
- $\int_{\mathbb{R}} f(x)dx = 1$
- for each set $B \in \mathscr{B}(\mathbb{R})$

$$\mathbb{P}(X \in B) = \int_B f(x)dx$$

2.2 Basic notions on stochastic processes

A *stochastic process* $\{X_t, t \in T\}$ is a collection of random variables. That is, for each $t \in T$, X_t is a random variable defined on a probability space. The index t is often interpreted as time and as a consequence, X_t is usually referred to as the state of the process at time t. Formally:

Definition 5 *A stochastic process is a family of random variables* $X = \{X_t, t \in T\}$ *defined on the probability space* $(\Omega, \mathscr{A}, \mathbb{P})$.

The set T is called the *index set* of the process. When T is a countable set, the stochastic process is said to be a *discrete-time* process. If T is an interval of the real line, the stochastic process is said to be a *continuous-time* process.

The *state space* of a stochastic process is defined as the set of all possible values that the random variables X_t can assume. Let us denote the state space with S. It is possible that each random variable X_t has its own state space S_t. In this case, S will be defined as $S = \cup_{t \in T} S_t$.

Given that, the random variables are defined as measured functions from Ω to $\mathbb{R}$, a stochastic process can be interpreted as a collection:

$$\{X_t(\omega), \omega \in \Omega, t \in T\}$$

Fixed ω, if we vary $t \in T$, we obtain a trajectory of the stochastic process, while if we fix t and vary $\omega \in \Omega$ we have the random variable that describes the process at time t.

2.2.1 The Random Walk

The first and most simple example of a stochastic process is the random walk. An example of a simple random walk is the case of a particle that moves at

unit intervals, a step forward or a step backward with movements that are mutually independent.

Let us give a more formal definition of this process. For each unit of time $n = 1, 2, 3, \ldots$, let X_n be a dichotomous random variable that assumes value equal to 1 with a probability p and a value of -1 with a probability of $q = 1 - p$:

$$X_n = \begin{cases} +1 & p \\ -1 & q \end{cases}$$

Let us also assume that X_n are independent. This means that, at each instant of time n[1], the particle can move a step forward or backward with a probability of p and $q = 1 - p$, respectively. Moreover the movement of the particle at each unit of time does not depend on the previous or future moves.

From this random variable, it can be defined, for each unit of time n:

$$S_n = X_1 + X_2 + \cdots + X_{n-1} + X_n$$

This is also a random variable that represents the position (or level) of the particle after n steps. From the above definition, it can be easily seen that:

$$S_n = X_1 + X_2 + \cdots + X_{n-1} + X_n = S_{n-1} + X_n, \quad for\ n = 1, 2, 3, \ldots$$

The last equation shows how the random walk might be considered as a particular case of an homogeneous Markov Chain.

[1]For convention, we will use the index $t \in \mathbb{R}$ in the continuous case and the index $n \in \mathbb{N}$ when the observed time is a countable set.

2.3 Markov Chains and semi-Markov models

In this section we will present a general description of a particular class of stochastic processes. Their main characteristic is a specific type of dependence that will be presented shortly.

Let us consider a sequence of real random variables $X_0, X_1, \ldots, X_n, \ldots,$ that we indicate in a more compacted form as $\{X_n, n \geq 0\}$. Each X_n can assume values in a finite state space that we will denote with S.

Definition 6 (Markov Property) *A stochastic process $\{X_n, n \geq 0\}$, where X_n is defined in a finite set S, satisfies the Markov property when:*

$$\mathbb{P}(X_{k+n} = j|(X_n = i_n \cap X_{n-1} = i_{n-1} \cap \cdots \cap X_0 = i_0)) = \mathbb{P}(X_{n+k} = j|X_n = i_n)$$

for each n and k and for each state $j, i_0, \ldots, i_{n-1}, i_n \in S$

The Markov property states that the future state of the process X_{n+k} depends solely on the current value X_n and not on the previous values $X_0, X_1, \ldots, X_{n-1}$. This is commonly known as a "loss of memory" of the process.

Definition 7 (Markov Chain) *A discrete stochastic process $X_n, n \geq 0$, defined in a finite state space S, is a Markov Chain if it satisfies the Markov property such that, for each $n \geq 0$ and for $i_0, i_1, \ldots, \imath_n \in S$:*

$$\begin{aligned}\mathbb{P}(X_{n+k} = j|X_n = i_n, X_{n-1} = i_{n-1}, \ldots, X_0 = i_0) &= \mathbb{P}(X_{n+k}|X_n = i_n) \\ &= p_{ij}^{(k)}(n) \quad (2.1)\end{aligned}$$

where $p_{ij}^{(k)}(n)$ is called the *transition probability* and it denotes the probability of the process in state i at time n to go to the state j at time $n + k$.

A *finite-state Markov chain* is a Markov chain in which S is finite or countable. In Eq. 2.1, if we consider $k = 1$, we have the *one-step* transition probability that is usually denoted with $p_{ij}(n)$ instead of $p_{ij}^{(1)}(n)$. If we consider the transition probability to be independent of the time n such that:

$$p_{ij}(n) = p_{ij}$$

we have the *time homogeneous Markov chain.* This concept can be easily extended to the case of $k \geq 1$. In fact, if a Markov chain is time homogeneous, then the probability $p_{ij}^{(k)}(n)$, $\forall k > 1$, does not depend on time n.

If an homogeneous Markov chain has a finite number of states, the one step probability transition can be represented through a square stochastic matrix.

Definition 8 *Let $\{X_n, n \geq 0\}$ be a Markov chain with a finite state space. The transition probability matrix, denoted in general with P, is a stochastic matrix whose generic element is given by:*

$$P_{ij} = \mathbb{P}(X_n = j | X_{n-1} \in [0, 1]), \quad i, j \in S.$$

The transition matrix should satisfy the following properties:

- $P_{ij} \geq 0$ for $i, j \in S$
- $\sum_{j \in S} P_{ij} = 1$ for $i, j \in S$

The first property states that each element of the matrix is non negative while the second states that the sum of all the elements of each row is equal to 1. This condition expresses the fact that in a Markov chain, at each unit time n, there is necessary a transition to one of the states.

2.3.1 Transition probability after n steps

In the study of Markov chains, we are often interested in knowing the probability of passing from one state to another in a certain lapse of time i.e. after a certain number of steps k:

$$P_{ij}^{(k)} = \mathbb{P}(X_k = j | X_0 = i) \quad k = 2, 3, \ldots \tag{2.2}$$

The process can go from state i to state j by crossing different intermediate states. In this case it is useful the following equation:

Theorem 1 *(Chapman-Kolmogorov Equation) Let $\{X_n, n \geq 0\}$ be a Markov chain. For each pair of integers n and m and for each pair of states it results that:*

$$P_{ij}^{(n+m)} = \sum_{r \in S} P_{ir}^{(n)} P_{rj}^{(m)} \tag{2.3}$$

The Chapman-Kolmogorov equation, expressed in a matricial form, can be useful to show that in a time homogeneous Markov chain with a finite number of states, the transition matrix after n steps is equal to the n-th power of the transition matrix.

Theorem 2 *Let $\{X_n, n \geq 0\}$ be a time homogeneous Markov chain. It results that:*

$$P^{(k)} = P^k \tag{2.4}$$

From the Chapman-Kolmogorov equation, the following important properties can be derived:

- $P_{ij}^{(n+m)} \geq P_{ik}^{(n)} P_{kj}^{(m)} \quad \forall i, j, k \in S$
- $P_{ij}^{(n*m)} \geq \left[P_{ij}^{(n)}\right]^m \quad \forall n, m \geq 1 \quad \forall i, j, k \in S$

It is possible to calculate also the distribution of the process at a certain time unit $n \geq 0$. Let

$$\pi_0(i) \equiv \mathbb{P}(X_0 = i), \quad i \in S \left[\sum_{j \in S} \pi_0(j) = 1 \right]$$

be the initial state probability vector of the chain. We can define the probability at a generic time n the same way:

$$\pi_n(i) \equiv \mathbb{P}(X_n = i), \quad i \in S \left[\sum_{j \in S} \pi_(j) = 1 \right]$$

The probability distribution at time 0 can be used to obtain the transition probability after n steps through the following theorem:

Theorem 3 *Given an homogeneous Markov chain with a transition probability matrix P and initial distribution $\boldsymbol{\pi_0}$, it results that:*

$$\boldsymbol{\pi_n = \pi_0 P^n}$$

and, in general:

$$\boldsymbol{\pi_n = \pi_k P^{n-k}}$$

Often it is useful to analyze the asymptotic behavior of the process, i.e. $P_{ij}^{(n)}$, when $n \to \infty$. Let us first define an invariant [2] distribution for a Markov chain.

Definition 9 *A distribution π is an invariant distribution of a Markov chain if:*

$$\pi = \pi P$$

This means that, given an initial distribution π, after one step (or after n arbitrary steps) the marginal probability distribution is always π:

$$\boldsymbol{\pi P^n = \pi P P^{n-1} = \pi P^{n-1} = \cdots = \pi P = \pi} \tag{2.5}$$

[2]The terms equilibrium, stationary or steady state are sometimes used to mean the same thing

Chapter 3

Literature review on Markov processes in finance and insurance

3.1 Markov processes in the actuarial field

Markov chains (MC) and semi-Markov processes (SMP) have been used in a vast variety of disciplines such as physics, engineering, environmental sciences etc. Semi-Markov processes have been introduced in the fifties independently by Levy (1954), Takács (1954) and Smith (1953) and can be considered as a generalization of a Markov Process in which the waiting time distribution function can be of any type. This characteristic makes the SMP more flexible in addressing different problems. Recently, there has been an increasing attention in the application of Markov processes in finance and actuarial sciences.

Probably, the first works on semi-Markov models applied to the actuarial sciences are those of Janssen (1966), in which an Homogeneous semi-Markov

model (HSMP) is used, and Hoem (1969). In Iosifescu Manu (1972), the Non Homogeneous semi-Markov models (NHSMP) are defined and studied for the first time. In Hoem (1972), the author describes a number of demographic and actuarial examples which may be modeled through the time Non Homogeneous semi-Markov model. Further, in Janssen and De Dominicis (1984) the basic definitions and properties of non homogeneous semi-Markov processes with a finite number of states are given. The model is applied to the study of a social pension problem and, in particular, to the case of the calculation of premiums for insurance against professional diseases.

Other applications of semi-Markov in insurance problems are given in Sahin and Balcer (1979) and in Balcer and Sahin (1986). In the paper of Balcer and Sahin (1983), the author develops a stochastic theory to describe a pension scheme dynamics based on a semi-Markov reward process(HSMRP). Under a number of simplifying assumptions, the employment patterns are described by a two-state semi-Markov process identifying the individual as having (state 1) or not having (state 2) a coverage by a pension plan. The reward function is then constructed in order to convert pensionable service to pension benefits and pension costs. A number of basic functions are constructed and analyzed in terms of the impact on lifetime pension benefits to plan types, vesting rules, coverage rate and portability.

In Janssen and Manca (1997), it is presented for the first time the Generalized Discrete Time Non Homogeneous semi-Markov Reward (GDTNHSMRP) model for pension funds in which, by introducing two time variable, it was possible to consider simultaneously the age and seniority. The authors, through this stochastic time model, take into account economic, financial and demographic evolution factors. The choice of discrete time is justified by the fact that the management of a pension fund is mainly on a yearly basis, the

non homogeneity is observed on the age and seniority of the member since all the other parameters are age dependent and lastly, the semi-Markov model is chosen because transition states, transition times, costs and contributions have to be modeled simultaneously. The non-homogeneous reward processes were defined, with an application in health insurance, in De Dominicis and Manca (1986).

These first works have been further developed in Janssen and Manca (2002, 2006) where it is shown how to apply continuous time semi-Markov reward processes in multiple life insurance. In Stenberg et al. (2006b), the previous results are further improved by developing a method for the calculation of not only the expectations but also of higher order moments for the accumulated rewards. The method is applied to a database of persons with silicosis problems. Six states are defined according to the disability degree and transition between states are supposed to occur after a visit to the doctor. Also, a technique for checking the "semi-Markov", i.e. that the sojourn times are not geometrically distributed, is introduced.

NHSMP transition probabilities can be generalized by using a backward environment, i.e. the transition probabilities are also dependent on the time of entrance into a given state. An application of these models in insurance can be found in Stenberg et al. (2006a, 2007). In D'Amico et al. (2009a) it is presented for the first time a general formula of a discrete time NHSMP with initial and final backward stochastic processes. The model is applied to the study of a sample of insured population with a policy of LTC (Long-Term Care). A complete study of continuous time homogeneous backward semi-Markov process was reported in Howard (1971) and successively in Limnios and Oprisan (2012). Further, in D'amico et al. (2013) it is presented a reward semi-Markov model with an initial and final backward recurrence applied

to the study of disability. The reward structure permits to determine the equations for the prospective and retrospective mathematical reserves.

3.2 Markov models in finance

The use of Markov processes in modeling financial time series can be sub divided in two main branches. The first one uses a switching between states technique while the second one is based on models that endogenously consider stylized facts of time series.

Since the first paper of Hamilton (1989) suggested a Markov switching technique as a method for modeling non-stationary time series, this approach has successively been widely used in the study of stocks and foreign exchange markets analysis. The first to employ a Markov-Switching autoregressive process (MS-AR) have been Turner et al. (1989) and Chu et al. (1996). In the first paper, a model of the stock market in which the excess return is drawn from a mixture of two normal densities is introduced. The market is assumed to switch between two states, which in turn determine which of the two normal distributions is used to generate the excess return for the specific period. The state is assumed to be generated by a first-order Markov process. In the second paper, the linkage between the stock market volatility and regime shifts in the stock market returns is analyzed. A Markov-Switching model is applied to the time series of monthly stock market returns and variation in volatility in different return regimes is examined. The authors find that the stock returns are best characterized by a model that contains six regimes. They find a non linear and asymmetric relationship between returns and volatility, i.e. the volatility is higher in negative return regimes than in positive return regimes.

This approach has been further developed in Schaller and Norden (1997). The authors find a strong evidence of switching behavior in stock market returns for a variety of different specifications (more precisely they considered switching in means, in variance and in both mean and variance). By apply-

ing new test to the results of Turner et al. (1989), they show that the new hypothesis of no switching can be rejected in the period since World War II. In Nishiyama (1998), a Markov-switching model is applied to examine discrete shifts in regime in equity markets from 1976 to 1991 in five countries (Germany, Japan, USA, United Kingdom and Canada). The author concludes that stock returns are time varying, that the conditional variance is state-dependent and that markets vary widely in frequency of regime shifts and persistence.

Other application can be found in Ismail and Isa (2008), which uses a two regime MS-AR model to capture regime-switching behavior in both the mean and the variance of the Malaysian equity market. They conclude that the MS-AR model is able to capture the timing of regime shifts occurring during the period 1974–2003 and generated by the 1974 oil shock, 1987 stock market crash and the Asian financial crisis of 1997. Wang and Theobald (2008) implement an MS-AR model to investigate the regime switching behavior of six Asian emerging stock markets over the period 1970–2004. They conclude that there are two or three volatility states for the selected stock markets. Moreover, they find that switching between regimes is associated with international, as well as country specific, events that lead to fluctuating levels of confidence within these markets. Using a similar approach, Moore and Wang (2007) study stock returns for new EU member states of Czech Republic, Hungary, Poland, Slovenia and Slovakia. Evidence is found that the entry to EU appears to be associated with a reduction of volatility in unstable emerging markets.

Other authors proposed more advanced techniques. In Cai (1994), a Markov-ARCH model which incorporates the features of both Markov switching regime model and ARCH model is developed. The model is used to ex-

amine the issue of volatility persistence in the monthly excess returns of the three-month T-bill using the period 1964-1991. This model has been used also in Hamilton and Susmel (1994), in Edwards and Susmel (2003) and in Brunetti et al. (2008). In the latter, exchange rate turmoil are analyzed by using a Markov-switching GARCH model. The authors distinguish between two different regimes: a "calm" regime characterized by low exchange rate changes and low volatility, and a "turbulent" regime characterized by high exchange rate devaluation and high volatility.

In Kanas (2005) a Markov Switching Vector Autoregressive model (MS-VAR) is used to explore whether there are linkages in volatility regimes between the Mexican currency market and six emerging equity markets, namely, Mexico, Brasil, Argentina, Thailand, Hong Kong, and Hungary. The authors conclude that the MS-VAR-based correlation coefficients between the Mexican currency market and each equity market do not result significantly higher during the high-volatility regime, thereby indicating that the volatility regime dependence can be interpreted as evidence of interdependence rather than contagion. Kanas and Kouretas (2007) examine the short-run and long-run relationship between the parallel and official markets for US dollars in Greece using a bivariate Markov Switching Vector Error Correction Model (MS-VECM).

In Henry (2009) the impact of volatility in the London interbank market on equity returns with data observed over the period January 1980 - August 2007. The model used for equity returns is a two regime Markov-Switching Exponential GARCH model in which one regime is consistent with a high-mean, low variance state and the other regime appears to be a low-mean, high variance state. In the first case, which appears to be the dominant one, volatility responds persistently but symmetrically to news while in the

second one the conditional variance of returns responds asymmetrically and without persistence to news.

In more recent studies, Markov-switching regimes methodology has been used in Mandilaras and Bird (2010) to model movements and contagion in the foreign exchange markets of the European Monetary System. The authors find that most foreign exchange markets correlations increase during the crisis state. In order to detect contagion in the EMS, they considered Denmark and Italy as the source countries (Denmark for rejecting the Maastricht Treaty and Italy because of accumulated competitiveness problems). They found that the Denmark indeed exported its volatility to Ireland and Belgium and that contagion effects appear to exist between Italy and Denmark as well. In Walid et al. (2011) a Markov-switching EGARCH model is used to investigate the dynamic linkage between stock price volatility and exchange rate changes for four emerging countries over the period 1994-2009 by using weekly data. The authors conclude that there is a strong evidence that the relationship between stock and foreign exchange markets is regime dependent.

As mentioned, the models described so far belong to the first sub branch of a more general approach based on Markov processes for financial series. In the following paragraphs models that belong to the second sub-branch will be briefly presented while in the next section a particular model of this type will be described in details.

The first application of semi-Markov processes in the theory of Option pricing can be found in Janssen et al. (1997). The authors show a new model as an alternative of the classical CRR (Cox-Ross-Rubinstein) by using a discrete time Markov approach and a finite number of possible values of the embedded asset. The case of an European and an American option are considered. Also, the authors consider both the case in which the un-

derlying asset follows an homogeneous Markov chain and an homogeneous semi-Markov process, the latter considered both in the continuous and in the discrete time. In Kijima (2013), an extension of the Binomial Model to include the possibility that the transition probabilities are state dependent is considered. In D'Amico et al. (2009b) it is assumed that the log returns of the underlying asset follows a discrete time and finite state space HSMP. The values of European and American options, as well as their bare risks, are determined. The prices and risks obtained depend explicitly on the waiting-time distributions of the asset and they are duration dependent.

The possibility of using the semi-Markov approach for the study of intra-day asset price dynamics has been hypothesized in D'Amico and Petroni (2012a). The authors examined the possibility of using a semi-Markov model for the study of high frequency financial data. In particular, a time homogeneous semi-Markov process was used to model the dynamics of intra-day price returns while a Markov-chain was used to model the overnight returns. The authors showed that these models are able to reproduce some stylized empirical facts such as for example the absence of autocorrelations in returns and the gain/loss asymmetry. In the paper, the authors also showed that the autocorrelation in the square of returns was higher compared to the Markov model but it was still too small compared to the empirical one.

In D'Amico and Petroni (2011), a new model, called the "Indexed semi-Markov Model" or ISMC, is proposed for modeling asset price returns. More precisely, the intra-day returns (up to one minute frequency) are described by a discrete time homogeneous semi-Markov process where a memory index is introduced in order to take account of market volatility. The authors hypothesize that the kernel of the semi-Markov process depends on the level of volatility in the market. The index process was defined as a moving average

of the reward accumulation process linked to the Markov Renewal Process of the price returns. The authors concluded that the semi-Markov kernel is influenced by the past volatility and that the model was able to reproduce quite well the behavior of market returns: uncorrelated returns and long correlation of the square of returns.

Since that first work, the ISMC model has been improved further. In D'Amico and Petroni (2012b) the Weighted-Index semi-Markov model was introduced where the Index process was calculated as an exponentially weighted index while in D'Amico and Petroni (2014) a bivariate model where the indexed semi-Markov process is used to reproduce simultaneously two stocks is applied.

In the following Section, the main finding of the latter works will be presented and the ISMC model will be described in details.

3.3 The model for price returns

In the following section we will briefly summarize the main characteristics of the Markov model proposed in D'Amico and Petroni (2012a) for the modeling of price returns. The authors hypothesize that the process of asset price returns, and in particular intra-day price returns, can be described through the semi-Markov approach while the overnight returns can be described by a discrete time homogeneous Markov chain. This different choice is supported by the fact that the intra-day returns are determined while the market is opened and prices can be influenced almost immediately by a possible new flow of information while for the later (the overnight returns) the accumulated flow of information during the closing time of the market is reflected in the opening price of the successive day.

If $S(t)$ is assumed to describe the process of asset prices of quoted firms then let $Z(t)$ be the intra-day return at time t while we let $X(t)$ be the overnight return. The time variable $t \in \{0, 1, \ldots, nd\}$ where n is the number of unit periods during the day (i.e. minutes) and d is the number of days. Then, the process of price returns can be defined as follows:

$$W(t) = \begin{cases} Z(t) & \text{if } (k-1)n < t < nk \\ X(t) & \text{if } t = nk \end{cases} \tag{3.1}$$

where $k = 1, 2, ..., d$
and $Z(t) = \frac{S(t+1)-S(t)}{S(t)}$ is calculated over a time interval of length 1,
also, $X(t) = \frac{S(t+1)-S(t)}{S(t)}$ for $t = nk$, $k = 1, 2, ..., d$.

As mentioned above, we let $Z(t)$ be a discrete time semi-Markov model defined in the finite state space

$$E = \{-z_{min}\Delta, \ldots, -2\Delta, -\Delta, 0, \Delta, 2\Delta, \ldots, z_{max}\Delta\}$$

and kernel $\mathbf{b} = (b_{ij}(\gamma))$, $\forall i, j \in E$ and $\gamma \in \mathbb{N}$.

The kernel has the following probabilistic interpretation:

$$\begin{aligned} &P[Z_{n+1} = z, T_{n+1} - T_n \leq \gamma | \sigma(Z_h, T_h),\, h \leq \gamma, Z_n = i] = \\ &P[Z_{n+1} = z, T_{n+1} - T_n \leq \gamma | Z_n = i] = b_{ij}(\gamma) \end{aligned} \tag{3.2}$$

and it results $p_{ij} = \lim_{\gamma \to \infty} b_{ij}(\gamma)$; $i, j \in E$, $\gamma \in \mathbb{N}$ where $\mathbf{P} = (p_{ij})$ is the transition probability matrix of the embedded Markov chain $Z(n)$.

On the other hand, we consider $X(t)$ as a discrete time homogeneous Markov chain with the same state space and transition probability matrix $\mathbf{T} = (t_{i,j})_{i,j \in E}$.

Let $M_t(\tau)$ be the accumulation factor from t to $t + \tau$ and it is given by

$$M_t(\tau) = \prod_{k=0}^{\tau-1} (1 + W(t + k)) \,. \tag{3.3}$$

which takes value in the set

$$SP_\tau = \{x \in \mathbb{R} : x = \prod_{k=0}^{\tau-1} (1 + i(t + k)) \,, i(t + k) \in E\}.$$

The relation between $M_t(\tau)$ and the price $S(t)$ is given by the following equation:

$$M_t(\tau) = \frac{S(t + \tau)}{S(t)}. \tag{3.4}$$

In order to compare results obtained through the Markov modeling and real data two characteristics have been considered: the first passage time distribution (ftp) and the autocorrelation function.

We recall that the first passage time for an investment made at time t at price $S(t)$, is defined as the time interval $\tau = t' - t$, $t' > t$, where the relation $M_t(\tau) \geq \rho$ is fulfilled for the first time. We will denote the first passage time as $\lambda_\rho(t)$. Then

$$\lambda_\rho(t) = \min\{\tau \geq 0; M_t(\tau) \geq \rho\}.$$

Since it is assumed that the semi-Markov process $Z(t)$ is time homogeneous, the first passage time can be denoted with $\lambda_\rho(t) = \lambda_\rho$. For each time t, let

$$R_i(v, t||\rho) = P(\lambda_\rho > t|(i, v))$$

where $i \in E$ denotes the state of the return and $v \in \mathbb{N}$ the time length of being in this state, both of them considered at time zero.

Let us define by $R_{i,j}(v, t||w, \rho)$, $\forall w \in SP_t$, $\forall \rho \in \mathbb{R}$, the probability

$$P(\lambda_\rho > t, W(t) = j, M_0(t+1) = w|(i, v)),$$

obviously

$$R_i(v, t||\rho) = \sum_{j \in E} \sum_{x \in SP_t, x < \rho} R_{i,j}(v, t||x, \rho). \tag{3.5}$$

The following equations for finding the first passage time distribution have been derived [1]. First, let us define the distribution function

$$H_i(t) = \mathbb{P}[T_{n+1} - T_n \leq t|Z_n = i] = \sum_{j \in E} b_{ij}(t) \tag{3.6}$$

which represents the unconditional waiting time distribution in state i.

Four different cases are identified. In the first case, the first passage time distribution $R_{i,j}(v, T||w, \rho)$ for time t belonging to the first day is determined as:

$$\begin{aligned} R_{i,j}(v, t||w, \rho) &= 1_{\{(1+i\Delta)^t < \rho\}} \delta_{ij} \left(\frac{1 - H_i(t+v)}{1 - H_i(v)} \right) + \\ &\sum_{a \in E} \sum_{m=1}^{t} \frac{b_{ia}(v+m)}{1 - H_i(v)} 1_{\{(1+i\Delta)^m < \rho\}} \cdot \\ &R_{a,j} \left(0, t - m || \frac{w}{(1+i\Delta)^m}, \frac{\rho}{(1+i\Delta)^m} \right). \end{aligned} \tag{3.7}$$

In the second case, the opening of the second day is considered and thus we have $t = n$. In this situation we should take care for the transition at

[1] For a more detailed description and for the proofs of the equations the reader can refer to D'Amico and Petroni (2012a)

time $t = n$ which is due to the Markov chain $X(t)$. To obtain a formula for the first passage time distribution until time n it is sufficient to consider all possible states for the return and for the accumulation factor at time $n-1$ and to use equation (3.7).

$$R_{i,j}(v, n||w, \rho) = \sum_{\overline{w} \in SP^{\rho}_{n-1}} \sum_{a \in E} t_{a,j} 1_{\{(1+j\Delta)=\frac{w}{\overline{w}}\}} R_{i,a}(v, n-1||\overline{w}, \rho) \tag{3.8}$$

By similar computations it is possible to have the first passage time distribution for time $t = nd$. The relation is the following:

$$R_{i,j}(v, nd||w, \rho) = \sum_{a \in E} \sum_{\overline{w} \in SP_{(n-1)d}}^{\rho} R_{i,a}(v, (n-1)d||\overline{w}, \rho) R_{a,j}(0, n||\frac{w}{\overline{w}}, \frac{\rho}{\overline{w}}). \tag{3.9}$$

Formula (3.9) is obtained by conditioning on all possible states of the return process $W(t)$ and on all possible values of the accumulation factor $M_0(t)$ at time $t = (n-1)d$ (the closing of day $n-1$).
The last case, when $(n-1)d < t < nd$, can be obtained by using jointly the formulas (3.7) and (3.9). The resulting relation is the following:

$$\begin{aligned} &R_{i,j}(v, t||w, \rho) = \\ &\sum_{a \in E} \sum_{\overline{w} \in SP_{(n-1)d}}^{\rho} R_{i,a}(v, (n-1)d||\overline{w}, \rho) R_{a,j}(0, t-(n-1)d||\frac{w}{\overline{w}}, \frac{\rho}{\overline{w}}). \end{aligned} \tag{3.10}$$

Formula (3.10) is obtained by conditioning on the states of the return process and on the values of the accumulation factor process at time $t = (n-1)d$ and then by using formula (3.7).

Formulae (3.7), (3.8), (3.9) and (3.10) allow us to compute the probability $R_{i,j}(v, t||w, \rho)$ for all times t. It should be noted that if ρ is not too big, it is highly probable that the accumulation factor process exceeds ρ within the day. In this case probabilities (3.8), (3.9) and (3.10) will be equal to zero. Consequently, the first passage time distribution will have non zero values

only for $1 \leq t \leq n-1$. In this case (3.7) satisfies the following simpler equation:

$$R_i(v,t||\rho) = 1_{\{(1+i\Delta)^t<\rho\}}\left(\frac{1-H_i(t+v)}{1-H_i(v)}\right)+ \\ \sum_{a\in E}\sum_{m=1}^{t}\frac{b_{ia}(v+m)}{1-H_i(v)}1_{\{(1+i\Delta)^m<\rho\}}R_a(0,t-m||\frac{\rho}{(1+i\Delta)^m}). \tag{3.11}$$

which is obtained from (3.7) through relation (3.5).

3.3.1 The Indexed semi-Markov model

In D'Amico and Petroni (2011), the authors propose a new model called Indexed semi-Markov model (or ISMC) as a generalization of the semi-Markov process that is able to represent higher-order dependencies between successive observations of a state variable. Models based on Markov processes that have the characteristic of longer memory have previously been defined in Limnios and Oprisan (2003) where high-order semi-Markov processes have been used. ISMC on the other hand is presented as a more parsimonious model. This new model should better reproduce the long-term dependence in the square of returns of financial time series. It seems useful to describe in detail the main characteristics of this model.

Let $(\Omega, \mathbf{F}, P)$ be a probability space and let us consider the stochastic process:

$$J_{-(m+1)}, J_{-m}, J_{-(m-1)}, ..., J_{-1}, J_0, J_1, ...$$

with a finite state space $E = \{1, 2, ..., S\}$. Let us also consider:

$$T_{-(m+1)}, T_{-m}, T_{-(m-1)}, ..., T_{-1}, T_0, T_1, ...$$

Let J_n describe the price return process at the n-th transition and let T_n describe the time in which the n-th transition of the price return process occurs.

Also we define another stochastic process:

$$U_{-(m+1)}, U_{-m}, U_{-(m-1)}, ..., U_{-1}, U_0, U_1, ...$$

with values in $\mathbb{R}$. The random variable U_n describes the value of the index process at the n-th transition. U_n^m is defined as follows:

$$U_n^m = \frac{1}{T_n - T_{n-(m+1)}} \sum_{k=0}^{m} \int_{T_{n-1-k}}^{T_{n-k}} f(J_{n-1-k}, s) ds, \tag{3.12}$$

where $f : E \times \mathbb{R} \to \mathbb{R}$ is a Borel measurable bounded function and $U_{-(m+1)}^m, ..., U_0^m$ are known and non-random.

The process U_n^m can be interpreted as a moving average of the accumulated reward process with the function f as a measure of the rate of reward per unit time. The function f depends on the state of the system J_{n-1-k} and on the time s.

To better explain the construction of the index process, let us consider the simplest case where $m = 1$ and $f(J_n, s) = (J_n)^2$, i.e. the function f is equal to the square of returns. We have that:

$$U_n^1 = \frac{1}{T_n - T_{n-2}} \Big((J_{n-1})^2 \cdot (T_n - T_{n-1}) + (J_{n-2})^2 \cdot (T_{n-1} - T_{n-2}) \Big), \tag{3.13}$$

which expresses a moving average of order $m + 1 = 2$ executed on the series of the square of returns with weights given by the fractions

$$\frac{T_n - T_{n-1}}{T_n - T_{n-2}}; \quad \frac{T_{n-1} - T_{n-2}}{T_n - T_{n-2}}. \tag{3.14}$$

To construct an indexed model, a dependence structure between the variables has to be specified:

$$\begin{aligned} &P[J_{n+1} = j,\, T_{n+1} - T_n \leq t | \sigma(J_h, T_h, U_h^m),\, h = -m, ..., 0, ..., n, J_n = i, U_n^m = v] \\ &= P[J_{n+1} = j,\, T_{n+1} - T_n \leq t | J_n = i, U_n^m = v] := Q_{ij}^m(v; t), \end{aligned} \tag{3.15}$$

where $\sigma(J_h, T_h, U_h^m)$, $h \leq n$ is the natural filtration of the three-variate process.

The matrix of functions $\mathbf{Q}^m(v;t) = (Q_{ij}^m(v;t))_{i,j\in E}$ is called the *indexed semi-Markov kernel*.

The joint process (J_n, T_n), which is embedded in the indexed semi-Markov kernel, depends on the moving average process U_n^m which acts as a stochastic index. Moreover, the index process U_n^m depends on (J_n, T_n) through the functional relationship (3.12).

Given the three-dimensional process $\{J_n, T_n, U_n^m\}$ and the indexed semi-Markov kernel $\mathbf{Q}^m(v;t)$, let us define:

$$\begin{aligned} &N(t) = \sup\{n \in \mathbb{N} : T_n \leq t\}; \\ &Z(t) = J_{N(t)}; \\ &U^m(t) = \frac{1}{t - T_{(N(t)-\theta)-m}} \sum_{k=0}^{m} \int_{T_{(N(t)-\theta)-k}}^{t\wedge T_{(N(t)-\theta)+1-k}} f(J_{(N(t)-\theta)-k}, s)ds, \end{aligned} \tag{3.16}$$

where $T_{N(t)} \leq t < T_{N(t)+1}$ and $\theta = 1_{\{t=T_{N(t)}\}}$.

The stochastic processes defined in (3.16) represents, respectively:

- The number of transitions up to time t
- The state of the system (price return) at time t
- The value of the index process (moving average of a function of price return) up to t

The process $Z(t)$ defined above is referred to as an indexed semi-Markov process. The process $U^m(t)$ is a generalization of the process U_n^m where time t can be a transition or a waiting time.

It is simple to realize that $\forall m$, if $t = T_n$ we have that $U^m(t) = U_n^m$. Let, then

$$p_{ij}^m(v) := P[J_{n+1} = j | J_n = i, U_n^m = v].$$

be the transition probabilities of the embedded indexed Markov chain. It denotes the probability that the next transition is in state j given that at current time the process entered in state i and the index process is v. It is simple to realize that

$$p_{ij}^m(v) = \lim_{t\to\infty} Q_{ij}^m(v;t). \tag{3.17}$$

Let $H_i^m(v;\cdot)$ be the sojourn time cumulative distribution in state $i \in E$:

$$H_i^m(v;t) := P[T_{n+1} - T_n \leq t | J_n = i,\, U_n^m = v] = \sum_{j\in E} Q_{ij}^m(v;t). \tag{3.18}$$

It expresses the probability to make a transition from state i with sojourn time less or equal to t given the indexed process is v. The conditional waiting time distribution function G expresses the following probability:

$$G_{ij}^m(v;t) := P[T_{n+1} - T_n \leq t \mid J_n = i, J_{n+1} = j, U_n^m = v]. \tag{3.19}$$

It is simple to establish that

$$G_{ij}^m(v;t) = \begin{cases} \frac{Q_{ij}^m(v;t)}{p_{ij}^m(v)} & \text{if } p_{ij}^m(v) \neq 0 \\ 1 & \text{if } p_{ij}^m(v) = 0. \end{cases} \tag{3.20}$$

To properly assess the probabilistic behavior of the system, we introduce the transition probability function:

$$\phi^m_{(i_{-(m+1)},i_{-m},\ldots i_0;j)}(t_{-(m+1)}, t_{-m}, ..., t_0; t, V) :=$$
$$P[Z(t) = j, U^m(t) \leq V | J_0 = i_0, ..., J_{-(m+1)} = i_{-(m+1)}, T_0 = t_0, ..., T_{-(m+1)} = t_{-(m+1)}]. \tag{3.21}$$

In D'Amico and Petroni (2011) the authors show that the transition probability function of the indexed semi-Markov process satisfies a renewal-type equation[2].

[2]For a more detailed description of the propositions and proofs refer to the cited paper.

3.3.2 The WISMC and the Bivariate ISMC

The model described in the above paragraph has been further improved in two successive papers, D'Amico and Petroni (2012b) and D'Amico and Petroni (2014). In this section, the main characteristics and the most relevant improvements achieved in these two works will be presented.

Firstly, let us describe the Weighted Indexed semi-Markov Chain (or WISMC). Let us recall that the main objective of these models is to reproduce the long-term dependence in the square of returns in the most efficient way. This, as mentioned above, is an important characteristic of financial time series.

Let us assume that the value of the financial asset under study is described by the time varying asset price $S(t)$. The return at time t calculated over a time interval of length 1 is defined as $\frac{S(t+1)-S(t)}{S(t)}$. The return process changes value in time, then we denote by $\{J_n\}_{n\in\mathbb{N}}$ the stochastic process with finite state space $E = \{1, 2, ..., s\}$ and we assume that it describes the value of the return process at the n-th transition.

Let us consider the stochastic process $\{T_n\}_{n\in\mathbb{N}}$ with values in $\mathbb{N}$ that describes the time in which the n-th transition of the price return process occurs.

Let us consider also the stochastic process $\{U_n^\lambda\}_{n\in\mathbb{N}}$ with values in $\mathbb{R}$. As defined in the above paragraph, the random variable U_n^λ describes the value of the index process at the n-th transition.

We recall that, in case of the ISMC, the process $\{U_n\}$ was defined as a moving average of the reward process. Here, motivated by the application to financial returns, we consider a more flexible index process defined as follows:

$$U_n^\lambda = \sum_{k=0}^{n-1} \sum_{a=T_{n-1-k}}^{T_{n-k}-1} f(J_{n-1-k}, a, \lambda), \tag{3.22}$$

where $f : E \times \mathbb{N} \times \mathbb{R} \to \mathbb{R}$ is a Borel measurable bounded function and U_0^λ is known and non-random.

The process U_n^λ can be interpreted as an accumulated reward process with the function f as a measure of the weighted rate of reward per unit time. The function f depends on the current time a, on the state J_{n-1-k} visited at current time and on the parameter λ that represents the weight.

In next section a specific functional form of f will be selected in order to produce a real data application.

To construct the WISMC model, we have to specify a dependence structure between the variables. Toward this end we adopt the following assumption:

$$\begin{aligned} &\mathbb{P}[J_{n+1} = j,\ T_{n+1} - T_n \leq t | \sigma(J_h, T_h, U_h^\lambda),\ h = 0, ..., n, J_n = i, U_n^\lambda = v] \\ &= \mathbb{P}[J_{n+1} = j,\ T_{n+1} - T_n \leq t | J_n = i, U_n^\lambda = v] := Q_{ij}^\lambda(v; t), \end{aligned} \tag{3.23}$$

where $\sigma(J_h, T_h, U_h^\lambda)$, $h \leq n$ is the natural filtration of the three-variate process.

The matrix of functions $\mathbf{Q}^\lambda(v; t) = (Q_{ij}^\lambda(v; t))_{i,j \in E}$ is called *weighted-indexed semi-Markov kernel*.

The joint process (J_n, T_n) depends on the process U_n^λ, the latter acts as a stochastic index. Moreover, the index process U_n^λ depends on (J_n, T_n) through the functional relationship (3.12).

Observe that if

$$\mathbb{P}[J_{n+1} = j, T_{n+1} - T_n \leq t | J_n = i, U_n^\lambda = v] = \mathbb{P}[J_{n+1} = j, T_{n+1} - T_n \leq t | J_n = i]$$

for all values $v \in \mathbb{R}$ of the index process, then the weigthed indexed semi-

Markov kernel degenerates in an ordinary semi-Markov kernel and the WISMC model becomes equivalent to classical semi-Markov chain model.

The triple of processes $\{J_n, T_n, U_n^\lambda\}$ describes the behavior of the system only in correspondence of the transition times T_n. To describe the behavior of our model at whatever time t which can be a transition time or a waiting time, we need to define additional stochastic processes.

Given the three-dimensional process $\{J_n, T_n, U_n^\lambda\}$ and the weighted indexed semi-Markov kernel $\mathbf{Q}^\lambda(v;t)$, we define by

$$\begin{aligned} N(t) &= \sup\{n \in \mathbb{N} : T_n \leq t\}; \\ Z(t) &= J_{N(t)}; \\ U^\lambda(t) &= \sum_{k=0}^{N(t)-1+\theta} \sum_{a=T_{N(t)+\theta-1-k}}^{(t\wedge T_{N(t)+\theta-k})-1} f(J_{N(t)+\theta-1-k}, a, \lambda), \end{aligned} \tag{3.24}$$

where $\theta = 1_{\{t>T_{N(t)}\}}$.

As seen in the case of ISMC, the stochastic processes defined in (3.24) represent the number of transitions up to time t, the state of the system (price return) at time t and the value of the index process (weighted moving average of function of price return) up to t, respectively. We refer to $Z(t)$ as a weighted indexed semi-Markov process.

The process $U^\lambda(t)$ is a generalization of the process U_n^λ where time t can be a transition or a waiting time. It is simple to realize that if $t = T_n$ we have that $U^\lambda(t) = U_n^\lambda$.

Let

$$p_{ij}^\lambda(v) := \mathbb{P}[J_{n+1} = j | J_n = i, U_n^\lambda = v],$$

be the transition probabilities of the embedded indexed Markov chain. It denotes the probability that the next transition is in state j given that at current time the process entered in state i and the index process is equal to

v. It is simple to realize that

$$p_{ij}^{\lambda}(v) = \lim_{t\to\infty} Q_{ij}^{\lambda}(v;t). \tag{3.25}$$

Let $H_i^{\lambda}(v;\cdot)$ be the sojourn time cumulative distribution in state $i \in E$:

$$H_i^{\lambda}(v;t) := \mathbb{P}[T_{n+1} - T_n \leq t | J_n = i,\, U_n^{\lambda} = v] = \sum_{j\in E} Q_{ij}^{\lambda}(v;t). \tag{3.26}$$

It expresses the probability to make a transition from state i with sojourn time less or equal to t given the indexed process is v.

The conditional waiting time distribution function G expresses the following probability:

$$G_{ij}^{\lambda}(v;t) := \mathbb{P}[T_{n+1} - T_n \leq t \mid J_n = i, J_{n+1} = j, U_n^{\lambda} = v]. \tag{3.27}$$

It is simple to establish that

$$G_{ij}^{\lambda}(v;t) = \begin{cases} \frac{Q_{ij}^{\lambda}(v;t)}{p_{ij}^{\lambda}(v)} & \text{if } p_{ij}^{\lambda}(v) \neq 0 \\ 1 & \text{if } p_{ij}^{\lambda}(v) = 0. \end{cases} \tag{3.28}$$

It is possible to derive for the WISMC model the same results in terms of an explicit renewal type equation such as the ones submitted to describe the probabilistic behavior of ISMC.

In D'Amico and Petroni (2014), the WISMC model is extended to the bivariate case. The results obtained can easily be extended to the case of a multivariate setting. In the bivariate model, two stocks are considered and it is assumed that each of the two stocks is modeled via a WISMC model. Let us denote as usual $J_n^i, T_n^i, U_n^{\lambda_i}$, and $Z^i(n)$ the price return, the time, the value of the index and the state of the return process, respectively, all considered at the n-th transition. Also, let i denote the stock considered and in the case of the bivariate we have $i = 1, 2$.

In order to simplify the definition of the bivariate WISMC, the backward recurrence time process for the stock i is defined, for each time $t \in \mathbb{N}$ by $B^i(t) = t - T_{N^i(t)}$, where $N^i(t)$ is the counting process associated to the stock i. This way, the bivariate model can be simplified and be defined in terms of $(Z^i(t), B^i(t), U^{\lambda_i}(t))$.

In defining the model, the three following assumptions have to be made. Let us denote with $\mathbf{Z}(n) = (Z^1(n), Z^2(n))$, $\mathbf{B}(n) = (B^1(n), B^2(n))$, $\mathbf{U}^\lambda(n) = (U^{\lambda_1}(n), U^{\lambda_2}(n))$, $\mathbf{j} = (j_1, j_2)$, $\mathbf{i} = (i_1, i_2)$, $\mathbf{d} = (d_1, d_2)$ and $\mathbf{u} = (u_1, u_2)$.

Assumption 1:

$$\begin{aligned}&\mathbb{P}[\mathbf{Z}(n+1) = \mathbf{j}, \mathbf{B}(n+1) = \mathbf{d} | \sigma(\mathbf{Z}(h), \mathbf{B}(h)), 0 \leq h \leq n, \mathbf{Z}(n) = \mathbf{i}, \mathbf{B}(n) = \mathbf{u}] \\ &= \mathbb{P}[\mathbf{Z}(n+1) = \mathbf{j}, \mathbf{B}(n+1) = \mathbf{d} | \mathbf{Z}(n) = \mathbf{i}, \mathbf{B}(n) = \mathbf{u}, \mathbf{U}^\lambda(n) = \mathbf{v}]\end{aligned} \tag{3.29}$$

Assumption 1 basically states that knowing $(\mathbf{Z}(n) = \mathbf{i}, \mathbf{B}(n) = \mathbf{u}, \mathbf{U}^\lambda(n) = \mathbf{v})$ is sufficient to give the conditional distribution of $(\mathbf{Z}(n+1), \mathbf{B}(n+1))$ independently of the past values of the variables.

It results that:

$$\begin{aligned}&\mathbb{P}[\mathbf{Z}(n+1) = \mathbf{j}, \mathbf{B}(n+1) = \mathbf{d} | (\mathbf{Z}(n) = i, \mathbf{B}(n) = u), \mathbf{U}^\lambda(n) = v] \\ &= \mathbb{P}[Z^1(n+1) = j_1, B^1(n+1) = d_1 | (Z^2(n+1) = j_2, B^2(n+1) = d_2, \\ &\qquad \mathbf{Z}(n) = \mathbf{i}, \mathbf{B}(n) = \mathbf{u}, \mathbf{U}^\lambda(n) = \mathbf{v})] \\ &* \mathbb{P}[Z^2(n+1) = j_2, B^2(n+1) = d_2 | \mathbf{Z}(n) = \mathbf{i}, \mathbf{B}(n) = \mathbf{u}, \mathbf{U}^\lambda(n) = \mathbf{v}]\end{aligned} \tag{3.30}$$

In order to compute 3.30, the following assumption is needed:

Assumption 2:

$$\begin{aligned}&\mathbb{P}[Z^2(n+1) = j_2, B^2(n+1) = d_2 | \mathbf{Z}(n) = \mathbf{i}, \mathbf{B}(n) = \mathbf{u}, \mathbf{U}^\lambda(n) = \mathbf{v}] \\ &= \mathbb{P}[Z^2(n+1) = j_2, B^2(n+1) = d_2 | Z^2(n) = i_2, B^2(n) = u_2, U^{\lambda_2}(n) = v_2] \\ &=: p^2_{(i_2,u_2)(j_2,d_2)}(v_2)\end{aligned} \tag{3.31}$$

The second assumption affirms that the next state of return and the next duration of the second stock depend only on these two variables observed at a previous time, i.e. the second stock evolves by its own while the evolution of the first stock depends on the second one. Obviously, such a hierarchy can be inverted.

Assumption 3:

$$\begin{aligned}
&\mathbb{P}[Z^1(n+1) = j_1, B^1(n+1) = d_1 | Z^2(n+1) = j_2, B^2(n+1) = d_2, \\
&\qquad \mathbf{Z}(n) = \mathbf{i}, \mathbf{B}(n) = \mathbf{u}, \mathbf{U}^\lambda(n) = \mathbf{v}] \\
&= \mathbb{P}[Z^1(n+1) = j_1, B^1(n+1) = d_1 | sgn(Z^2(n+1)) = s, \\
&\qquad Z^1(n) = i_1, B^1(n) = u_1, U^{\lambda_1}(n) = v_1] \\
&=: \tilde{p}^1_{(i_1,u_1)(j_1,d_1)}(v_1; s)
\end{aligned} \tag{3.32}$$

where $sgn(Z^2(n+1))$ is the sign of $Z^2(n+1)$ and it can assume the value $+, 0, -$ according to the fact that the stock number 2 exhibits a positive, constant or negative return, respectively. This assumption simplifies drastically the model while still preserving the cross correlation between the two stocks.

From Assumption 1, 2 and 3 the joint one step transition probability of the two stocks can be calculated as the following product:

$$\tilde{p}^1_{(i_1,u_1)(j_1,d_1)}(v_1; s) * p^2_{(i_2,u_2)(j_2,d_2)}(v_2) \tag{3.33}$$

The probabilities $p^2_{(i_2,u_2)(j_2,d_2)}(v_2)$ are represented as a function of the Weighted Indexed semi-Markov kernel. The probabilities $\tilde{p}^1_{(i_1,u_1)(j_1,d_1)}(v_1; s)$ can be evaluated from the data by using the following estimator:

$$\frac{N_L^{1,2}(i_1, u_1, v_1; j_1, d_1, s_2)}{N_L^{1,2}(i_1, u_1, v_1; s_2)} \tag{3.34}$$

where L is the length of the bivariate series of stock returns;

$N_L^{1,2}(i_1, u_1, v_1; j_1, d_1, s_2) = \sum_{t=1}^{L} 1_{\{} Z^1(t) = j_1, B^1(t) = d_1, sgn(Z^2(t)) =$

$s_2, Z^1(t-1) = i_1, B^1(t-1) = u_1, U^{\lambda_1}(t-1) = v_1\}$;

$N_L^{1,2}(i_1, u_1, v_1; s_2) = \sum_{j_1 \in E} \sum_{d_1 \in \mathbb{N}} N^{1,2}(i_1, u_1, v_1; j_1, d_1, s_2)$.

Chapter 4

Markov processes for the salary evolution

4.1 Introduction

We previously introduced a particular type of stochastic processes which have in common the "Markov property". As mentioned in the previous Chapters, semi-Markov chains (SMC) are a wide class of this type of stochastic process that generalize at the same time both Markov chains and renewal processes. Their main advantage is that of using whatever type of waiting time distribution to model the time of having a transition from one state to another.

This major flexibility has on the other hand a weakness, i.e. the amount of data necessary to estimate the parameters of the model should be very large. Since pension funds have many participants and data are kept for a long time, it seems then natural to try to verify the semi-Markovian hypothesis on future streams of salaries.

In this Chapter, we propose a semi-Markov chain to model the salary levels of participants in a pension scheme. The aim of the models is to

understand the evolution in time of the salary of active workers in order to implement it in the construction of the actuarial technical balance sheet. It is worth mentioning that the level of the contributions in a pension scheme is directly proportional to the incomes of the active workers (in almost all cases it is a percentage of the worker's incomes). As a consequence, an adequate modeling of the salary evolution is essential for the determination of the contributions paid to the fund and thus for the determination of the fund's sustainability.

The models are applied to a large dataset of a real compulsory Italian pension fund of the first pillar. The semi-Markovian hypothesis will be tested by using a statistic test developed in Stenberg et al. (2006b).

4.2 A Markov approach for the salary evolution

The aim of this work is to test the possibility of modeling the yearly salary evolution for an active insured person through a Markov process. In particular, we are interested in finding out if the salary evolution can be modeled through a semi-Markov model with discrete backward recurrence times. As mentioned above, a similar model has been used in modeling disability insurance.

Let us then briefly describe this type of Markov process and its main characteristics. We will use the same definitions seen in the previous Chapters with regard to a semi-Markov model defined in a probability state space $(\Omega, \mathbf{F}, \mathbb{P})$. We suppose that the SMC represents the salary level of an "active" member of the pension scheme. Let us then consider the following two

sequences of random variables:

$$J_n : \Omega \to E; T_n : \Omega \to \mathbb{N}$$

Obviously, J_n and T_n denote, the state and the time of the n-th transition of the system, respectively.

We assume that (J_n, T_n) is a Markov Renewal Process on the state space $E \times \mathbb{N}$ with kernel $Q_{ij}(t),\ i, j \in E, t \in \mathbb{N}$.

The kernel has the following probabilistic interpretation:

$$\begin{aligned} &\mathbb{P}[J_{n+1} = j, T_{n+1} - T_n \leq t | \sigma(J_h, T_h),\ h \leq t, J_n = i] = \\ &\mathbb{P}[J_{n+1} = j, T_{n+1} - T_n \leq t | J_n = i] = Q_{ij}(t) \end{aligned} \tag{4.1}$$

From the kernel defined above, we can obtain the transition probability matrix of the embedded Markov chain J_n, $\mathbf{P} = (p_{ij})$ as:

$$p_{ij} = \lim_{t \to \infty} Q_{ij}(t);\ i, j \in E,\ t \in \mathbb{N}$$

Obviously, it results that:

$$\mathbb{P}[J_{n+1} = j \mid J_n = i] = p_{ij}$$

Furthermore, we can introduce the probability to have the next transition in state j at time t given the starting at time zero from state i

$$\begin{aligned} b_{ij}(t) &= \mathbb{P}[J_{n+1} = j, T_{n+1} - T_n = t | J_n = i] = \\ &= \begin{cases} Q_{ij}(t) - Q_{ij}(t-1) & \text{if } t > 0 \\ 0 & \text{if } t = 0 \end{cases} \end{aligned} \tag{4.2}$$

The distribution functions, that represents the unconditional waiting time distribution in a generic state i, is defined as follows:

$$H_i(t) = \mathbb{P}[T_{n+1} - T_n \leq t | J_n = i] = \sum_{j \in E} Q_{ij}(t) \tag{4.3}$$

The Radon-Nikodym theorem assures for the existence of a function $G_{ij}(t)$ such that:

$$G_{ij}(t) = \mathbb{P}[T_{n+1} - T_n \leq t | J_n = i, J_{n+1} = j] = \begin{cases} \frac{Q_{ij}(t)}{p_{ij}} & \text{if } p_{ij} \neq 0 \\ 1 & \text{if } p_{ij} = 0 \end{cases} \tag{4.4}$$

The function defined above denotes the waiting time distribution function (also denoted as the sojourn time distribution function) in state i given that, in the next transition, the process will be in the state j.

It should be noted that the sojourn time distribution $G_{ij}(\cdot)$ can be any type of distribution function. In cases where $G_{ij}(\cdot)$ is geometrically or an exponentially distributed, we obtain the discrete or the continuous time Markov chain, respectively.

It is possible to define the homogeneous semi-Markov chain (HSMC) $Z(t)$ as

$$Z(t) = J_{N(t)}, \quad \forall t \in \mathbb{N} \tag{4.5}$$

where $N(t) = \sup\{n \in \mathbb{N} : T_n \leq t\}$. The process $Z(t)$ represents the state of the system for each waiting time t.

Let us denote the transition probabilities of the HSMP by

$$\phi_{ij}(t) = \mathbb{P}[Z(t) = j | Z(0) = i, T_0 = 0]$$

They satisfy the following evolution equation[1]

$$\phi_{ij}(t) = \delta_{ij}(1 - H_i(t)) + \sum_{k \in E} \sum_{\tau=1}^{t} b_{ik}(\tau)\phi_{kj}(t - \tau). \tag{4.6}$$

Once the semi-Markov process is defined, it is necessary to introduce the discrete backward recurrence time process linked to the SMP.

[1]To solve Eq. (4.6) there are well known algorithms in the SMP literature (see Barbu et al. (2004)).

For each time $t \in \mathbb{N}$, let us define the following stochastic process:

$$B(t) = t - T_{N(t)}. \tag{4.7}$$

This process is called the discrete backward recurrence time process.

If the semi-Markov process $Z(t)$ indicates the state of the system at time t, $B(t)$ indicates the time since the last jump.

The joint stochastic process $(Z(t), B(t), t \in \mathbb{N})$ with values in $E \times \mathbb{N}$ is a Markov process. That is:

$$\begin{aligned} &P[Z(T)=j, B(T)\leq v'|\sigma(Z(h), B(h)), h\leq t, Z(t)=i, B(t)=v] \\ &= P[Z(T) = j, B(T) \leq v'|Z(t) = i, B(t) = v]. \end{aligned}$$

To save space the event $\{Z(0) = i, B(0) = v\}$ can be denoted in a more compact form by (i, v).

In the sequel we will make use of the following probabilities:

$$\begin{aligned} {}^b\phi^b_{ij}(v; v', t) &= P[Z(t) = j, B(t) = v'|(i, v)]; \\ {}^b\phi_{ij}(v; t) &= P[Z(t) = j|(i, v)]. \end{aligned} \tag{4.8}$$

Our next step is to compute ${}^b\phi^b_{ij}(v; v', t)$ as a function of the semi-Markov kernel. For all states $i, j \in E$ and times $h, v, t \in \mathbb{N}$ such that $H_i(v) < 1$ we have that

$$\begin{aligned} {}^b\phi^b_{ij}(v; v', t) =& \frac{\delta_{ij}[1 - H_i(t + v)]}{[1 - H_i(v)]} 1_{\{v'=t+v\}} + \\ & \sum_{k\in E} \sum_{s=1}^{t} \frac{b_{ik}(s + v)}{[1 - H_i(v)]} \, {}^b\phi^b_{kj}(0; v', t - s), \end{aligned} \tag{4.9}$$

Notice that ${}^b\phi^b_{ij}(0; v', t)$ satisfies the following system of equations

$${}^b\phi^b_{ij}(0; v', t) = \delta_{ij}[1 - H_i(t + v)] + \sum_{k\in E} \sum_{s=1}^{t} b_{ik}(s + v) \, {}^b\phi^b_{kj}(0; v', t - s).$$

It results that ${}^b\phi_{ij}(v;t) = \sum_{v'=0}^{t+v} {}^b\phi_{ij}^b(v;v',t)$. Consequently:

$$\begin{aligned} {}^b\phi_{ij}(v;t) &= \delta_{ij}\frac{[1-H_i(t+v)]}{[1-H_i(v)]} \\ &+ \sum_{k\in E}\sum_{s=1}^{t}\frac{b_{ik}(s+v)}{[1-H_i(v)]}\phi_{kj}(t-s) \end{aligned} \tag{4.10}$$

4.3 Empirical application

4.3.1 Database Description

The model described in the previous section has been applied to the stream of salaries of the active participants of an Italian National Institute of Social security. This is one of the biggest Private Schemes concerning first pillar pensions for Professionals in Italy. The database contains information regarding more then 27 thousand active participants recorded for a time span of 30 years, starting from 1981 till 2011.

The following tables summarize some of the most relevant descriptive statistics of the dataset. In Table 4.1, the age distribution of the active participants of is shown. The age is calculated at the end of the observation period, i.e. at 31st December 2011. Also, the histograms of the age distribution by gender as well as for the total population of the active workers are shown in Figure 4.1.

Since the participants considered may have entered the Pensions Scheme at different years, it seems relevant to also calculate the seniority of each active member of the dataset at the end of the observation period. The seniority is thus calculated considering the date of first subscription of the member to the scheme and results are shown in Table 4.2.

Since the database consists of streams of salaries that refer to a long time span, it is important to also consider the monetary reevaluation of wages. In order to perform this, the index FOI calculated by ISTAT and obtained from *www.istat.it* has been used. In particular, wages of active workers have been reevaluated up to the year 2011. Table 4.3 shows the coefficients used for the reevaluation for each year.

The salaries of the participants are collected by the Institute because the

Classes of Age	**Male**	**Female**	**Total**
$[0, 25]$	72	59	131
$(25, 30]$	777	726	1503
$(30, 35]$	1772	1742	3514
$(35, 40]$	2316	2258	4574
$(40, 45]$	2520	2190	4710
$(45, 50]$	2865	2009	4874
$(50, 55]$	2574	1327	3901
$(55, 60]$	1914	750	2664
$(60, 65]$	912	198	1110
$(65, 70]$	126	34	160
$(70, 75]$	58	10	68
$(75, 80]$	30	6	36
$(80, \dots)$	29	2	31
Total	**15.965**	**11.311**	**27.276**

Table 4.1: Age class distribution of active participants

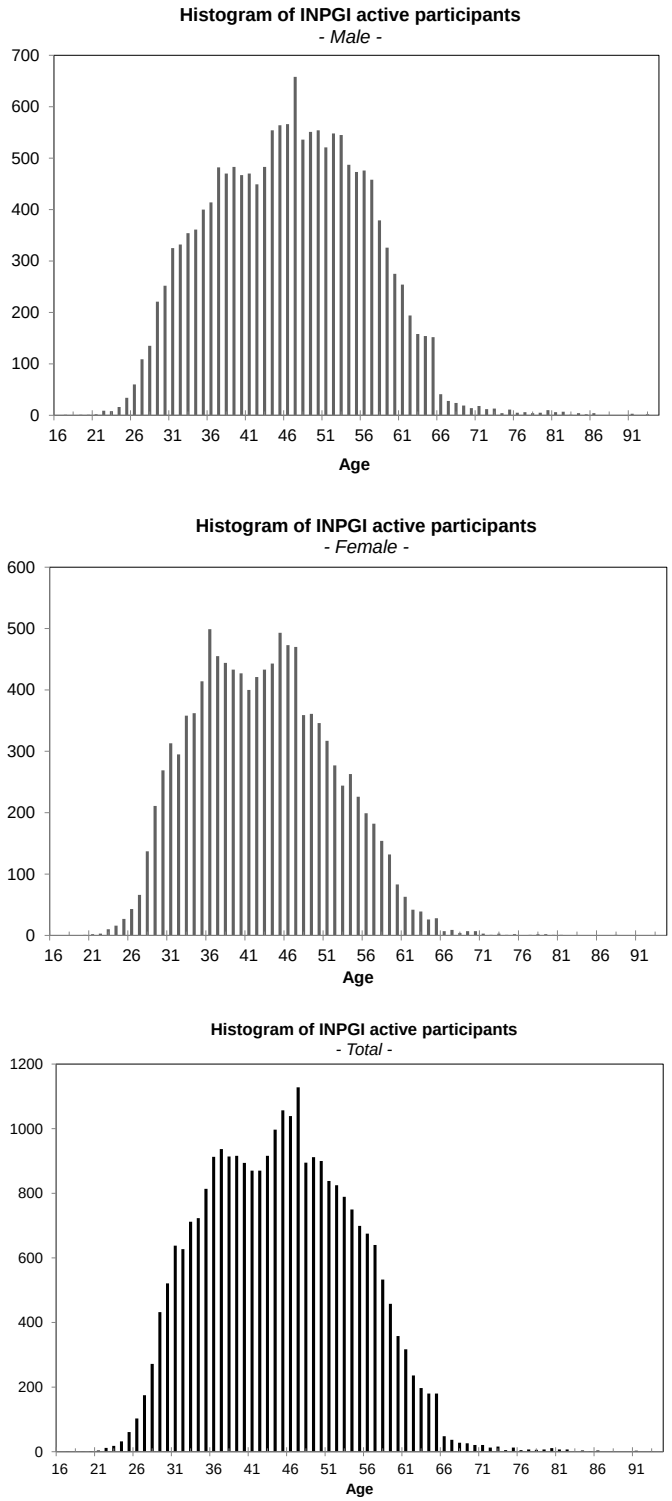

Figure 4.1: Histogram of active workers by age

Classes of seniority	**Male**	**Female**	**Total**
$[0, 5]$	2600	2457	5057
$(5, 10]$	3419	3031	6450
$(10, 15]$	3199	2375	5574
$(15, 20]$	1913	1292	3205
$(20, 25]$	2997	1642	4639
$(25, 30]$	974	342	1316
$(30, 35]$	643	151	794
$(35, 40]$	195	18	213
$(40, \dots)$	25	3	28
Total	**15.965**	**11.311**	**27.276**

Table 4.2: Seniority class distribution of active participants

amount of the contribution paid by each single participant is proportionally dependent on his\her incomes. Given that the database reports only annual salaries collected by the Scheme for contribution purposes, we calculated the monthly salaries by taking into account the effective months of contributions paid by each participant. Graph 4.2 shows the mean salary value of all participants for the time span considered while graph 4.3 shows the histogram of the average monthly salaries. The first two bins of the histogram are equal to 500 while from the third and on they are equal to 2000. This choice helps better analyze lower classes of salaries, i.e. salaries less than 2000, which is the most populated.

Year	***FOI***	**Year**	***FOI***	**Year**	***FOI***
1981	*3.821*	1991	*1.696*	2001	*1.225*
1982	*3.284*	1992	*1.609*	2002	*1.196*
1983	*2.856*	1993	*1.544*	2003	*1.167*
1984	*2.583*	1994	*1.486*	2004	*1.145*
1985	*2.378*	1995	*1.410*	2005	*1.125*
1986	*2.241*	1996	*1.357*	2006	*1.103*
1987	*2.142*	1997	*1.334*	2007	*1.085*
1988	*2.041*	1998	*1.310*	2008	*1.051*
1989	*1.915*	1999	*1.290*	2009	*1.043*
1990	*1.805*	2000	*1.258*	2010	*1.027*

Source: ISTAT

Table 4.3: FOI Index

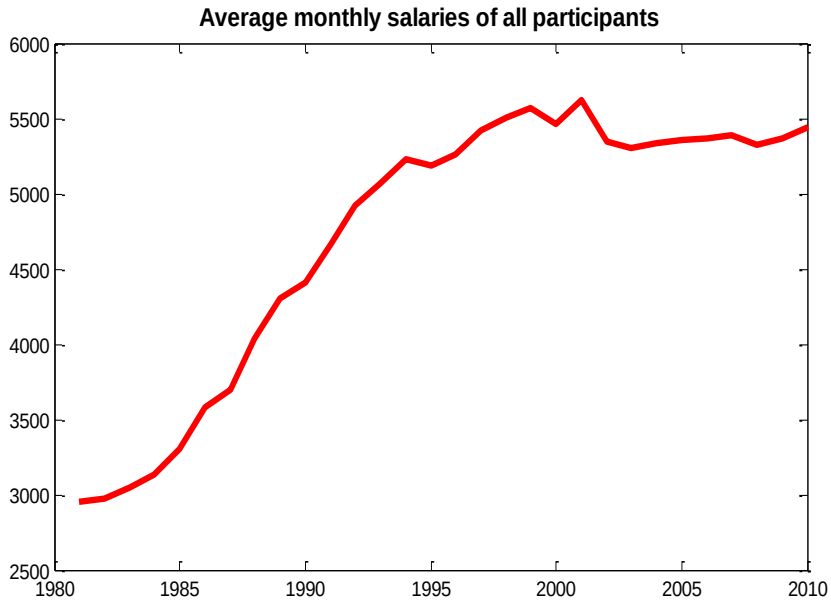

Figure 4.2: Average monthly salaries of active participants

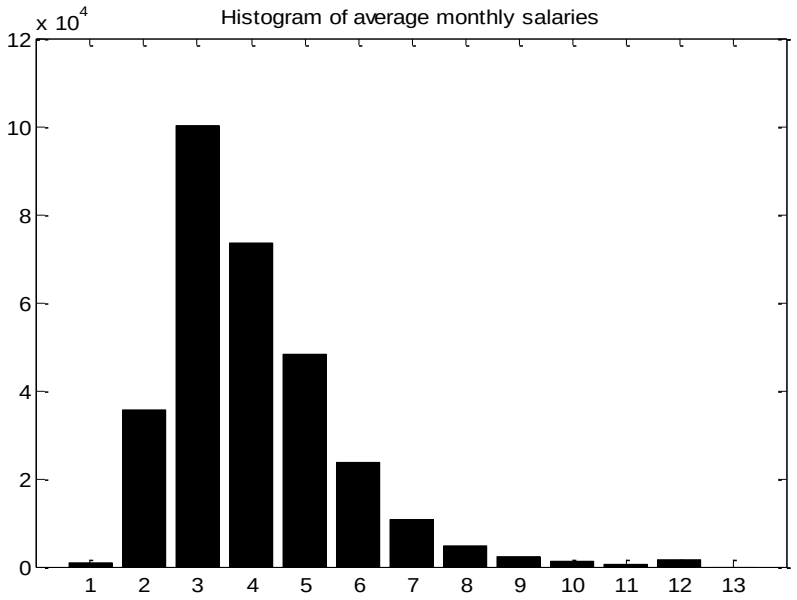

Figure 4.3: Histogram of monthly salaries of active participants

4.3.2 Modeling the salary evolution

In the following paragraph we will try to implement a Markov processes to describe the salary evolution of the active participants of the Pension Scheme. In particular, we are interested in finding whether the model described in Section 4.2, i.e. a semi-Markov model with backward reward process, is more suitable than a simple Markov chain to model salaries.

From the analysis of the graphs shown in the above paragraph, and in particular of the histogram of the average monthly salaries, we chose to divide salaries into 11 states. The first ten states are constructed considering salaries from 0 to 20000 and forming groups by a step of 2000, i.e. the first state considers the interval $[0, 2000)$, the second one $[2000, 4000)$ and so on. The last state is a residual class for salaries higher than 20000.

The salary evolution can be modeled considering transitions from the above defined classes through a Markov process, which as mentioned before

considers geometrically distributed waiting times, or through a semi-Markov process which permits the use of whatever type of distribution of the waiting times.

First, we estimated the transition probability matrices $\hat{p}_{ij}$ of a Markov chain through its maximum likelihood estimator given by the equation:

$$\hat{p}_{ij} = \frac{\sum_{n=1}^{\omega} \mathbf{1}_{\{J_{n-1}=i, J_n=j\}}}{\sum_{n=1}^{\omega} \mathbf{1}_{\{J_{n-1}=i\}}} = \frac{N_{ij}}{N_i} \tag{4.11}$$

where $\omega \in \mathbb{N}$ is the observation period and N_{ij}, N_i denote, respectively, the number of times the process transitioned from a generic state i to another generic state j and the total number of times the process transitioned in the generic state i. The estimated matrix is:

$$\hat{p}_{ij} = \begin{pmatrix}
0.676 & 0.310 & 0.012 & 0.002 & 0.001 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 \\
0.038 & 0.797 & 0.157 & 0.006 & 0.001 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 \\
0.003 & 0.069 & 0.789 & 0.133 & 0.005 & 0.001 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 \\
0.001 & 0.008 & 0.060 & 0.806 & 0.115 & 0.007 & 0.002 & 0.000 & 0.000 & 0.000 & 0.000 \\
0.000 & 0.003 & 0.008 & 0.078 & 0.774 & 0.119 & 0.011 & 0.003 & 0.001 & 0.001 & 0.001 \\
0.000 & 0.004 & 0.006 & 0.012 & 0.099 & 0.724 & 0.125 & 0.020 & 0.006 & 0.002 & 0.003 \\
0.000 & 0.002 & 0.004 & 0.006 & 0.018 & 0.132 & 0.649 & 0.136 & 0.031 & 0.008 & 0.013 \\
0.001 & 0.006 & 0.007 & 0.005 & 0.011 & 0.037 & 0.171 & 0.534 & 0.156 & 0.039 & 0.032 \\
0.000 & 0.005 & 0.007 & 0.004 & 0.015 & 0.020 & 0.061 & 0.167 & 0.450 & 0.160 & 0.113 \\
0.000 & 0.007 & 0.007 & 0.005 & 0.004 & 0.023 & 0.027 & 0.051 & 0.188 & 0.395 & 0.293 \\
0.001 & 0.007 & 0.006 & 0.004 & 0.011 & 0.007 & 0.018 & 0.026 & 0.036 & 0.066 & 0.817
\end{pmatrix}$$

The second step is that of modeling the salary evolution through a semi-Markovian process. In order to do so, the distribution of the waiting times has to be estimated.

Let us firstly define the probability mass function of the sojourn times as follows:

$$g_{ij}(t) = \mathbb{P}[T_{n+1} - T_n = t | J_n = i, J_{n+1} = j] = \begin{cases} G_{ij}(t) - G_{ij}(t-1) & \text{if } t > 1 \\ G_{ij}(1) & \text{if } t = 1 \end{cases} \tag{4.12}$$

We considered a discrete time step of one year and calculated the waiting times for up to 20 years. Through an algorithm we estimated the sojourn times on the data using the following:

$$\hat{g}_{ij}(t) = \frac{\sum_{n=1}^{\omega} \mathbf{1}_{\{T_n - T_{n-1} = t, J_{n-1} = i, J_n = j\}}}{\sum_{n=1}^{\omega} \mathbf{1}_{\{J_{n-1} = i, J_n = j\}}} \tag{4.13}$$

In order to verify the semi-Markov hypothesis, a test proposed by Stenberg et al. (2006a) can be applied. Obviously, the model can be considered semi-Markovian if the sojourn times are not geometrically distributed.

Under the geometrical hypothesis the equality $g_{ij}(1)(1-g_{ij}(1)) - g_{ij}(2) = 0$ must hold. Then a sufficiently strong deviation from this equality has to be interpreted as an evidence in favor of the semi-Markov model.

The test-statistic is the following:

$$\hat{S}_{ij} = \frac{\sqrt{N(i,j)}\big(\hat{g}_{ij}(1)(1-\hat{g}_{ij}(1)) - \hat{g}_{ij}(2)\big)}{\sqrt{\hat{g}_{ij}(1)(1-\hat{g}_{ij}(1))^2(2-\hat{g}_{ij}(1))}}. \tag{4.14}$$

where $N(i,j)$ denotes the number of transitions from state i to state j observed in the sample and $\hat{g}_{ij}(x)$ is the empirical estimator of the probability $g_{ij}(x)$.

The results of the test statistic are shown in Table 4.4. It reports the score as calculated from Eq. 4.14 for the eleven states of the salaries.

This statistic, under the geometrical hypothesis H_0 (or markovian hypothesis), has approximately the standard normal distribution. This means that large values of the test statistic suggest the rejection of the Markovian hypothesis in favor of the more general semi-Markov one. We applied this procedure to our data to execute tests at a significance level of 95%.

In Table 4.5 we show the results of the test applied to the waiting time distribution functions where R means the null hypothesis is *Rejected* while A means that the null hypothesis of geometrical distribution of the waiting times is *Accepted*.

i \ j	1	2	3	4	5	6	7	8	9	10	11
1	n.a.	-3.59	3.88	0.58	0.57	0.43	n.a.	n.a.	n.a.	n.a.	n.a.
2	3.86	n.a.	-3.49	-0.15	2.09	1.41	-0.75	-0.19	1.83	n.a.	0,04
3	1.93	9.89	n.a.	2.73	1.84	1.50	1.28	-1.34	-0.15	0.82	n.a.
4	2.28	7.13	8.82	n.a.	3.94	-0.08	1.39	-0.56	0.69	0.47	-0.19
5	0.85	2.48	2.47	8.25	n.a.	3.94	2.10	0.53	-0.01	1.36	1.46
6	n.a.	1.06	1.47	1.70	5.78	n.a.	3.54	1.26	-0.81	1.25	0.14
7	n.a.	0.97	0.54	0.43	4.63	4.00	n.a.	1.89	1.19	0.79	3.08
8	n.a.	-0.40	-0.56	-0.65	-0.13	2.46	3.44	n.a.	3.81	-1.08	0.72
9	n.a.	-0.46	n.a.	-0.52	1.51	1.86	1.56	1.96	n.a.	-0.45	0.88
10	n.a.	n.a.	n.a.	n.a.	n.a.	-0.28	-0.79	3.05	1.42	n.a.	0.24
11	n.a.	2.86	2.39	1.73	1.26	1.21	1.84	0.11	0.36	1.97	n.a.

Table 4.4: Test scores for classes of salaries

i \ j	1	2	3	4	5	6	7	8	9	10	11
1	n.a.	R	R	A	A	A	n.a.	n.a.	n.a.	n.a.	n.a.
2	R	n.a.	R	A	R	A	A	A	A	n.a.	A
3	A	R	n.a.	R	A	A	A	A	A	A	n.a.
4	R	R	R	n.a.	R	A	A	A	A	A	A
5	A	R	R	R	n.a.	R	R	A	A	A	A
6	n.a.	A	A	A	R	n.a.	R	A	A	A	A
7	n.a.	A	A	A	R	R	n.a.	A	A	A	R
8	n.a.	A	A	A	A	R	R	n.a.	R	A	A
9	n.a.	A	n.a.	A	A	A	A	R	n.a.	A	A
10	n.a.	n.a.	n.a.	n.a.	n.a.	A	A	R	A	n.a.	A
11	n.a.	R	R	A	A	A	A	A	A	R	n.a.

Table 4.5: Decision for the null hypothesis for classes of salaries

As shown in Table 4.5, the geometric hypothesis is rejected for 31% of the waiting time distributions. There is, thus, evidence that we should be inclined in choosing the more general model of a semi-Markov process to that of a simple Markov chain.

We also considered the case of the worker's job qualification for the construction of the states of the salary process. The database furnished the necessary information regarding the different levels of qualification for each participant at each year of the observation period. In order to have an estimation as robust as possible, only qualifications with more then 5000 participants have been considered.

The hypothesis under such modeling is that the transition from one qualification to the other, which would indirectly imply a transition in the salary level, can be described by a Markov process. The procedure applied is quite similar to the one described above. First, we will estimate the transition probability matrix of the Markov chain through the same algorithm described above. The estimated matrix $\hat{p}_{ij}$ is:

$$\begin{pmatrix}
0.90 & 0.03 & 0.00 & 0.01 & 0.00 & 0.02 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.03 \\
0.02 & 0.92 & 0.00 & 0.01 & 0.00 & 0.01 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.03 \\
0.00 & 0.11 & 0.85 & 0.01 & 0.00 & 0.01 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.01 \\
0.00 & 0.03 & 0.05 & 0.89 & 0.00 & 0.01 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.01 \\
0.00 & 0.01 & 0.01 & 0.11 & 0.84 & 0.01 & 0.00 & 0.00 & 0.01 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.01 \\
0.00 & 0.01 & 0.00 & 0.03 & 0.02 & 0.91 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.01 \\
0.01 & 0.01 & 0.00 & 0.03 & 0.02 & 0.47 & 0.41 & 0.01 & 0.00 & 0.02 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.01 \\
0.00 & 0.01 & 0.01 & 0.01 & 0.00 & 0.09 & 0.00 & 0.78 & 0.00 & 0.02 & 0.00 & 0.00 & 0.00 & 0.00 & 0.02 & 0.05 \\
0.00 & 0.02 & 0.02 & 0.02 & 0.01 & 0.01 & 0.00 & 0.00 & 0.91 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.01 \\
0.00 & 0.01 & 0.00 & 0.02 & 0.02 & 0.38 & 0.00 & 0.01 & 0.00 & 0.52 & 0.01 & 0.00 & 0.00 & 0.00 & 0.01 & 0.02 \\
0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.01 & 0.00 & 0.00 & 0.00 & 0.00 & 0.95 & 0.00 & 0.00 & 0.00 & 0.00 & 0.03 \\
0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.97 & 0.00 & 0.00 & 0.00 & 0.02 \\
0.00 & 0.01 & 0.00 & 0.01 & 0.00 & 0.07 & 0.32 & 0.02 & 0.00 & 0.03 & 0.00 & 0.00 & 0.50 & 0.00 & 0.01 & 0.02 \\
0.00 & 0.01 & 0.00 & 0.01 & 0.01 & 0.05 & 0.00 & 0.01 & 0.00 & 0.61 & 0.01 & 0.00 & 0.00 & 0.24 & 0.01 & 0.04 \\
0.00 & 0.01 & 0.00 & 0.00 & 0.01 & 0.03 & 0.00 & 0.01 & 0.00 & 0.14 & 0.00 & 0.00 & 0.01 & 0.63 & 0.13 & 0.04 \\
0.01 & 0.01 & 0.00 & 0.01 & 0.00 & 0.04 & 0.00 & 0.01 & 0.00 & 0.02 & 0.05 & 0.00 & 0.00 & 0.01 & 0.01 & 0.81
\end{pmatrix}$$

Secondly, we suppose that a semi-Markov model can better explain the transitions between states since it gives the possibility of using whatever type

of waiting time distribution.

The values of the statistic test are obtained in the same way as seen above and are reported in Table 4.6. We fixed a level of significance of 95% and the results of the statistic test are given in Table 4.7. The null hypothesis of geometrical distribution is rejected for 25% of the waiting time distributions considered and thus we conclude that the semi-Markov process is more suitable to model the salary evolution of participants with states formed on the basis of the the active workers' qualification.

i \ j	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	n.a.	0.59	-0.50	-0.69	2.32	1.40	1.41	0.02	0.06	0.83	-0.43	n.a.	n.a.	n.a.	0.24	1.08
2	-0.10	n.a.	1.53	0.16	-0.08	1.96	0.82	-0.13	1.17	3.90	0.20	-0.30	n.a.	n.a.	1.22	2.72
3	1.67	-0.12	n.a.	0.95	0.68	1.38	n.a.	0.73	0.81	1.35	-0.25	n.a.	n.a.	0.82	0.27	0.90
4	-2.49	-0.95	-3.78	n.a.	0.85	3.27	0.19	-0.46	0.05	1.14	1.80	n.a.	n.a.	n.a.	0.00	1.41
5	1.24	1.79	0.97	-0.75	n.a.	2.87	1.97	1.48	-4.14	2.55	0.63	n.a.	n.a.	0.82	0.99	-0.24
6	2.30	0.93	-0.43	-2.36	-0.29	n.a.	5.11	1.58	-0.93	1.03	1.55	0.39	4.05	2.54	0.44	1.36
7	-1.44	-0.77	-0.57	-2.96	-1.94	-24.71	n.a.	0.09	-0.23	-1.99	-0.61	-2.27	0,00	n.a.	-0.52	-0,33
8	0.01	1.28	-0.02	1.87	0.88	0.83	3.19	n.a.	0.37	-0.73	-0.39	-0.55	n.a.	-0.15	-1.20	-0.56
9	0.14	1.04	-0.72	0.92	5.83	0.36	n.a.	0.56	n.a.	n.a.	-0.82	n.a.	n.a.	n.a.	-0.82	1.13
10	-1.52	-2.36	0.60	-3.81	-5.88	-48.41	2.19	0.31	-1.29	n.a.	-2.71	n.a.	n.a.	-1.62	-1.93	-3.28
11	-3.44	1.08	1.15	0.35	n.a.	-0.14	n.a.	-0.47	n.a.	0.23	n.a.	-0.19	n.a.	n.a.	-0.82	1.28
12	n.a.	n.a.	n.a.	-0.10	n.a.	-2.35	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	0.72
13	-4.36	-7.13	-2.98	-8.26	-5.84	-19.77	-87.59	-8.11	-2.49	-8.39	-0.52	-1.41	n.a.	1.17	0.15	-8.14
14	0.23	0.78	-1.14	-0.33	0.75	-0.02	-0.68	-0.61	n.a.	-1.88	-1.22	-0.99	n.a.	n.a.	-0.07	0.43
15	-0.28	-0.16	n.a.	-1.94	-0.44	-0.60	-0.79	-0.19	-1.15	-9.53	-1.12	n.a.	n.a.	-2.18	n.a.	-1.12
16	-0.05	1.38	-0.78	0.56	1.87	4.18	0.10	2.33	-0.21	2.42	-31.66	0.62	-0.47	5.05	1.56	n.a.

Table 4.6: Test scores for qualification states

i \ j	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	n.a.	A	A	A	R	A	A	A	A	A	A	n.a.	n.a.	n.a.	A	A
2	A	n.a.	A	A	A	R	A	A	A	R	A	A	n.a.	n.a.	A	R
3	A	A	n.a.	A	A	A	n.a.	A	A	A	A	n.a.	n.a.	A	A	A
4	R	A	R	n.a.	A	R	A	A	A	A	A	n.a.	n.a.	n.a.	A	A
5	A	A	A	A	n.a.	R	R	A	R	R	A	n.a.	n.a.	A	A	A
6	R	A	A	R	A	n.a.	R	A	A	A	A	A	R	R	A	A
7	A	A	A	R	A	R	n.a.	A	A	R	A	R	A	n.a.	A	A
8	A	A	A	A	A	A	R	n.a.	A	A	A	A	n.a.	A	A	A
9	A	A	A	A	R	A	n.a.	A	n.a.	n.a.	A	n.a.	n.a.	n.a.	A	A
10	A	R	A	R	R	R	R	A	A	n.a.	R	n.a.	n.a.	A	A	R
11	R	A	A	A	n.a.	A	n.a.	A	n.a.	A	n.a.	A	n.a.	n.a.	A	A
12	n.a.	n.a.	n.a.	A	n.a.	R	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	A
13	R	R	R	R	R	R	R	R	R	R	A	A	n.a.	A	A	R
14	A	A	A	A	A	A	A	A	n.a.	A	A	A	n.a.	n.a.	A	A
15	A	A	n.a.	A	A	A	A	A	A	R	A	n.a.	n.a.	R	n.a.	A
16	A	A	A	A	A	R	A	R	A	R	R	A	A	R	A	n.a.

Table 4.7: Decision for the null hypothesis for qualification states

4.4 Conclusions

In this chapter, we proposed a semi-Markov Chain with backward recurrence time to model salary lines. This is a more general model compared to the simple Markov chain since any type of distribution can be used for the waiting times while the latter allows only the geometrical distribution.

The model was applied to a dataset of active participants to an Italian Pension Scheme of the First Pillar. More than 27 thousand workers have been observed over a period time of 30 years. Salaries of participants have been clustered to form the states of a stochastic process in two ways: the first one considers classes of salaries of 2000 while the second one is based on the qualification of the workers. In both cases, the transition probability matrices of the embedded Markov chain have been estimated. Then, the semi-Markov hypothesis is tested by a statistical test applied in the recent literature.

We fixed a significance level of 95% and performed the statistic test on the distribution of the waiting times. The null hypothesis of geometrical distribution was rejected for nearly 30% and 25% of the tests, in the two cases considered, and thus we concluded that the semi-Markov model should be preferred to a simple Markov chain to model salary evolution of participants.

Chapter 5

A change-point approach in IMC

5.1 Introduction

In Section 3.3 a review of semi-Markov processes in the application of high frequency price return modeling was shown. In particular, a new model called the Indexed semi-Markov process was introduced. In this model, a discretization of the Index process was necessary in order to determine a limited number of states, each representing a particular level of volatility. In the above mentioned papers, the authors chose five states of volatility and defined the border values of each class based on the analysis of the histogram of the Index process.

In this Chapter, we propose a method for the discretization of the state space of an Indexed semi-Markov process by using the change-point approach. The change-point approach is used in econometric studies where the process analyzed undergoes a significant change at some point in time, marking a discontinuity with the past. In Polansky (2007), the problem of the change-

point estimation in case of a Markov processes is discussed. The estimation of the transition probability matrices before and after the change are performed. Through the likelihood theory, the author develops a method for the estimation in case the change point is unknown. Also, the possibility to detect more then one change point and a method for selecting the most parsimonious model is shown.

In this Chapter, we try to use the change-point approach to determine the number of states and the border values of each state of the Index process in a generic Indexed Markov model. The chapter is organized as follows: in Section 5.2 a general description of Markov model where an Index process is introduced is provided, in Section 5.3 the discretization of the state space of the Index process based on the change-point approach is formalized and in Section 5.4 the mathematical framework for predicting the state of the Index after one step is shown. Lastly, in Section 5.5 an empirical applications of the methodology and results are evidenced.

5.2 A generalized model of Indexed Markov Chains

The aim of this work is to analyze the discretization of the index process. Since the method applied, which will be fully described in the following paragraphs, does not depend on the type of Markov process chosen but rather on the dependence of the price return processes with the index process it seemed then useful to consider the most simple type of Markov processes for the description of the price return dynamics. As a consequence, and without any loss of generality, let us consider a simplified indexed Markov model where $T_n = n$ for $\forall n \in \mathbb{N}$, i.e. the price return process is described by

a Markov chain rather than a semi-Markov process and we will denote this process as Indexed Markov Chain (IMC) rather than ISMC. The approach used for the discretization of the index process can then be easily extended to the case of more complex models.

Consider the stochastic process $\{J_n\}_{n\in\mathbb{N}}$ that describes the value of the return process at the n-th transition and the stochastic process $\{V_n^m\}$ as the value of the index process at the n-th transition with memory equal to m such that:

$$V_n^m = \frac{\sum_{k=1}^{m-1} f(J_{n-k})}{m} \tag{5.1}$$

We assume that:

$$\begin{aligned} &\mathbb{P}(J_{n+1} = j | J_n = i, V_n^m = v, J_{n-1}, V_{n-1}^m) \\ &= \mathbb{P}(J_{n+1} = j | J_n = i, V_n^m = v) \\ &= p_{ij}(v) \end{aligned} \tag{5.2}$$

Equation (5.2) states that the value of the price return process at the $n+1$ transition depends on the value of process in the previous n-th transition and the value of the index process in the previous n-th transition. From now on we will refer to this model and we will use the hypothesis in equation 5.2 to find the optimal discretization of the index process.

The goal of this work is to try and answer the following questions:

(a) If we consider only two levels of volatility of the price return process, how do we identify the border values of Index process?

(b) Could there be more than two levels of volatility? How can we identify the optimal number of volatility levels?

(c) In case of more than two levels of volatility, how can we identify the border values of the Index process for each level?

5.3 Change-point approach in IMC

In this paper, the change point approach will be used to determine the optimal number of states of the index process V_n^m as well as the border values for each state. This will be obtained by using the fact that the dynamics of the price return process J_n depend on the value of the index process V_n^m.

We defined E as the state space of the price return dynamics J_n such that for $\forall n, J_n(\omega) \in E$. Let E_f be the state space of the index process such that $V_n^\lambda \in E_f$.

Let $\overline{E}_f$ and $\underline{E}_f$ be the maximum and the minimum, respectively, of the space state E_f. Then:

$$V_n^m \in [\underline{E}_f, \overline{E}_f] \tag{5.3}$$

First, let us suppose that there are mainly two levels of volatility in the market: low volatility and high volatility. What we are interested in is to model the price dynamics process J_n through two different Markov processes described by two different transition probability matrices. Let $\underline{P}(\psi_1)$ be the transition probability matrix of the price return process in case of low volatility and $\overline{P}(\psi_1)$ be the transition probability matrix in case of high volatility.

Since the level of the volatility is considered through the index process then, the state space $[\underline{E}_f, \overline{E}_f]$ of the Index process has to be subdivided into two discrete states. Let $\psi_1 \in [\underline{E}_f, \overline{E}_f]$ be the value of the Index model that determines a change in the dynamics of the price return process J_n such that:

(a) The interval $[\underline{E}_f, \psi_1]$ represents the low volatility case. If the Index process $V_n^\lambda < \psi_1$, than we suppose that the price return process J_n is described by the transition probability matrix $\underline{P}(\psi_1)$

(b) The interval $[\psi_1, \overline{E}_f]$ represent the high volatility case. If the Index

process $V_n^\lambda \geq \psi_1$, than we suppose that the price return process J_n is described by the transition probability matrix $\overline{P}(\psi_1)$

5.3.1 Case of a known change point

If the change point ψ_1 is known, the transition probability matrices can be estimated through the maximum likelihood estimators $\hat{\overline{P}}$ and $\hat{\underline{P}}$ whose (i,j)-th elements are given as follows:

$$\hat{\overline{P}}_{ij}(\psi_1) = \frac{\sum_{n=1}^{T-1} \mathbf{1}_{\{J_{n-1}=i, J_n=j, V_{n-1}\geq\psi_1\}}}{\sum_{n=1}^{T-1} \mathbf{1}_{\{J_{n-1}=i, V_{n-1}\geq\psi_1\}}} = \frac{\overline{N}_{ij}(\psi_1)}{\overline{N}_i(\psi_1)} \tag{5.4}$$

$$\hat{\underline{P}}_{ij}(\psi_1) = \frac{\sum_{n=1}^{T-1} \mathbf{1}_{\{J_{n-1}=i, J_n=j, V_{n-1}<\psi_1\}}}{\sum_{n=1}^{T-1} \mathbf{1}_{\{J_{n-1}=i, V_{n-1}<\psi_1\}}} = \frac{\underline{N}_{ij}(\psi_1)}{\underline{N}_i(\psi_1)} \tag{5.5}$$

where $T \in \mathbb{N}$ is the time of the last observation of the process.

In order to test the hypothesis that the two matrices are statistically equal and, as a consequence, there is no significant change in the dynamics of the price return process due to the different levels of the states of the index process a statistical test can be developed:

$$\begin{cases} H_0 : \overline{P} = \underline{P} \\ H_1 : \overline{P} \neq \underline{P} \end{cases} \tag{5.6}$$

Once the two hypothesis are set, a distance measure has to be defined. As shown in Polansky (2007), a convenient distance measure can be developed through the log of the likelihood ratio. Let $\mathfrak{L}(\overline{P}(\psi_1), \underline{P}(\psi_1), \underline{x})$ be the likelihood function of a sample $\underline{x}$ observed on the period $[0, T]$ with a change point ψ_1. Then,

$$\begin{aligned} &\mathfrak{L}(\overline{P}(\psi_1), \underline{P}(\psi_1), \underline{x}) = \mathbb{P}(J_0 = x_0, J_1 = x_1, \ldots, J_T = x_T) \\ &= \prod_{i,j\in E} (\overline{P}_{ij}(\psi_1))^{\overline{N}_{ij}(\psi_1)} * \prod_{i,j\in E} (\underline{P}_{ij}(\psi_1))^{\underline{N}_{ij}(\psi_1)} \end{aligned} \tag{5.7}$$

Let L be the log-likelihood function:

$$
\begin{aligned}
&L(\overline{P}(\psi_1), \underline{P}(\psi_1), \underline{x}) = log(L) \\
&= \sum_{ij} \left(\overline{N}_{i,j}(\psi_1) * log\left(\frac{\overline{N}_{i,j}(\psi_1)}{\overline{N}_i(\psi_1)} \right) \right) + \sum_{ij} \left(\underline{N}_{i,j}(\psi_1) * log\left(\frac{\underline{N}_{i,j}(\psi_1)}{\underline{N}_i(\psi_1)} \right) \right) \\
&=: \overline{L} + \underline{L}
\end{aligned}
\tag{5.8}
$$

A distance measure can then be calculated from the log-likelihood functions:

$$\mathfrak{D} = 2 * [L(\overline{P}(\psi_1)) + L(\underline{P}(\psi_1)) - L(P)] \tag{5.9}$$

where

- $L(\overline{P}(\psi_1)) + L(\underline{P}(\psi_1))$ is the log of the likelihood value for the model under the alternative hypotheses
- $L(P)$ is the log of the likelihood value for model under the null hypotheses. In this case, we suppose that there is no change point in the observed process and thus we estimate the dynamics of the price return process $\{J_n\}$ by using a single transition probability matrix P of the Markov Chain.

5.3.2 Case of an unknown change point

In our case, the change point is not known and it also needs to be estimated. In cases in which the change point is not known, the parameter ψ_1 becomes un unknown parameter of the log-likelihood function seen above. Usually, the maximum likelihood estimator for ψ_1, in this cases, does not exist in closed form but it can be estimated through an iterative method;

$$\hat{\psi}_1 = arg\ max\{\hat{\psi}_1 \in [\underline{E}_f, \overline{E}_f] : \overline{L}_M + \underline{L}_M\} \tag{5.10}$$

where $\overline{L}_M$ and $\underline{L}_M$ are the maximum values of the two likelihood functions observed conditioned on $\hat{\psi}_1$.

Operatively, the methodology is the following:

1. Define on the state space n discrete and distinct points $\psi \in [\underline{E}_f, \overline{E}_f]$.
2. For each possible $\psi \in [\underline{E}_f, \overline{E}_f]$
   (a) Set $\psi_1 = \psi$
   (b) Estimate $\overline{P}(\psi_1)$ with Eq. (5.4) and $\underline{P}(\psi_1)$ with Eq. (5.5)
   (c) Compute $L(\psi_1) = \overline{L}(\psi_1) + \underline{L}(\psi_1)$ as in Eq. (5.8)
3. Fix $\hat{\psi}_1 = arg\ max\{\psi_1 \in [\underline{E}_f, \overline{E}_f] : L(\psi_1)\}$

The results will be the value of the index process that determines a change in the price return process, the estimated transition probability matrices of the price return process that reflect the two different dynamics and the value of the log-likelihood function. Once these outputs are obtained, a test of the hypothesis as shown in (5.6) can be performed.

Unfortunately, the theoretical distribution of the statistic test under H_0 is also not known but it can be approximated by using the bootstrap methodology described as follows. Given a sample $\underline{x}$, a single transition probability matrix P is estimated from the data (i.e. without considering the change point). B trajectories of the same length of the data are simulated from the transition matrix P. For each simulation, the test statistic defined in Eq. (5.9) is calculated. We denote this values as D_B. The theoretical distribution of the statistic test D can then be approximated by the kernel distribution of the simulated statistic test D_B. Once the level of confidence α for the test is fixed, the critical value d_α can be approximated by the $1 - \alpha$ percentile

of the simulated statistic test D_B. Hence, if the statistic test on the sample data $\hat{D}(\underline{x}) \geq d_\alpha$, the null hypothesis cannot be accepted.

The empirical p-value of the test can be then calculated as:

$$p - value = \frac{1}{B+1} * \left[1 + \sum_{i=1}^{B} \mathbf{1}_{D_B(\underline{x}) \geq \hat{D}(\underline{x})}\right] \tag{5.11}$$

5.3.3 Case of more than one unknown change points

The case of two or more change points can be extended from the estimation methods of a single change point shown so far.

Let $\psi_1 < \psi_2 < \cdots < \psi_k$ be k change points defined in the interval $[\underline{E}_f, \overline{E}_f]$. Each possible combination (s) of the k change points divides the interval in $k+1$ sub-intervals:

$$\begin{cases} [\underline{E}_f, \psi_1] & =: I_1^{(s)} \\ [\psi_1, \psi_2] & =: I_2^{(s)} \\ \vdots & \\ [\psi_{r-1}, \psi_r] & =: I_r^{(s)} \\ \vdots & \\ [\psi_k, \overline{E}_f] & =: I_{k+1}^{(s)} \end{cases}$$

For each of the (s) possible partitions of interval in $k+1$ sub-intervals $\left\{I_r^{(s)}\right\}_{r=1}^{k+1}$, $k+1$ transition probability matrices $\hat{P}^{(s)}$ have to be estimated whose (i,j) elements are:

$$\hat{P}_{i,j}^{(s)} = \frac{\sum_{n=1}^{T-1} \mathbf{1}_{J_{n-1}=i, J_n=j, V_{n-1} \in I_r^{(s)}}}{\sum_{n=1}^{T-1} \mathbf{1}_{J_{n-1}=i, V_{n-1} \in I_r^{(s)}}} =: \frac{N_{i,j;r}^{(s)}}{N_{i;r}^{(s)}} \tag{5.12}$$

The likelihood function of the sample $\underline{x}$ observed on the period $[0, T]$ for

the generic partition (s) with $k+1$ change points is:

$$\mathfrak{L}^{(s)}(P_1^{(s)}, P_2^{(s)}, \ldots, P_{k+1}^{(s)}; \underline{x}) = \prod_{r=1}^{k+1} \prod_{i,j \in E} \left(P_{i,j;r}^{(s)}\right)^{N_{i,j;r}^{(s)}} \tag{5.13}$$

The log-likelihood function can then be calculated as

$$L^{(s)} = log(\mathfrak{L}^{(s)}) = L_1^{(s)} + L_2^{(s)} + \cdots + L_{k+1}^{(s)} = \sum_{r=1}^{k+1} L_r^{(s)} \tag{5.14}$$

where $L_r^{(s)}$ is the log-likelihood of the generic (s) partition in reference to the sub-interval I_r. As seen above, the values of the k change points can be calculated in an iterative way. We set n discrete and distinct values of $\psi \in [\underline{E}_f, \overline{E}_f]$. Notice that, for k change points, we have $\binom{n-1}{k}$ possible combinations of the vector of change points $(\psi_1, \psi_2, \ldots, \psi_k)$. The methodology is basically the same as that for one change point but in this case we have to consider all the possible combinations of the set of k change points $(\psi_1, \psi_2, \ldots, \psi_k)$. As seen above, the estimated values can be calculated through:

$$(\hat{\psi}_1, \hat{\psi}_2, \ldots, \hat{\psi}_k) = arg\ max\{(\psi_1 < \psi_2 < \cdots < \psi_k) \in [\underline{E}_f, \overline{E}_f] : L^{(s)}\} \tag{5.15}$$

A test can be performed in this case where:

$$\begin{cases} H_0 : P_0 = P_1 = \cdots = P_k \\ H_1 : P_i \neq P_j \qquad\qquad for\ some\ i \neq j \end{cases} \tag{5.16}$$

The test statistic can be calculated as follows:

$$\mathfrak{D}^{(s)} = 2 * \left[\sum_{r=1}^{k+1} L_r^{(s)} - L(P) \right] \tag{5.17}$$

where $L(P)$ is the log-likelihood value in case there is no change point on the sample $\underline{x}$ and thus a single transition matrix P is estimated. The bootstrap methodology can be used to obtain an approximation of the theoretical distribution of the statistic test.

If the number of change points is unknown, it also needs to be estimated and thus measures such as AIC or BIC can be used.

The AIC objective function is given by:

$$AIC(k) = 2 * c * (c-1) * (k+1) - 2 * \sum_{r=1}^{k+1} L_r^M \tag{5.18}$$

where

- L^M is the value of the log-likelihood of the maximum likelihood estimate of $\psi_1, \psi_2, \ldots, \psi_k$ conditioned on the fact that there are k change points
- c is the number of states of the Markov Chain
- k is the number of change points.

The estimated number of change points is given by:

$$\hat{k}_{AIC} = arg\ min\{k \in \{1, 2, \ldots, n\} : AIC(k)\} \tag{5.19}$$

The BIC objective function is given by:

$$BIC(k) = 2 * log(n) * c * (c-1) * (k+1) - 2 * \sum_{r=1}^{k+1} L_r^M \tag{5.20}$$

where

- L^M is the value of the log-likelihood of the maximum likelihood estimate of $\psi_1, \psi_2, \ldots, \psi_k$ conditioned on the fact that there are k change points
- c is the number of states of the Markov Chain
- k is the number of change points.

- n is the length of the sample $\underline{x}$

The estimated number of change points is given by:

$$\hat{k}_{BIC} = arg\ min\{k \in \{1, 2, \ldots, n\} : BIC(k)\} \tag{5.21}$$

5.4 Predicting the value of the index in the next step

Suppose we have identified k change points of our Indexed Markov chain $\{J_n\}_{n\in\mathbb{N}}$ such that:

$$\psi_1 < \psi_2 < \cdots < \psi_k \ ; \ \psi_i \in \mathbb{R} \ \forall i = 1, 2, \ldots, k$$

As mentioned, the state space of the chain is $E = \{1, 2, \ldots, S\}$ and we have estimated $k + 1$ transition probability matrices. Let us indicate such matrices in general with:

$$P(v) = (p_{ij}(v))_{i,j\in E, v\in\mathbb{R}}$$

and

$$p_{ij}(u) = p_{ij}(v)$$

if

$$\exists a \in \{1, 2, \ldots, k-1\}$$

such that

$$\psi_a < u \leq \psi_{a+1} \ ; \ \psi_a < v \leq \psi_{a+1}$$

This means that if u, v belong to the same interval

$$I_{a+1} := (\psi_a, \psi a + 1]$$

where

$$I_1 = (-\infty, \psi_1]$$

$$\psi_{k-1} = +\infty$$

then

$$P(u) = P(v)$$

At time s, the information needed in order to apply the model is the vector of the past states of the return process $i^s_{s-m+1} := \{i_{s-m+1}, i_{s-m+2}, \dots, i_{s-1}, i_s\}$. Once the vector i^s_{s-m+1} is known, we can calculate the value of the index process:

$$V^m_s = \frac{\sum_{k=0}^{m-1} f(J_{s-k})}{m} = \frac{\sum_{k=0}^{m-1} f(i_{s-k})}{m} \tag{5.22}$$

Let us define:

$$T_{i^s_{s-m+1}}(I_a) := \inf\{^{n\in\mathbb{N}}_{n\geq s}\colon V^m_n \in I_a | J^s_{s-m+1} = i^s_{s-m+1}\} \tag{5.23}$$

as the first time (successive of the current time s) in which the Index process enters in the generic interval $I_a = (\psi_{a-1}, \psi_a)$.

Also, let

$$g_{i^s_{s-m+1}}(I_a; s+n) = \mathbb{P}(T_{i^s_{s-m+1}}(I_a) = s+n | J^s_{s-m+1} = i^s_{s-m+1}) \tag{5.24}$$

be the probability distribution of the first entrance time in I_a of the Index process. Let us assume that

$$g_{i^s_{s-m+1}}(I_a; s) = 0 \ ; \ \forall I_a \ , \ \forall i^s_{s-m+1}$$

Proposition 1 *(explicit formula for g)*

$$g_{i^s_{s-m+1}}(I_a; s+n) = \sum_{(i_{s+1},\dots,i_{s+m-1}),i_{s+n}}^{E^c,E} \prod_{r=1}^{n} P_{i_{s+r-1},i_{s+r}} \frac{\sum_{k=0}^{m-1} f(i_{s+r-1-k})}{m} \tag{5.25}$$

where the symbol

$$\sum_{(i_{s+1},\dots,i_{s+m-1}),i_{s+n}}^{E^c,E} := \sum_{i_{s+1}\in E^c_{i^0_{s-m+1}}(I_a)} \sum_{i_{s+2}\in E^c_{i^1_{s-m+2}}(I_a)} \dots$$
$$\dots \sum_{i_{s+m-1}\in E^c_{i^{s+n-2}_{s-m+(n-1)}}(I_a)} \sum_{i_{s+n}\in E^c_{i^{s+n-1}_{s-m+n}}(I_a)}$$

and

$$E_{i^s_{s-m+1}}(I_a) := \{k \in E : \frac{\sum_{k=0}^{m-1} f(i_{s+1-k})}{m} \in I_a\}$$

also, $E^c_{i^s_{s-m+1}}(I_a)$ is the complementary of $E_{i^s_{s-m+1}}(I_a)$.

Note that $E_{i^s_{s-m+1}}(I_a)$ is the sub-interval of the space state E that, through the estimation of i^s_{s-m+1}, let the Index process enter in the interval I_a with the next transition.

Proof 1 *For $n = 1$, Proposition 1 is true. In fact:*

$$\begin{aligned}
&g_{i^s_{s-m+1}}(I_a; s+1) = \mathbb{P}(T_{is-m+1^s}(I_a) = s+1 | J^s_{s-m+1} = i^s_{s-m+1}) \\
&= \sum_{i_{s+1} \in E} \mathbb{P}(T_{i^s_{s-m+1}}(I_a) = s+1, J(s+1) = i_{s+1} | J^s_{s-m+1} = i^s_{s-m+1}) \\
&= \sum_{i_{s+1} \in E} \mathbb{P}(V^m_{s+1} \in (I_a), J(s+1) = i_{s+1} | J^s_{s-m+1} = i^s_{s-m+1}) \\
&= \sum_{i_{s+1} \in E} \mathbb{P}(V^m_{s+1} \in (I_a) | J^{s+1}_{s-m+1} = i^{s+1}_{s-m+1}) \\
&* \mathbb{P}(J(s+1) = i_{s+1} | J^s_{s-m+1} = i^s_{s-m+1})) \\
&= \sum_{i_{s+1} \in E_{i^s_{s-m+1}}(I_a)} 1 * \mathbb{P}_{i_s, i_{s+1}} \left(\frac{\sum_{k=0}^{m-1} f(i_{s-k})}{m} \right)
\end{aligned} \tag{5.26}$$

that is equal to Proposition 1 for $n = 1$.

Now, let us suppose that Proposition 1 is true for $n-1$ and that it is also true that:

$$g_{i^s_{s-m+1}}(I_a; s+m-1) = \sum_{(i_{s+1}, \ldots, i_{s+m-2}, i_{s+m-1})}^{E^c, E} \prod_{r=1}^{(s+m-1)-s} \mathbb{P}_{i_{s+r-1}, i_{s+r}} \left(\frac{\sum_{k=0}^{m-1} f(i_{s-k})}{m} \right)$$

We can calculate:

$$
\begin{aligned}
& g_{i^s_{s-m+1}}(I_a; s+n) = \mathbb{P}(T_{i^s_{s-m+1}}(I_a) = s+n | J^s_{s-m+1} = i^s_{s-m+1}) \\
&= \sum_{i_{s+1} \in E^c_{i^s_{s-m+1}}(I_a)} \mathbb{P}(T_{i^s_{s-m+1}}(I_a) = s+n, J(s+1) = i_{s+1} | J^s_{s-m+1} = i^s_{s-m+1}) \\
&= \sum_{i_{s+1} \in E^c_{i^s_{s-m+1}}(I_a)} \mathbb{P}(T_{i^s_{s-m+1}}(I_a) = s+n | J(s+1) = i_{s+1}, J^s_{s-m+1} = i^s_{s-m+1}) \\
&\qquad * \mathbb{P}(J(s+1) = i_{s+1} | J^s_{s-m+1} = i^s_{s-m+1}) \\
&= \sum_{i_{s+1} \in E^c_{i^s_{s-m+1}}(I_a)} \mathbb{P}(T_{i^s_{s-m+1}}(I_a) = s+n | J^s_{s-m+1} = i^s_{s-m+1}) \\
&\qquad * \mathbb{P}_{i_s, i_{s+1}}\left(\frac{\sum_{k=0}^{m-1} f(i_{s-k})}{m}\right) \\
&= \sum_{i_{s+1} \in E^c_{i^s_{s-m+1}}(I_a)} g_{i^{s+1}_{s-m+2}}(I_a; s+n) * \mathbb{P}_{i_s, i_{s+1}}\left(\frac{\sum_{k=0}^{m-1} f(i_{s-k})}{m}\right) \\
&= (from\ inductive\ hypothesis) \\
&= \sum_{i_{s+1} \in E^c_{i^s_{s-m+1}}(I_a)} \left[\sum_{(i_{s+2}, \ldots, i_{s+m-2}), i_{s+m-1}}^{E^c, E} \right. \\
&\prod_{r=1}^{s+n-(s+1)} \left. \mathbb{P}_{i_{s+1-r-1}, i_{s+1+r}}\left(\frac{\sum_{k=0}^{m-1} f(i_{s+1,r-1-k})}{m}\right) \right] * \mathbb{P}_{i_s, i_{s+1}}\left(\frac{\sum_{k=0}^{m-1} f(i_{s-k})}{m}\right) \\
&\qquad = Propostion\ 1
\end{aligned}
\tag{5.27}
$$

5.5 Empirical application of the change-point in IMC

5.5.1 Database description

The methodology described so far has been applied to the study of intra-day prices of a quoted Italian firm. The sample data of prices starts on January 1st, 2007 and ends on December 31st 2010. The dataset is obtained from $www.borsaitaliana.it$ and it contains tick-by-tick quotes of the traded stocks. The data have been re-sampled to have 1 minute frequency according to the procedure described below. Let us consider a single day k with $1 \leq k \leq d$ where d is number of traded days in the time series. Since we consider four full years of trading, the number of trading days is $d = 1076$.

The stock market in Italy operates as follows: first the opening price is fixed at a random time in the first minute after 9 am; immediately after that a continuous trading of the stock starts and it ends just before 5:25 pm; finally the closing price is fixed just after 5:30 pm. Let us denote with $S(t)$ the time varying asset price of the stock under study. Therefore, we can define $S(k_1)$ as the price of the last trading before 9:01:00 am, $S(k_2)$ as the price of the last trading before 9:02:00 am and so on. If there are no transactions during a generic minute z with $1 \leq z \leq n$ considered, the price $S(k_z)$ is set equal to $S(k_{z-1})$. The price is considered to be unchanged also in the case where the title is suspended and reopened in the same day. Let us denote with $S(k_n)$ the price of the last trading before 5:25:00 pm. As a consequence, the number of prices for each day is $n = 507$. It should be noticed that before the 28th of September 2009, the continuous trading started at 9:05 am and therefore, for the time period from the 1st of January 2007 to the 28th of September 2009, $n = 502$. Also, the title may have a delay in the opening and/or a

closing in advance, e.g. a title suspended but not reopened. In these cased, the affective number of trading minutes is considered and as a result n might be less then 502 or less then 507 depending of the period before or after the 28th of September 2009, respectively. Finally, the number of price data for the stock under study is approximately 500 thousand.

The price returns at time t are calculated as

$$R(t) = \frac{S(t+1) - S(t)}{S(t)}$$

with $R(t) \in \mathbb{R}$ and are shown in Figure 5.1.

The returns are then converted into a series of states denoted by $\{J_n\}_{n\in\mathbb{N}}$ with values in E space of states. The index value for each value of J_n is then calculated using Eq. (**??**) and choosing a memory $m = 100$. The index process is denoted by $\{V_n^m\}_{n\in\mathbb{N}}$. As mentioned above, the value of the index denotes the level of the volatility of the prices: higher the value of the index, higher the price volatility.

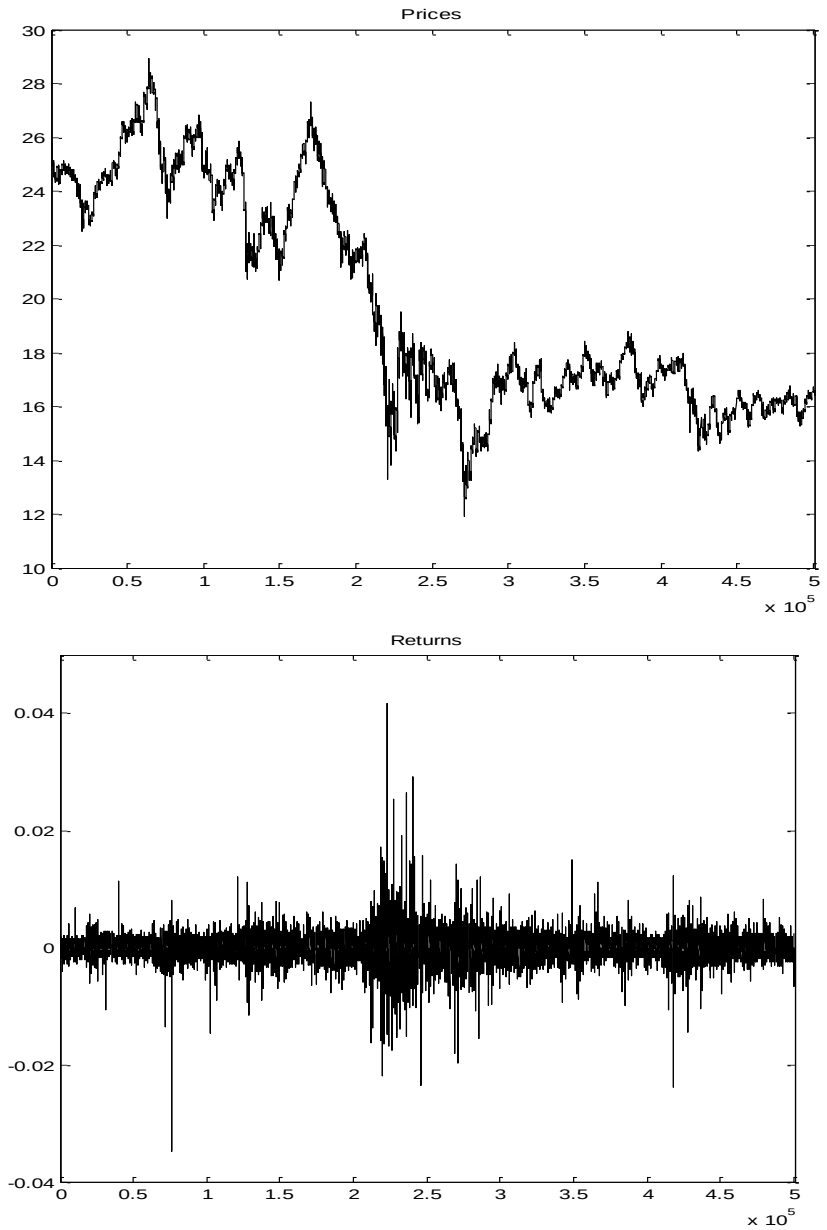

Figure 5.1: Prices and returns of the traded stocks

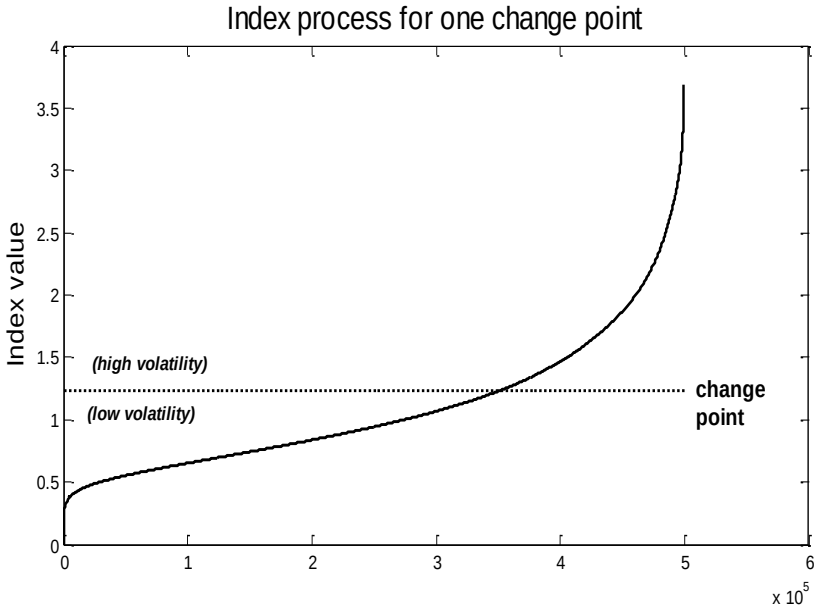

Figure 5.2: Discretization of the index process in case of one change point

5.5.2 The index process in case of one change-point

Firstly, let us suppose that there can be only two levels of the volatility: high and low. As a consequence, the index process should also be discretized in these two states. In order to detect the optimal value of the index that subdivides the process in two states we use the fact that the price return process $\{J_n\}$ presents different dynamics, i.e. different transition matrices, based on the level of the volatility. Thus, in our work the change point is identified as the value of the index which would maximize the differences in the price return dynamics.

Since there is no closed form for identifying the change point we operate algorithmically by using the maximum likelihood procedure described in Section 5.3. The estimated change point and the two probability matrices of the price return process in each state of the index are calculated for each of the stocks. Figure 5.2 shows the discretization of the index process in two states

by using the change point identified through the maximum log-likelihood procedure.

A well known characteristic of financial time series is the "volatility clustering" which simply states that periods of high/low volatility in the market tend to be followed by periods of high/low volatility. It seems useful to point out that this characteristic is confirmed by analyzing the following two transition probability matrices:

$$\hat{\overline{P}}(\psi_1) = \begin{pmatrix} 0.173 & 0.151 & 0.207 & 0.218 & 0.251 \\ 0.129 & 0.196 & 0.267 & 0.255 & 0.153 \\ 0.137 & 0.209 & 0.300 & 0.221 & 0.133 \\ 0.150 & 0.246 & 0.275 & 0.204 & 0.125 \\ 0.237 & 0.215 & 0.218 & 0.159 & 0.171 \end{pmatrix}$$

;

$$\hat{\underline{P}}(\psi_1) = \begin{pmatrix} 0.067 & 0.162 & 0.312 & 0.338 & 0.121 \\ 0.031 & 0.183 & 0.391 & 0.347 & 0.048 \\ 0.033 & 0.236 & 0.466 & 0.234 & 0.031 \\ 0.049 & 0.339 & 0.397 & 0.185 & 0.030 \\ 0.110 & 0.338 & 0.316 & 0.170 & 0.066 \end{pmatrix}$$

In fact, it can be noticed that in the case of high volatility the transition matrix $\hat{\overline{P}}(\psi_1)$ presents higher probabilities of transiting in the most extreme states than the transition matrix $\hat{\underline{P}}(\psi_1)$. This simply confirms the fact that in a high volatility market, the probability of experiencing large variations of the asset price (i.e. strongly positive returns followed by strongly negative returns and vice versa) is higher than the case of low volatility.

Once the matrices are estimated, the next step is that of determining whether they are statistically different. For this purpose we calculated two indices that can be used as a measure of the difference between the two

Table 5.1: % Root square mean deviation and % mean absolute deviation

Matrices	% RSMD	% MAD
$\hat{\overline{P}}(\psi_1)$, $\hat{\underline{P}}(\psi_1)$	49.3%	44.2%

matrices. In particular, we used the *percentage root mean square deviation (% RSMD)* and the *percentage mean absolute deviation (%MAD)*. The formulas are given in Eq. 5.28 and Eq. 5.29, respectively.

$$\%RSMD = \sqrt{\frac{\sum_{ij}(\overline{p}_{ij} - \underline{p}_{ij})^2}{n}} * \frac{n * 100\%}{\sum_{ij} \underline{p}_{ij}} \tag{5.28}$$

$$\%MAD = \frac{\sum_{ij} |\overline{p}_{ij} - \underline{p}_{ij}|}{\sum_{ij} \underline{p}_{ij}} * 100\% \tag{5.29}$$

Results are shown in Table 5.1. Since, the values of these two indices are a measure of the differences of the estimated matrices, higher the value of the index, higher the distance between the two matrices and higher the probability that there actually is a difference in the process.

The next step is to test statistically whether the two transition probability matrices are different. In order to construct the test, we use the procedure described in Section 5.3.

As mentioned before, the distribution of the statistic test under the null hypothesis is not known and thus it has to be approximated through the bootstrap methodology. We have simulated *1000* trajectories of the same length of the data from a single transition probability matrix estimated considering the whole dataset. The histogram of the simulated statistic test as well as the kernel fitting are derived and are shown in Figure 5.3.

We chose a level of significance of 5%. The results of the test for a single change point of the stock are shown in Table 5.2. The p-value strongly

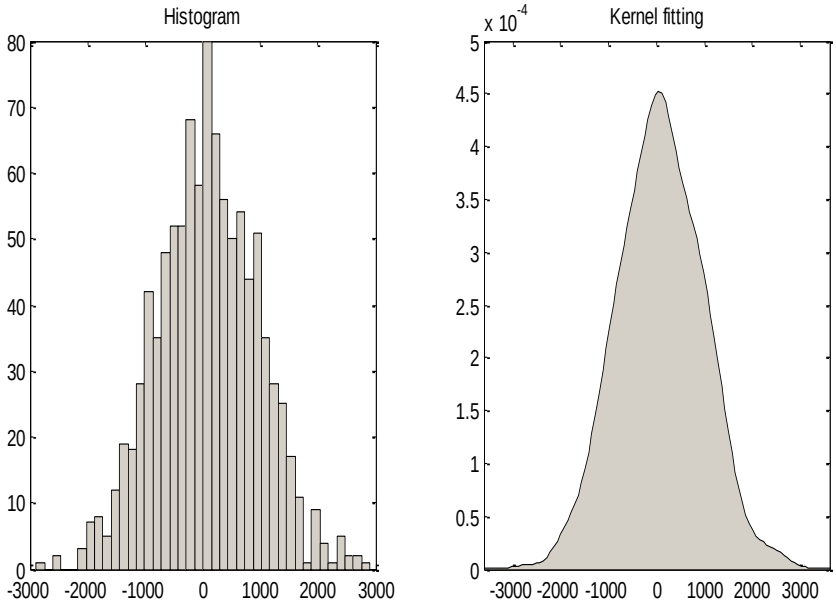

Figure 5.3: Histogram and kernel fitting of the simulated statistic test

Table 5.2: Results of the statistic test on one change point

	ψ_1	$\mathfrak{D}$	$\mathfrak{D}^*(0.95)$	$\mathfrak{D}^*(0.99)$	Empirical p-value
Value	1.23	32400	3290	3580	0.000

suggests to not accept the null hypothesis and thus we conclude that there is a statistical difference between the two transition probability matrices.

5.5.3 Identifying the optimal number of change points

The case of more then one change point has also been considered. The procedure is similar to the case of one change point described so far.

Two change points. In the case of two change points, we suppose that there are three levels of volatility in the market, i.e. low, medium and high. In this case, in order to find the maximum of the log-likelihood function, all the possible combinations of the two change points have to be considered.

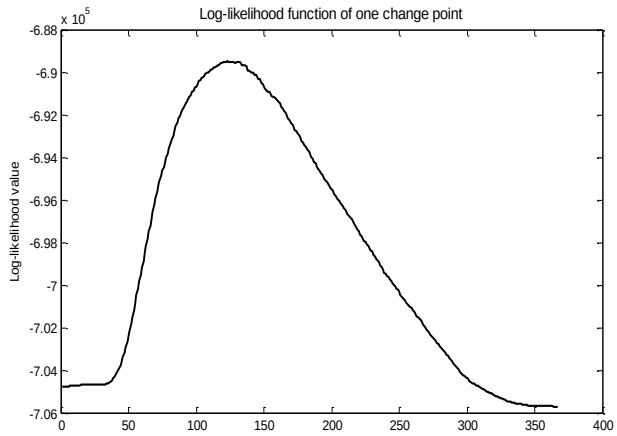

Figure 5.4: Log-likelihood function: one change point

Table 5.3: Results of the statistic test on two change point

	ψ_1	ψ_2	$\mathfrak{D}$	$\mathfrak{D}^*(0.95)$	$\mathfrak{D}^*(0.99)$	Empirical p-value
Value	0.90	1.60	42000	3300	3590	0.000

As a consequence, the log-likelihood is a bivariate function of the two change points. In Figure 5.4, the log-likelihood function in case of one change point (the univariate function) is shown. The representation of the bivariate

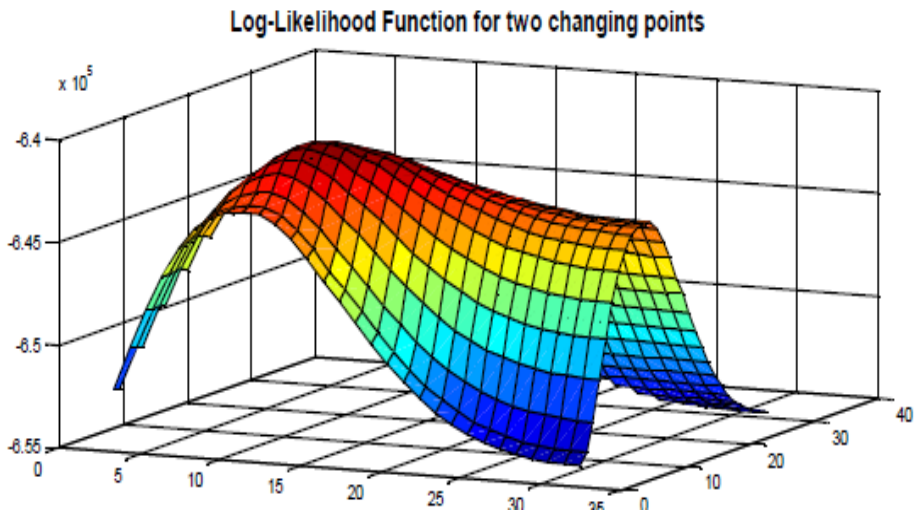

Figure 5.5: Log-likelihood function: two change points

case, i.e. the log-likelihood function in case of two change points is shown in Figure 5.5.

From the maximization of the log-likelihood function we obtained the values of the two change points (see Table 5.3) and the three transition probability matrices in each of the classes of volatility.

The estimated transition probability matrices are the following:

$$\hat{P}(1) = \begin{pmatrix} 0.055 & 0.151 & 0.332 & 0.361 & 0.101 \\ 0.021 & 0.168 & 0.416 & 0.360 & 0.035 \\ 0.024 & 0.230 & 0.496 & 0.228 & 0.022 \\ 0.036 & 0.352 & 0.424 & 0.167 & 0.021 \\ 0.094 & 0.357 & 0.339 & 0.156 & 0.054 \end{pmatrix}; \hat{P}(2) = \begin{pmatrix} 0.096 & 0.172 & 0.273 & 0.300 & 0.159 \\ 0.067 & 0.212 & 0.325 & 0.309 & 0.087 \\ 0.073 & 0.244 & 0.370 & 0.245 & 0.068 \\ 0.089 & 0.300 & 0.331 & 0.219 & 0.061 \\ 0.142 & 0.302 & 0.281 & 0.183 & 0.092 \end{pmatrix}$$

$$\hat{P}(3) = \begin{pmatrix} 0.202 & 0.140 & 0.184 & 0.187 & 0.287 \\ 0.172 & 0.178 & 0.232 & 0.216 & 0.202 \\ 0.184 & 0.181 & 0.254 & 0.203 & 0.178 \\ 0.191 & 0.208 & 0.238 & 0.187 & 0.176 \\ 0.272 & 0.185 & 0.195 & 0.147 & 0.201 \end{pmatrix}$$

The statistic test shows that also in this case the null hypothesis cannot be accepted. We calculated the the *percentage root mean square deviation (% RSMD)* and the *percentage mean absolute deviation (%MAD)* (Table 5.4 and Table 5.5 respectively) and performed the statistic test (see Table 5.3).

Table 5.4: % Root square mean deviation: two change points

	$\hat{P}(1)$	$\hat{P}(2)$	$\hat{P}(3)$
$\hat{P}(1)$	0.0%	28.4%	73.0%
$\hat{P}(2)$	28.4%	0.0%	47.8%
$\hat{P}(3)$	73.0%	47.8%	0.0%

Table 5.5: % Mean absolute deviation: two change points

	$\hat{P}(1)$	$\hat{P}(2)$	$\hat{P}(3)$
$\hat{P}(1)$	0.0%	25.8%	65.3%
$\hat{P}(2)$	25.8%	0.0%	45.2%
$\hat{P}(3)$	65.3%	45.2%	0.0%

Three change points. The procedure for more than two change points is substantially the same and thus we summarize the results obtained for the case of three and four change points.

Table 5.6: Values of the Index process in case of three change points

	ψ_1	ψ_2	ψ_3
Value	0.80	1.30	2.00

The estimated transition probability matrices are the following:

$$\hat{P}(1) = \begin{pmatrix} 0.051 & 0.145 & 0.346 & 0.368 & 0.090 \\ 0.017 & 0.157 & 0.429 & 0.367 & 0.030 \\ 0.020 & 0.226 & 0.511 & 0.224 & 0.019 \\ 0.032 & 0.357 & 0.435 & 0.159 & 0.017 \\ 0.086 & 0.361 & 0.345 & 0.153 & 0.055 \end{pmatrix} ; \hat{P}(2) = \begin{pmatrix} 0.077 & 0.171 & 0.294 & 0.321 & 0.137 \\ 0.049 & 0.211 & 0.347 & 0.323 & 0.070 \\ 0.055 & 0.250 & 0.396 & 0.249 & 0.050 \\ 0.069 & 0.316 & 0.354 & 0.214 & 0.047 \\ 0.122 & 0.327 & 0.297 & 0.179 & 0.075 \end{pmatrix}$$

$$\hat{P}(3) = \begin{pmatrix} 0.140 & 0.167 & 0.232 & 0.249 & 0.212 \\ 0.116 & 0.202 & 0.278 & 0.266 & 0.138 \\ 0.126 & 0.217 & 0.308 & 0.229 & 0.120 \\ 0.138 & 0.257 & 0.284 & 0.212 & 0.109 \\ 0.194 & 0.246 & 0.255 & 0.175 & 0.130 \end{pmatrix} ; \hat{P}(4) = \begin{pmatrix} 0.234 & 0.125 & 0.158 & 0.162 & 0.321 \\ 0.218 & 0.160 & 0.191 & 0.182 & 0.249 \\ 0.232 & 0.150 & 0.216 & 0.173 & 0.229 \\ 0.240 & 0.175 & 0.195 & 0.163 & 0.227 \\ 0.311 & 0.157 & 0.163 & 0.133 & 0.236 \end{pmatrix}$$

Table 5.7: % Root square mean deviation: three change points

	$\hat{P}(1)$	$\hat{P}(2)$	$\hat{P}(3)$	$\hat{P}(4)$
$\hat{P}(1)$	0.0%	24.2%	51.4%	93.0%
$\hat{P}(2)$	24.2%	0.0%	29.7%	73.6%
$\hat{P}(3)$	51.4%	29.7%	0.0%	44.5%
$\hat{P}(4)$	93.0%	73.6%	44.5%	0.0%

Table 5.8: % Mean absolute deviation: three change points

	$\hat{P}(1)$	$\hat{P}(2)$	$\hat{P}(3)$	$\hat{P}(4)$
$\hat{P}(1)$	0.0%	21.8%	46.1%	83.4%
$\hat{P}(2)$	21.8%	0.0%	26.8%	70.0%
$\hat{P}(3)$	46.1%	26.8%	0.0%	43.0%
$\hat{P}(4)$	83.4%	70.0%	43.0%	0.0%

Four change points. The results of the algorithm in case of four change points are shown below.

Table 5.9: Values of the Index process in case of four change points

	ψ_1	ψ_2	ψ_3	ψ_4
Value	0.70	1.00	1.40	2.10

The estimated transition probability matrices are the following:

$$\hat{P}(1) = \begin{pmatrix} 0.046 & 0.143 & 0.351 & 0.378 & 0.082 \\ 0.014 & 0.141 & 0.445 & 0.374 & 0.026 \\ 0.016 & 0.219 & 0.532 & 0.217 & 0.016 \\ 0.027 & 0.365 & 0.448 & 0.146 & 0.014 \\ 0.084 & 0.360 & 0.358 & 0.147 & 0.051 \end{pmatrix} ; \hat{P}(2) = \begin{pmatrix} 0.064 & 0.162 & 0.314 & 0.344 & 0.116 \\ 0.033 & 0.202 & 0.375 & 0.340 & 0.050 \\ 0.039 & 0.249 & 0.429 & 0.248 & 0.035 \\ 0.053 & 0.333 & 0.382 & 0.199 & 0.033 \\ 0.107 & 0.343 & 0.321 & 0.169 & 0.060 \end{pmatrix}$$

$$\hat{P}(3) = \begin{pmatrix} 0.091 & 0.171 & 0.279 & 0.304 & 0.154 \\ 0.067 & 0.215 & 0.321 & 0.308 & 0.089 \\ 0.073 & 0.248 & 0.367 & 0.245 & 0.067 \\ 0.086 & 0.300 & 0.332 & 0.222 & 0.060 \\ 0.135 & 0.312 & 0.283 & 0.182 & 0.088 \end{pmatrix} ; \hat{P}(4) = \begin{pmatrix} 0.150 & 0.166 & 0.224 & 0.238 & 0.222 \\ 0.127 & 0.199 & 0.269 & 0.256 & 0.149 \\ 0.138 & 0.208 & 0.298 & 0.222 & 0.134 \\ 0.150 & 0.245 & 0.276 & 0.206 & 0.123 \\ 0.206 & 0.236 & 0.245 & 0.172 & 0.141 \end{pmatrix}$$

$$\hat{P}(3) = \begin{pmatrix} 0.239 & 0.121 & 0.153 & 0.155 & 0.332 \\ 0.232 & 0.151 & 0.182 & 0.169 & 0.266 \\ 0.241 & 0.142 & 0.207 & 0.168 & 0.242 \\ 0.249 & 0.168 & 0.182 & 0.161 & 0.240 \\ 0.320 & 0.150 & 0.155 & 0.130 & 0.245 \end{pmatrix}$$

Table 5.10: % Root square mean deviation: four change points

	$\hat{P}(1)$	$\hat{P}(2)$	$\hat{P}(3)$	$\hat{P}(4)$	$\hat{P}(5)$
$\hat{P}(1)$	0.0%	20.1%	35.8%	59.4%	100.5%
$\hat{P}(2)$	20.1%	0.0%	16.8%	43.0%	86.1%
$\hat{P}(3)$	35.8%	16.8%	0.0%	27.1%	70.9%
$\hat{P}(4)$	59.4%	43.0%	27.1%	0.0%	44.1%
$\hat{P}(5)$	100.5%	86.1%	70.9%	44.1%	0.0%

Table 5.11: % Mean absolute deviation: four change points

	$\hat{P}(1)$	$\hat{P}(2)$	$\hat{P}(3)$	$\hat{P}(4)$	$\hat{P}(5)$
$\hat{P}(1)$	0.0%	17.2%	32.2%	53.4%	90.1%
$\hat{P}(2)$	17.2%	0.0%	15.2%	38.6%	80.6%
$\hat{P}(3)$	32.2%	15.2%	0.0%	25.2%	67.7%
$\hat{P}(4)$	53.4%	38.6%	25.2%	0.0%	42.6%
$\hat{P}(5)$	90.1%	80.6%	67.7%	42.6%	0.0%

Five change points. The results of the algorithm in case of five change points are shown below.

Table 5.12: Values of the Index process in case of five change points

	ψ_1	ψ_2	ψ_3	ψ_4	ψ_5
Value	0.20	0.70	1.00	1.40	2.10

Once the border values of the Index process are identified, the probability transition matrices of the price return process in each state of volatility can be estimated through the maximum likelihood estimators. The estimated matrices are shown below.

$$\hat{P}(1) = \begin{pmatrix} 0.000 & 0.143 & 0.428 & 0.286 & 0.143 \\ 0.000 & 0.244 & 0.356 & 0.378 & 0.022 \\ 0.002 & 0.011 & 0.973 & 0.013 & 0.001 \\ 0.093 & 0.326 & 0.395 & 0.163 & 0.023 \\ 0.000 & 0.750 & 0.000 & 0.250 & 0.000 \end{pmatrix} ; \hat{P}(2) = \begin{pmatrix} 0.045 & 0.144 & 0.354 & 0.377 & 0.080 \\ 0.014 & 0.140 & 0.446 & 0.375 & 0.025 \\ 0.017 & 0.223 & 0.524 & 0.220 & 0.016 \\ 0.027 & 0.366 & 0.450 & 0.144 & 0.013 \\ 0.081 & 0.358 & 0.363 & 0.147 & 0.051 \end{pmatrix}$$

$$\hat{P}(3) = \begin{pmatrix} 0.064 & 0.161 & 0.315 & 0.344 & 0.116 \\ 0.032 & 0.200 & 0.377 & 0.341 & 0.050 \\ 0.038 & 0.248 & 0.431 & 0.249 & 0.034 \\ 0.052 & 0.334 & 0.384 & 0.197 & 0.033 \\ 0.106 & 0.346 & 0.321 & 0.168 & 0.059 \end{pmatrix} ; \hat{P}(4) = \begin{pmatrix} 0.090 & 0.172 & 0.280 & 0.307 & 0.151 \\ 0.066 & 0.214 & 0.323 & 0.309 & 0.088 \\ 0.072 & 0.249 & 0.368 & 0.245 & 0.066 \\ 0.084 & 0.300 & 0.334 & 0.223 & 0.059 \\ 0.135 & 0.313 & 0.281 & 0.183 & 0.088 \end{pmatrix}$$

$$\hat{P}(5) = \begin{pmatrix} 0.150 & 0.165 & 0.225 & 0.239 & 0.221 \\ 0.126 & 0.199 & 0.269 & 0.257 & 0.149 \\ 0.136 & 0.209 & 0.299 & 0.224 & 0.132 \\ 0.149 & 0.246 & 0.276 & 0.207 & 0.122 \\ 0.204 & 0.236 & 0.247 & 0.172 & 0.141 \end{pmatrix} ; \hat{P}(6) = \begin{pmatrix} 0.238 & 0.122 & 0.154 & 0.156 & 0.330 \\ 0.230 & 0.152 & 0.183 & 0.171 & 0.264 \\ 0.241 & 0.143 & 0.207 & 0.168 & 0.241 \\ 0.248 & 0.169 & 0.184 & 0.160 & 0.239 \\ 0.319 & 0.151 & 0.156 & 0.130 & 0.244 \end{pmatrix}$$

Table 5.13: % Root square mean deviation: five change points

	$\hat{P}(1)$	$\hat{P}(2)$	$\hat{P}(3)$	$\hat{P}(4)$	$\hat{P}(5)$	$\hat{P}(5)$
$\hat{P}(1)$	0.0%	80.2%	85.2%	91.8%	106.3%	134.9%
$\hat{P}(2)$	80.2%	0.0%	19.5%	35.3%	59.1%	100.2%
$\hat{P}(3)$	85.2%	19.5%	0.0%	16.8%	43.0%	86.0%
$\hat{P}(4)$	91.8%	35.3%	16.8%	0.0%	27.2%	71.0%
$\hat{P}(5)$	106.3%	59.1%	43.0%	27.2%	0.0%	44.1%
$\hat{P}(6)$	134.9%	100.2%	86.0%	71.0%	44.1%	0.0%

Table 5.14: % Mean absolute deviation: five change points

	$\hat{P}(1)$	$\hat{P}(2)$	$\hat{P}(3)$	$\hat{P}(4)$	$\hat{P}(5)$	$\hat{P}(5)$
$\hat{P}(1)$	0.0%	51.3%	52.0%	59.4%	78.7%	110.2%
$\hat{P}(2)$	51.3%	0.0%	16.8%	31.9%	53.3%	90.1%
$\hat{P}(3)$	52.0%	16.8%	0.0%	15.3%	38.6%	80.4%
$\hat{P}(4)$	59.4%	31.9%	15.3%	0.0%	25.2%	67.8%
$\hat{P}(5)$	78.7%	53.3%	38.6%	25.2%	0.0%	42.6%
$\hat{P}(6)$	110.2%	90.1%	80.4%	67.8%	42.6%	0.0%

Optimal number of change points. In order to decide the optimal number of change points methods based on AIC and BIC as defined in Eq. (5.19) and Eq. (5.20), respectively can be used. As mentioned, it does not exist a closed form for the calculation of the optimal number of change points and thus we operate in an iterative way. This way the most parsimonious model can be identified as the one that minimizes the two indices mentioned before. It should be remarked that the algorithm reaches very slowly the minimum values of AIC or BIC. We decided to also calculated the improvement of

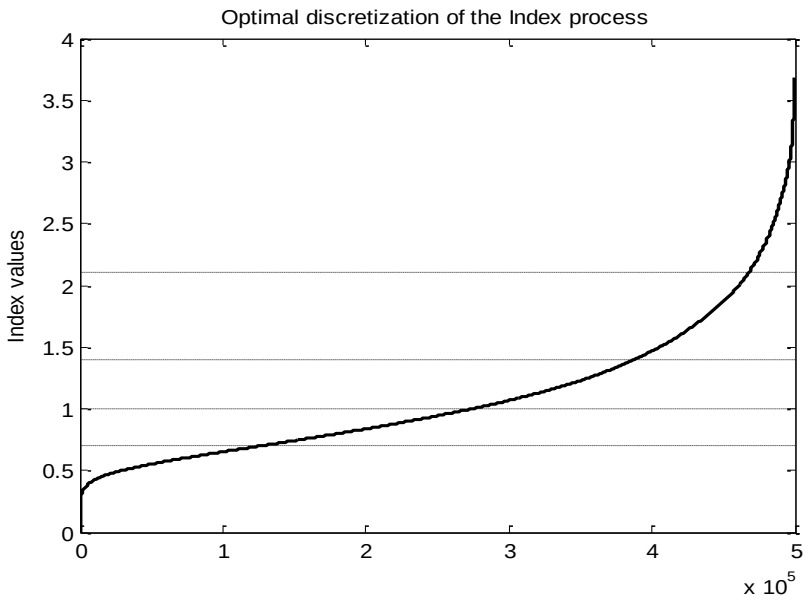

Figure 5.6: Optimal discretization of the Index process

an additional change point in terms of a percentage variation of the two indices. We considered a limit of less than 0.1% as a sign that the algorithm is reaching its minimum. Results are summarized in Table 5.15.

We identified the optimal number of change points to be four since the BIC index reached an improvement level of less than 0.1%. As a consequence, the Index process is divided in five states representing five levels of the volatility in the market: very low, medium-low, medium, medium-high and very high.

The border values of the index process obtained in the case of four change points have already been given in Table 5.9 in the above paragraph as well as the estimated transition probability matrices. Also, in Figure 5.6 the discretization of the Index process according to the estimated change points is shown.

Table 5.15: Values of AIC and BIC for various number of change points

k	**1**	**2**	**3**	**4**	**5**
$\mathfrak{D}$	32400	42000	46100	48300	49800
%Δ		29.6%	9.8%	4.8%	3.1%
AIC	1379000	1370000	1365000	1363000	1362000
%Δ		-0.7%	-0.4%	-0.1%	<-0.1%
BIC	1380000	1371000	1367000	1366000	1365000
%Δ		-0.7%	-0.3%	<-0.1%	<-0.1%

The autocorrelation function. Financial time series present a very important feature: the fact that while the returns are not autocorrelated and present i.i.d behavior, the square of returns or their absolute values are long rage correlated. It is important that the model describing such dynamics presents the same characteristic.

The autocorrelation of the square of returns for various time lags, which we will denote with τ, is given by the Eq. 5.30.

$$\Sigma(\tau) = \frac{Cov(R^2(t+\tau), R^2(t))}{Var(R^2(t))} \tag{5.30}$$

We compared the autocorrelation function of the square of returns of real data with simulated trajectories using the models defined above (see Figure 5.7)

The graph shows that for the model with one change point, the autocorrelation of the square of returns falls rapidly to zero. In choosing two change points, there is a significant improvement in the long range autocorrelation. In adding more change points, the function continues to improve even though at a very slow rate. It is worth noticing that for the model with four change

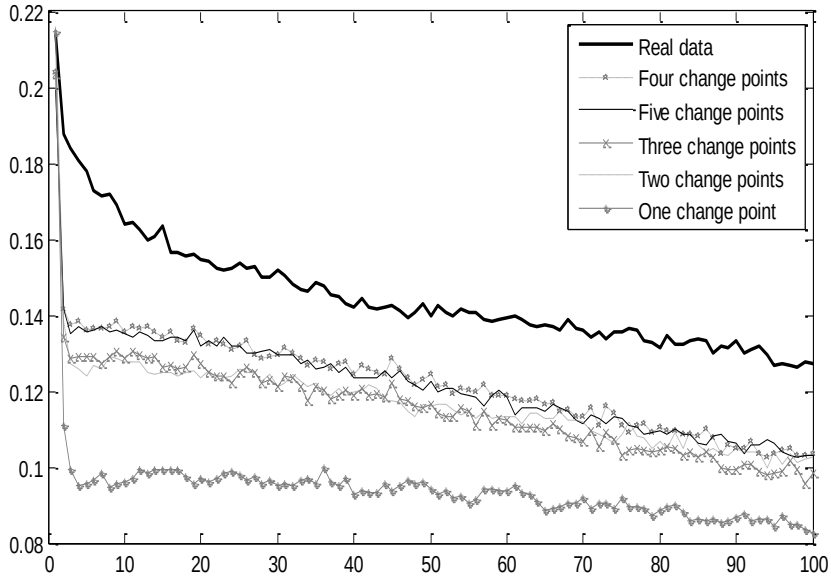

Figure 5.7: Autocorrelation of the square of returns for various number of change points of the Index process

points the autocorrelation function of the simulated data is the closest to the autocorrelation of the real data. Adding one more change point seems to worsen the result (even though the difference with the model with four change points is minimal).

5.6 Conclusions

In this Chapter the change point approach has been used to identify the optimal discretization of the Index process in the Indexed Markov Chain models. It is well known that financial time series present a long range autocorrelation of the square of return and thus models used to describe price returns should have this particular characteristic known also as 'a long memory'. Furthermore, financial time series present another peculiar feature often referred to as 'volatility clustering'. The latter states that there is a certain persistence of the price volatility in the market, i.e. high (low) volatility periods of price returns tend to be followed by high (low) volatility periods. Both these stylized facts of financial series are modeled in the IMC by the introduction of the Index process. In fact, the Index process is calculated as a function of the previous mth (accounting thus for the 'a long memory') of the square of returns (accounting this way for the 'volatility clustering') each exponentially weighted.

In order to construct the model, the Index process has to be discretized in a limited number of states. We applied the change point approach justified by the assumption that the price dynamics of the index are different on different levels of the volatility. We thus estimated, through the maximization of the log-likelihood function, the value of the index process which maximized the difference in the dynamics of the price return process. This was achieved in an iterative way through an algorithm implemented in MatLab.

We also treated the problem of finding the optimal number of change points to consider. In this case we used the methods based on AIC and BIC to find the most parsimonious model. This also has been achieved in the empirical study in an iterative way.

Lastly, we constructed the autocorrelation function of the square of re-

turns for the real data and for the different models with various change points.

From the empirical study, it resulted that the optimal number of change points was four which identified five states of the Index process each corresponding to a particular level of volatility: very low, medium -low, medium, medium-high and very high. The autocorrelation function showed that with four change points, the simulated data presented the closest behavior to the real data.

Chapter 6

Conclusions

In this thesis we tried to make one more step in the application of Markov processes in the actuarial and financial field. Two main problems have been dealt. The first one regarded the application of a Markov process for the description of the salary lines of participants in an Italian Pension Scheme of the First Pillar. A semi-Markov process with backward recurrence time was proposed. A statistic test has been applied in order to determine whether the null hypothesis of a geometrical distribution of the waiting times of the process should be accepted or not. The test showed that the null hypotheses was rejected for some of the waiting time distributions and thus we concluded that the semi-Markov process should be preferred to the simple Markov chain to model the transition in the states of the salary process.

In the financial application, we treated the *Indexed semi-Markov chain*, a new model that has been previously used to describe intra day price return dynamics. The peculiarity of this model is that, through the Index process, it manages two very known stylized facts of financial time series: the first one is the long memory of financial series and the second one is the volatility clustering. This is achieved by defining the Index as a function of the m-th

previous values of the price returns.

In order to transform the values obtained in states of a stochastic process a discretization of the Index is necessary. We proposed the method of change points as a new method to obtain the most efficient classes. This approach is justified by the fact that, for financial time series, the price dynamics in different levels of volatility in the market present different characteristics.

The theoretical framework has been set by considering various situation: the case of one known change point, of one unknown change point of more then one unknown change point. In all these cases the maximum likelihood estimations of the transition probability matrices are defined. The methodology has then been applied to study the dynamics of intra-day price returns of a quoted Italian firm. Different models, with different number of change points have been hypothesized. Tests, based on AIC and BIC, are performed in order to determine the most parsimonious model. We found out that the best discretization of the Index process is that of using four change points, which implied five levels of volatility in the market: very low, medium low, medium, medium high and very high. We also generated synthetic trajectories in order to calculate the autocorrelation of the square of returns for the real data as well as for the hypothesized models. The autocorrelation function showed that the model with four change points was the closest to the real data.

List of Figures

List of Tables

Bibliography

Balcer, Y. and Sahin, I. (1983). A stochastic theory of pension dynamics. *Insurance: Mathematics and Economics*, 2(3):179 – 197.

Balcer, Y. and Sahin, I. (1986). Pension accumulation as a semi-markov reward process, with applications to pension reform. In *Semi-Markov models*, pages 181–199. Springer.

Barbu, V., Boussemart, M., and Limnios, N. (2004). Discrete-time semi-markov model for reliability and survival analysis. *Communications in Statistics-Theory and Methods*, 33(11):2833–2868.

Brunetti, C., Scotti, C., Mariano, R. S., and Tan, A. H. (2008). Markov switching garch models of currency turmoil in southeast asia. *Emerging Markets Review*, 9(2):104–128.

Cai, J. (1994). A markov model of switching-regime arch. *Journal of Business & Economic Statistics*, 12(3):309–316.

Chu, C.-S. J., Santoni, G. J., and Liu, T. (1996). Stock market volatility and regime shifts in returns. *Information Sciences*, 94(1):179–190.

D'Amico, G., Guillen, M., and Manca, R. (2009a). Full backward non-homogeneous semi-markov processes for disability insurance models: a

catalunya real data application. *Insurance: Mathematics and Economics*, 45(2):173–179.

D'amico, G., Guillen, M., and Manca, R. (2013). Semi-markov disability insurance models. *Communications in Statistics - Theory and Methods*, 42(16):2872–2888.

D'Amico, G., Janssen, J., and Manca, R. (2009b). European and american options: The semi-markov case. *Physica A: Statistical Mechanics and its Applications*, 388(15):3181–3194.

D'Amico, G. and Petroni, F. (2011). A semi-markov model with memory for price changes. *Journal of Statistical Mechanics: Theory and Experiment*, 2011(12):P12009.

D'Amico, G. and Petroni, F. (2012a). A semi-markov model for price returns. *Physica A: Statistical Mechanics and its Applications*, 391(20):4867–4876.

D'Amico, G. and Petroni, F. (2012b). Weighted-indexed semi-markov models for modeling financial returns. *Journal of statistical mechanics: theory and experiment*, 2012(07):P07015.

D'Amico, G. and Petroni, F. (2014). Multivariate high-frequency financial data via semi-markov processes. *Markov Processes Relat. Fields*, (20):415–434.

De Dominicis, R. and Manca, R. (1986). Some new results on the transient behaviour of semi-markov reward processes. *Methods of Operations Research*, 54:387–397.

Edwards, S. and Susmel, R. (2003). Interest-rate volatility in emerging markets. *Review of Economics and Statistics*, 85(2):328–348.

Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica: Journal of the Econometric Society*, pages 357–384.

Hamilton, J. D. and Susmel, R. (1994). Autoregressive conditional heteroskedasticity and changes in regime. *Journal of Econometrics*, 64(1):307–333.

Henry, O. T. (2009). Regime switching in the relationship between equity returns and short-term interest rates in the uk. *Journal of Banking & Finance*, 33(2):405–414.

Hoem, J. M. (1969). Markov chain models in life insurance. *Blätter der DGVFM*, 9(2):91–107.

Hoem, J. M. (1972). Inhomogeneous semi-markov prosesses, select actuarial tables, and duration-dependence in demography. *Working papers from the central bureau of statistics of Norway.*

Howard, R. A. (1971). *Dynamic probabilistic systems: volume i: markov models.* John Wiley And Sons, Incorporated.

Iosifescu Manu, A. (1972). Non homogeneous semi-markov processes. *Stud. Lere. Mat*, 24:529–533.

Ismail, M. T. and Isa, Z. (2008). Identifying regime shifts in malaysian stock market returns. *International Research Journal of Finance and Economics*, 15:44–57.

Janssen, J. (1966). Application des processus semi-markoviens à un problème d'invaliditè. *Bulletin de l'Association Royale des Actuaries Belges.*

Janssen, J. and De Dominicis, R. (1984). Finite non-homogeneous semi-markov processes: Theoretical and computational aspects. *Insurance: Mathematics and Economics*, 3(3):157–165.

Janssen, J. and Manca, R. (1997). A realistic non-homogeneous stochastic pension fund model on scenario basis. *Scandinavian Actuarial Journal*, 1997(2):113–137.

Janssen, J. and Manca, R. (2002). General actuarial models in a semi-markov environment. *Proceedings of ICA Cancun*, 2002.

Janssen, J. and Manca, R. (2006). *Applied semi-Markov processes*. Springer Science & Business Media.

Janssen, J., Manca, R., and Di Biase, G. (1997). Markov and semi-markov option pricing models with arbitrage possibility. *Applied Stochastic Models and Data Analysis*, 13(2):103–113.

Kanas, A. (2005). Regime linkages between the mexican currency market and emerging equity markets. *Economic Modelling*, 22(1):109–125.

Kanas, A. and Kouretas, G. P. (2007). Regime dependence between the official and parallel foreign currency markets for us dollars in greece. *Journal of Macroeconomics*, 29(2):431–449.

Kijima, M. (2013). *Stochastic processes with applications to finance*. CRC Press.

Levy, P. (1954). Processus semi-markoviens. In *Proceedings of the International Congress of Mathematicians*, volume 3, pages 416–426.

Limnios, N. and Oprisan, G. (2003). An introduction to semi-markov processes with application to reliability. In *Stochastic Processes: Modelling*

and Simulation, volume 21 of *Handbook of Statistics*, pages 515 – 556. Elsevier.

Limnios, N. and Oprisan, G. (2012). *Semi-Markov processes and reliability*. Springer Science & Business Media.

Mandilaras, A. and Bird, G. (2010). A markov switching analysis of contagion in the ems. *Journal of International Money and Finance*, 29(6):1062–1075.

Moore, T. and Wang, P. (2007). Volatility in stock returns for new eu member states: Markov regime switching model. *International Review of Financial Analysis*, 16(3):282–292.

Nishiyama, K. (1998). Some evidence on regime shifts in international stock markets. *Managerial Finance*, 24(4):30–55.

Polansky, A. M. (2007). Detecting change-points in markov chains. *Computational Statistics & Data Analysis*, 51(12):6013–6026.

Sahin, I. and Balcer, Y. (1979). Stochastic models for pensionable service. *Operations Research*, 27(5):888–903.

Schaller, H. and Norden, S. V. (1997). Regime switching in stock market returns. *Applied Financial Economics*, 7(2):177–191.

Smith, W. L. (1953). Ii.—asymptotic renewal theorems. *Proceedings of the Royal Society of Edinburgh. Section A. Mathematical and Physical Sciences*, 64(01):9–48.

Stenberg, F., Manca, R., and Silvestrov, D. (2006a). Discrete time backward semi-markov reward processes and an application to disability insurance problems.

Stenberg, F., Manca, R., and Silvestrov, D. (2006b). Semi-markov reward models for disability insurance. *Theory of Stochastic Processes*, 12(3-4):239–254.

Stenberg, F., Manca, R., and Silvestrov, D. (2007). An algorithmic approach to discrete time non-homogeneous backward semi-markov reward processes with an application to disability insurance. *Methodology and Computing in Applied Probability*, 9(4):497–519.

Takács, L. (1954). Some investigations concerning recurrent stochastic processes of a certain type. *Magyar Tud. Akad. Mat. Kutató Int. Kozl*, 3:115–128.

Turner, C. M., Startz, R., and Nelson, C. R. (1989). A markov model of heteroskedasticity, risk, and learning in the stock market. *Journal of Financial Economics*, 25(1):3–22.

Walid, C., Chaker, A., Masood, O., and Fry, J. (2011). Stock market volatility and exchange rates in emerging countries: A markov-state switching approach. *Emerging Markets Review*, 12(3):272–292.

Wang, P. and Theobald, M. (2008). Regime-switching volatility of six east asian emerging markets. *Research in International Business and Finance*, 22(3):267–283.