Five Essays on Financial Economics: Econometric Methods and Credit Risk Valuation Techniques

Pedro José Serrano Jiménez

Departamento de Fundamentos del Análisis Económico II

Bilbao, 5 de noviembre de 2007

"Saluberrimi sunt sereni dies"

Auilo Cornelio Celso
(30 b.C.-50 a.C.)

Contents

List of Figures

List of Tables

Preface

This doctoral dissertation focuses on two important areas in Financial Economics that use to go in the same direction: the financial econometrics and the pricing of derivatives. We are interested in some questions concerning the adequacy of the existing econometric techniques to estimate the parameters of a certain model, and the model itself currently used for pricing complex derivatives instruments. We are inclined to believe that this study explores an alternative way to deal with this type of problems with the aim of providing several interesting questions and some helpful answers.

This dissertation consists in two main parts:

Part I: Econometric Methods. We study in detail a stochastic process recently proposed by Altmann *et al* (2007), describing its main statistical properties and suggesting and illustrating some estimation strategies that can be applied to this process.

Part II: Credit Risk Valuation Techniques. We also explore the current pricing techniques for a basket credit derivative named Collateral Debt Obligations, providing some additional insights.

This thesis is structured in five chapters grouped on these two blocks. Part I focuses on econometric methods and includes Chapters 1 to 3. This part studies the statistical properties and some estimation strategies of a stochastic process recently proposed by Altmann *et al* (2007). This new model is named the Shot-Noise Jump-Diffusion (SNJD, hereafter) process

and, basically, it introduces some new features into the family of Jump-Diffusion (JD) models initiated by Merton (1976).

Chapter 1 presents the SNJD model and some of the intuitions underlying this model. We also provide here several features of this model and some of the statistical tools that will be used on successive chapters. The chapter concludes by highlighting several of the economic implications of the SNJD model.

The second chapter in this part is devoted to the estimation of the SNJD model in the time domain. Basically, it comprises the estimation of this model by using the Generalized Method of Moments (GMM) technology. We review the GMM methodology and its properties, illustrating also the estimation of a JD process. The final part of this chapter develops an exhaustive Monte Carlo analysis of the estimation for the SNJD process.

Finally, this first part concludes with Chapter 3, which addresses the estimation of the SNJD model in the frequency domain, analyzing the adequacy of this methodology for the estimation of this type of process. We justify the econometric framework employed by providing some of the intuitions behind. Moreover, we perform an intensive Monte-Carlo study. This chapter finishes summarizing the main conclusions and suggesting some lines for further research.

Part II of this dissertation focuses on Credit Risk Valuation Techniques and comprises Chapters 4 and 5. One of the main goals of this part is to analyze the current pricing techniques of one of the fastest growing segments in the credit derivatives industry: the Collateral Debt Obligations, or CDO's. In this way, this part addresses in a simple and illustrative way some of the different alternatives that can be used to price and manage this type of derivatives.

Thus, Chapter 4 (the first one of this part) presents the standard model used in the industry for pricing CDOs. We generalize some of its assumptions and provide an empirical application with Moody's data.

This valuation block finishes with Chapter 5, which explores the CDO pricing model proposed recently in Longstaff and Rajan (2007). This chapter considers different multi-factor versions and, additionally, proposes some extensions considering jumps in the default process or a random loss distribution.

Acknowledgments

This thesis would not be possible without the help of my thesis advisor, Manuel Moreno. I am very grateful to him for his unconditional confidence in becoming involved in such a task. I have contracted a debt that I will not be able to pay back fully.

I am also very grateful to Professor Winfried Stute for providing many insights that had a major effect on my understanding and way of thinking. I am extremely thankful for having had the chance to work with him.

Next, I wish to thank to the Department of Fundamentos del Análisis Económico II (and, especially, to Jose María Usategui) for its instrumental and technical assistance in this research. They have provided me the opportunity of funding this thesis and several congresses and courses. I am strongly grateful to them for making possible this doctoral thesis.

Miguel Ángel Martínez also has an special place in this section. He has been the supporter in many difficult situations. Thanks, many thanks for being there.

I am extremely grateful to Alfonso Novales and Ángel León for their valuable comments and time, which has substantially improved this work. I also recognize comments from Juan Ignacio Peña, José María Vidal, and many of the attendants to several seminars and conferences. I also acknowledge to the Business Departments at University Carlos III, Madrid, Universitat Pompeu Fabra, Barcelone and University of Alicante for their hospitality during several visiting stays.

This thesis was improved by conversations with many of my colleagues

and friends of the Ph.D. Program in Banking and Quantitative Finance from current and previous editions. I am especially grateful to Agueda, Unai, Toni and Chema for sharing with them many questions, experiences and doubts.

Finalmente, quisiera agradecer el apoyo incondicional de mi familia: a mis abuelos, mis hermanos Mª Felicia, Emilio y Alfonso, y especialmente a mis padres, sin cuyo cariño y confianza nada de esto hubiera sido posible.

Jumilla, Noviembre de 2007

Introduction

This thesis dissertation deals, to a certain extent, with the interplay between financial models and the econometric tools used for their estimation. On one hand, we are interested in studying the economic features and implications of certain models ultimately designed for pricing purposes. On the other hand, we want to check the capability of some standard econometric tools for estimating the parameters of some posited models.

Naturally, this previous formulation leads to very different questions that should be treated separately. Attending to this feature, we have structured this dissertation in two parts that could be considered independently, taking into account the issues analyzed in each one.

The first part of this dissertation comprises several econometric issues and is structured in three chapters. This "econometric" part explores the availability of some standard econometric techniques for dealing with a model recently proposed by Altmann *et al* (2007) (named the Shot-Noise Jump-Diffusion (SNJD, hereafter). The SNJD process presents some new features with respect to the family of Jump-Diffusion (JD) models initially proposed by Merton (1976). Basically, the SNJD model tries to reflect that sudden, extreme changes in the evolution of asset prices may *fade away* on the long run. The intuition underlying this model is clear: If, for instance, a positive jump in a asset price happens, we expect a profit taking behave in financial markets, encouraging investors to sell their assets. In a similar way, an abnormal decrease in the asset price can encourage investors to buy such asset.

The importance of considering such behavior in asset prices could result in some interesting implications. For example, as it is mentioned by Ait-Sahalia (2004) for the case of JD process, decomposing the asset prices noise in two components (related to a continuous Brownian motion and a discontinuous jump, respectively) leads to different hedging strategies in a option pricing context. Moreover, the addition of a persistence component (the Shot noise part) related to the discrete term responsible of the jumps events could have unknown consequences on the option prices.

Another important issue concerns the capability of the SNJD model for generating serial persistence in asset returns that could lead to long-term memory patterns. This limitation of the current continuous-time models used in finance has been pointed out by Lo (1991), who refers to the impossibility of assessing, for example, the optimal decision sensitivity with respect to the investment horizon when the returns are long-range dependent.

From an econometric perspective, the estimation of the SNJD process involves some additional challenges. For example, we could worry about the necessary tools for its estimation (statistical distribution, moments, etc). Additionally, an important issue is to analyze its capability to fit market data better than some of the models previously proposed in the literature, Moreover, the availability of reliable estimation techniques of the SNJD process results on additional elements to test the presence of dependence patterns that can be present in financial data.

Chapter 1 focuses on the analysis of the SNJD model and its main purpose consists on providing the basic statistical tools for analyzing the SNJD process. We compute the characteristic function of the process as well as its main moments. We complete the study of this model computing its cross-moments and its spectrum. In addition to this, we also provide some numerical approximations and explore some of their economic implications.

Chapters 2 and 3 conduct the estimation in the time and frequency domain of the SNJD processes, respectively. The main issue addressed in these chapters concerns to the ability of the different proposed estimates to capture the parameters of a SNJD process.

Basically, Chapter 2 studies the performance of the GMM estimate proposed by Hansen (1982). We start checking the performance of the GMM methodology by using different samples not perturbed by Brownian noise, Secondly, we carry out an estimation of the JD processes with the purpose of being familiar with the GMM results for this type of processes. Finally, some results about the SNJD estimation are provided. In addition to this issue, the last section of this Chapter is devoted to analyze possible problems

of misspecification and overparametrization in the JD and SNJD models under SNJD and JD simulated samples, respectively.

Chapter 3 is devoted to the estimation of the SNJD process in the frequency domain. In more detail, this chapter presents two main tools for carrying out this task: firstly, the naive estimate, directly inspired on Bevan *et al* (1979), and an alternative technique, based on Whittle (1953). The first part of this chapter introduces the spectral approach and the intuitions underlying this method. We also discuss the estimates under study, presenting some of their features. To study the ability of the posited estimates to capture the SNJD parameters, we also perform an exhaustive Monte Carlo study.

The second part of this dissertation, named credit risk valuation techniques, includes two chapters and analyzes some of the current models for pricing basket credit derivatives. This part of the dissertation addresses in a simple and illustrative way some of the different alternatives proposed in the recent credit risk literature to deal with this type of derivatives.

The motivation for this part arises from the (large) differences between standard and credit derivatives. The idea is that, in comparison with standard derivatives, practitioners in credit derivatives markets are dealing with assets designed to satisfy different requirements and that imply hedging strategies when compared to standard derivatives.

In more detail, this part focuses on the pricing of credit risk derivatives. It should be clear that the pricing of derivatives depends on the process assumed for the underlying asset price. Then, it seems clear that the better the description for the underlying asset price, the more accurate the valuation of the derivative.

In the current successful world of credit derivatives, the classical concept of derivative still remains. Here, the underlying asset is usually a financial one (bond, loans, etc) and the holder of the credit derivative also has certain rights on this underlying. However, the main difference between standard derivatives and their credit analogues concerns the relevant issue for an investor. The idea is that a credit derivative investor does not worry on the asset price and her main interest is on the degree of creditworthiness. As a consequence, in the credit derivatives world, our interest moves to the modeling of the losses distribution. Of course, this leads us to new pricing approaches to price these new credit derivatives families.

In more detail, this part explores the valuation techniques for Collateral Debt Obligations, or CDOs, a basket credit derivative which has received an increasing attention by current research. Basically, we are interested

on the different alternatives posited in the literature to capture the loss distribution of a portfolio of credit names (loans, bonds, etc.). We approach to this issue by exploring the main current credit pricing methodologies available, grouped into the structural and intensity categories.

Jointly with the motivation aforementioned, as second motivation for this part, we can consider that, in recent years, the appearance of standardized products as the CDX and Itraxx credit indexes has contributed to the success of CDOs trading. Currently, the market for tranched credit products (CDOs, Itraxx tranches) is one of the fastest growing segments in the credit derivatives industry and its numbers may justify our study.

As we have told before, this part includes two chapters, 4 and 5. Chapter 4 addresses a simple question: within the structural framework, is it possible to relax some implications of the Vasicek (1991) (Gaussian one-factor) model, the pricing standard widely used in the industry? This chapter points out how some assumptions underlying this model (homogeneity, single factor, Normality) are probably too restrictive. Then, the aim of this chapter is to generalize this standard model by proposing a two by two model (two factors and two asset classes). We assume two driving factors (business cycle and industry) with independent t-Student distributions, respectively, and we allow the model to distinguish among portfolio assets classes. To illustrate the estimation of the model parameters, an empirical application with Moody's data is also included and potential relationships between default rates and macroeconomic variables are also analyzed.

Finally, Chapter 5 analyzes (from an intensity based approach) the CDO pricing model proposed recently in Longstaff and Rajan (2007) considering different multi-factor versions. This chapter wonders about the robustness of the results in this model replacing the posited intensity processes by more flexible extensions as a) considering jumps in the default process or b) a random loss distribution. In addition to this, we also discuss the usefulness of considering three factors in CDOs indexes pricing. Thus, a Monte Carlo simulation provides evidence that a three-factor version of this model with constant losses impact is flexible enough to reproduce the spreads given by the market. Finally, it is also shown that an alternative considering random losses can be helpful when dealing with one- and two-factor models.

To sum up, the main contributions of this dissertation can be summarized as follows:

- With regard to the first part, we analyze the SNJD model from a theoretical point of view, providing an analytical expression for the

distribution of the process and several of its statistical properties.

- To our knowledge, this is the first study on analyzing the economic implications of the SNJD model. This work is also pioneer on providing many of the statistical tools of the model necessary for its estimation.

- Additionally, from an empirical point of view, we propose different econometric frameworks for dealing with the SNJD process by exploring some standard estimation techniques in time and frequency domain.

- Our results seem to suggest that the GMM methodology is capable of estimating the parameters of a SNJD model. However, an observed bias is detected in the parameters associated with the total variance of the process. Apparently, the inclusion of more efficient weighting matrices in the estimations does not improve substantially the quality of the results obtained.

- We explore an alternative econometric methodology based on the spectral analysis to estimate the SNJD model parameters. Two basic spectral estimates are posited, the Naive and Whittle ones. In spite of their simplicity, our results reveals that the naive estimate could be an adequate tool for the estimation of the SNJD processes, under the sample studied. With regard to the Whittle estimate, some numerical problems concerning to the estimation procedure were encountered. However, many of its properties -consistence, asymptotical efficiency- leads to take into consideration this estimate in larger sample studies.

- On the whole, the frequency domain estimates are less biased than the time domain ones. With regard to their dispersion, this aspect varies across estimates and parameters. Although the comparison between time and frequency domain estimates of this part should not be taken as definite, we think that the obtained results support the idea that the spectral estimation should be considered as an additional tool for estimating continuous-time serial dependent processes. In line with this, our results seem to confirm that this procedure could complement other time domain alternatives as the GMM ones.

- In general, the numerical approximations involved in the different moment expressions appear to affect dramatically the results.

- With respect to the second part, we explore some of the current pricing techniques for a special type of credit basket derivative: the CDOs.

- We provide more flexible extensions to the Gaussian model proposed in Vasicek (1991) by considering a family of models that takes into account the existence of different asset classes or regions of correlation. We assume two driving factors (business cycle and industry) with independent t-Student distributions, respectively, and allow the model to distinguish between portfolio assets classes.

- We analyze the two asset classes models, extended their scope to the t-Student distribution. We also provide the econometric framework for assessing the parameters of the posited model. An empirical application with Moody's data has been also presented as an illustration of the methodology proposed. To the best of our knowledge, no similar study has been reported yet in this direction.

- In reference to Chapter 5, this is the first work on discussing and extending the recent proposal of Longstaff and Rajan (2007) to a more realistic intensity process set up, to the best of our knowledge.

- Our results seem to suggest that a three-factor version of Longstaff and Rajan (2007) with constant losses impact is flexible enough to reproduce the spreads given by the market. In addition to this, the inclusion of jumps to the default process results in an high arrival of credit events, as corroborated by the high values spreads for equity tranches.

- The alternative of random losses in the Longstaff and Rajan (2007) model can be helpful when dealing with one- and two-factor models, but it seems to be irrelevant in the case of three-factor models. Anyway, this last point can be developed as a subject for further research.

Part I

Econometric Methods

Chapter 1

The SNJD process

This chapter[1] analyzes the Shot Noise Jump-Diffusion model (SNJD, hereafter) posited by Altmann et *al* (2007). We study the main features of the classical Shot Noise process, the origin of the SNJD model. From a theoretical point of view, we provide an analytical expression for the distribution of the process and several of its statistical properties. Finally, some of the economic implications of the model are also analyzed.

1.1 Introduction

The irruption of diffusion processes has supposed a major breakout in the recent financial theory. From the seminal work of Black and Scholes (1973) up to our days[2], the continuous time techniques have revealed as a pow-

[1]I am sincerely grateful to W. Stute and M. Moreno for their contributions on this chapter. I would like also to thank to D. Nualart, A. Novales and A. León for his helpful comments. Finally, I also acknowledge financial support from the Plan Nacional de I+D+I (project BEC2003-02084) and especially to José M. Usategui. Previous drafts of this section have been presented in several congress as EuroWorking Group on Financial Modelling, 8th Italian-Spanish Meeting on Financial Mathematics, 9th Spanish-Italian Meeting on Financial Mathematics and XIV Foro de Finanzas. I have also benefited from the comments of participants in seminars at University of Basque Country, Universidad Complutense de Madrid, Universitat de les Illes Balears and Universidad Carlos III de Madrid.

[2]Black and Scholes (1973) is considered a real milestone in the modern Finance in continuous time. However, it was not the first in the usage of stochastic processes for modeling the prices behavior: in this context, the articles of Bachelier (1900), which introduces the Brownian motion, and Samuelson (1965), which uses the Geometric Brownian Motion for capturing the distribution of stock prices, are commonly cited as the basis of stochastic processes applied in Finance.

erful tool in a wide range of areas in finance: the derivatives pricing, risk management or asset allocation, among others, have been benefited from the tractability of diffusion processes technology and the reduced number of assumptions involved in.

Basically, a diffusion process comprises the evolution through time of a random variable constrained to a statistical distribution (Parzen, 1962). To put it another way, there is always a statistical distribution behind a diffusion process. The trade-off between the adequacy of the distribution to capture the behavior of the economic variable and its analytical tractability continues to be an important subject of active research on the current financial economics area.

Nowadays, the family of the stochastic processes named Affine Jump-Diffusion processes (AJD) seem to present a useful solution to this compromise. As it was pointed out by Duffie *et al* (2000), the AJD models provide analytical solutions to many problems in arbitrage-free pricing theory; moreover, the existence of a joint econometric theory for their estimation[3] leads to the current situation where the wide range use of ADJ process in financial economics is undoubting. To some extent degree, it would not be erroneous to consider that the AJD family comprises the most popular continuous time models in Finance[4]: for example, the Vasicek (1977) or the Cox-Ingersoll-Roll (1985) models for interest rates constitute some examples of the usage of the continuous time based engineering in capturing the behavior of economic variables.

In this context, the extension to new families of models is clearly needed when the original model is not able to reflect some empirical features usually presented in economic series. For instance, the case of the Geometric Brownian Motion (GBM) is often presented as paradigmatic: the most popular option pricing model introduced in Black and Scholes (1973) assumes that the price at a certain time of an asset follows a GBM. The assumption behind this process is that the underlying asset returns (prices) are normally (lognormally) distributed; however, the excess of kurtosis or the volatility clusters detected in financial prices series (see, for instance, Andersen *et al*, 2002) lead to misspricing errors in the options pricing (Navas, 2003).

Additional examples of drawbacks on current AJD literature could be

[3]See Singleton (2001) or Duffie *et al* (2000).

[4]Recent financial literature has also focused on the Lévy processes, which exhibits an infinite number of small jumps in any finite interval (Ait-Sahalia, 2004). Some useful references about this topic can be found on Chan (1999), Bandorff-Nielsen *et al* (2002) and references therein. The textbook of Schoutens (2003) also provides a very extensive introduction to this subject.

also reviewed: with reference to the interest rate models, Duffee (2002) points out the inner restrictions of standard affine models when it comes to compensate the risk faced by investors[5], based on the poor performance of AJD models on Treasury yields. On the other hand, Lo (1991) and references therein stress that the class of continuous time stochastic processes most commonly employed is inconsistent with the persistence patterns found on asset returns (Campbell *et al*, 1997), in view of the fact that current continuous-time models used in finance are not capable of generate serial dependence consistent with long-term memory in returns.

All things considered, is it still possible to learn something more about the empirical features with the existing continuous time models[6]? Up to a point, it seems clear that some questions remain open to the appearance of new models but, is it sufficiently justified the cost of adopting these models with the knowledge acquired? Needless to say, it is desirable that these new families would be based, as possible, on the past literature, as well as they should keep features as the variables parsimony and economic intuition.

This chapter is devoted to introduce a new family of models that fulfill this trade-off between economic intuition and analytical tractability. Inserted on the family of the Jump-Diffusion processes (JD, hereafter), we analyze the main features of the model recently posited by Altmann *et al* (2007): the SNJD process. This model is similarly based on traditional JD processes, where the total variation in the asset price is due to the joint effect of two types of changes: normal changes, due to the bid-ask crossing of common information, and abnormal ones, where the impact of the information on the price produces a non-marginal change in the asset price (see Merton (1976) for details). Basically, the SNJD process introduced by Altmann *et al* (2007) extends the model introduced in Merton (1976) by adding a term (called shot-noise function) to the Poisson process. As we will see, this additional term tries to reflect some situations that can arise in stock markets as, for instance, when the effects of a jump fade away on the long run[7].

This chapter contributes to the existing literature on diffusion process in several directions: from a theoretical point of view, we introduce a brief treatment of the Shot Noise in the scientific literature. We provide an analytical expression for the distribution of the process and several of its sta-

[5] An interesting discussion about this feature can be found in Singleton (2006).

[6] Paraphrase from Duffee (2002).

[7] An additional advantage of this type of models is that, as discussed in Bondesson (1988), many statistical distributions appear as marginal distributions for very simple shot-noise processes.

tistical properties, analyzing the implications of the model in both time and frequency domains. Finally, a general discussion of the economic significance of the model is also included. To our knowledge, this is the first study that analyzes in detail the applications of the Shot Noise process in Finance.

The structure of the chapter is as follows: firstly, we present a general treatment of the classical Shot Noise. Secondly, the SNJD is also presented. Section 3 and Section 4 provide the distribution of the processes, and its cross moments and population spectrum, respectively. We analyze on Section 5 the most basic example of SNJD process. Finally, Section 6 discusses the main economic implications of the SNJD model, and some conclusions are given on Section 7.

1.2 The model

We start introducing the idea of the classical Shot Noise process and, later, we present the SNJD process and some main features. Finally, we compare the SNJD process with the GBM and JD processes, analyzing the differences among these processes.

1.2.1 The Shot Noise process

Let H_t be an stochastic process of the form

$$H_t = \sum_{k=1}^{N_t} A_k h\left(t - \tau_k\right), \quad t \geq 0 \tag{1.1.2.1}$$

where $\{A_k,\ k = 1, \cdots, N_t\}$ are i.i.d. random variables, $\{N_t\}_{t\geq 0}$ is an homogeneous Poisson process with intensity λ, $h\left(t - \tau_k\right)$ represents the reaction to a possible event with magnitude A_k, and $\{\tau_k\}$ indicate the different moments in which the Poisson process appears. It is also assumed that the variables A_k are independent of the Poisson process, N_t.

Then, H_t is the classical *shot-noise process*. This type of process has been previously proposed in different situations as, for example, in modelling arrivals of electrons in a vacuum tube (see Rice, 1954) or earthquake aftershocks (see Daley and Vere-Jones, 1988). Figure 1.1 represents a simple path of a Shot Noise process.

An interesting feature of a shot-noise process is that many distributions appear as marginal ones for this kind of processes (see Bondesson, 1988). This is the reason why the function $h(\cdot)$ plays an important role in the

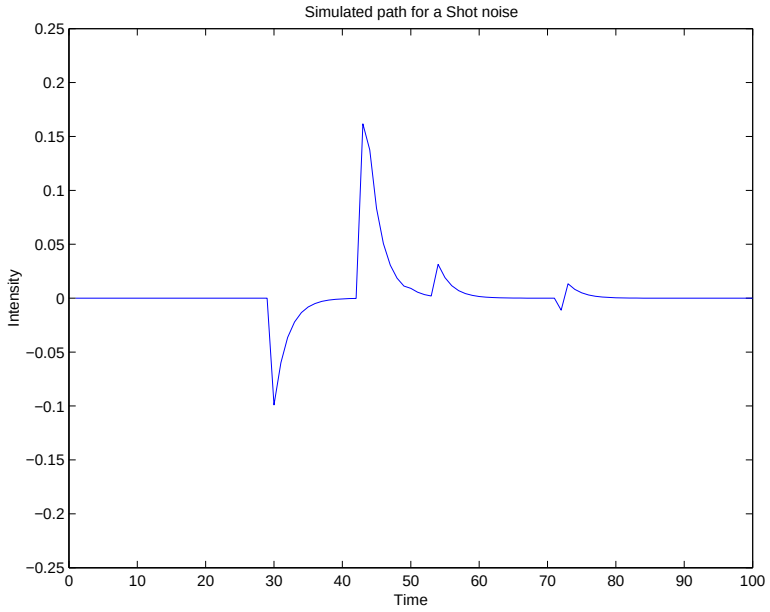


Figure 1.1: Representation of a path simulation for the Shot Noise process of expression (1.1.2.1). The response function is $h(t-\tau_k) = e^{-a(t-\tau_k)}$. Random variable A_k is lognormally distributed as $A = e^{-\beta^2/2+\beta\varepsilon} - 1$, with $\varepsilon \sim N(0,1)$ and mean $E[A_k] = 0$. Simulation parameters are $\lambda = 10.00$, $\beta = 0.10$ and $a = 0.50$.

process distribution. As an example, if the distribution of the variable $\{A_k\}$ is a mixture of exponential distributions and

$$h(x) = \begin{cases} 0 & x < 0 \\ ce^{-ax} & x \geq 0 \end{cases}$$

Bondesson (1988) shows that the shot-noise distribution is a generalized gamma convolution.

An additional characteristic of the Shot Noise process is that it is capable of introducing autocorrelation in data. Table 1.1 displays the autocorrelation values (ACF) for the classical Shot noise process of expression (1.1.2.1) with exponential decaying response function[8]. The parameter a controls the degree of decay of the process. Each column correspond to different values of this parameter. In order to compare with a known process, the last column includes the ACF for an AR(1) model. Different lags are disposed in

[8] Advancing some results that will be presented later, the ACF of this process is given by $Corr[H_t, H_{t+n\Delta t}] = e^{-a\, n\Delta t}$, and can be found in Ross (1996). More general expressions for the ACF of the Shot noise process are provided later in this chapter.

Table 1.1: Autocorrelation function for different values of the parameter a.

Lag (n)	$a = 0.2$	$a = 0.6$	$a = 1.5$	$AR(1), \phi = 0.6$
1	0.8187	0.5488	0.2231	0.6000
2	0.6703	0.3012	0.0498	0.3600
3	0.5488	0.1653	0.0111	0.2160
4	0.4493	0.0907	0.0025	0.1296
5	0.3679	0.0498	0.0006	0.0778
10	0.1353	0.0025	3.06×10^{-7}	0.0060
25	0.0067	3.06×10^{-7}	5.17×10^{-17}	2.84×10^{-6}
50	4.54×10^{-5}	9.35×10^{-14}	0.0000	8.08×10^{-12}
100	2.06×10^{-9}	8.75×10^{-27}	0.0000	6.53×10^{-23}

This table includes the theoretical autocorrelation coefficients $\rho(n)$ for different lags and values of the parameter a involved in the response function $h(t - \tau_k) = e^{-a(t-\tau_k)}$ included in eq. (1.1.2.1). Random variable A_k is lognormally distributed as $A = e^{-\beta^2/2+\beta\varepsilon} - 1$, with $\varepsilon \sim N(0, 1)$ and mean $E[A_k] = 0$. Simulation parameters are $\lambda = 0.04$, $\beta = 0.10$ and $a = 0.50$.

rows. This table is directly inspired on Campbell *et al* (1997), where we have employed the Shot Noise model instead of the fractionally differenced process.

Table 1.1 reflects that the parameter a modulates the persistence of the process in time: the more we increase parameter a, the more decay is the ACF at first lags. To put it another way, as a increases, we observe that serial correlation in time diminishes, as it is observed by direct comparison with the $AR(1)$ process column. On the whole, we can consider that the same pattern of decay is observed between the $AR(1)$ process column and the classical Shot noise.

To provide additional arguments to the serial correlation introduced by the model (1.1.2.1), Figure 1.2 exhibits the nonparametric estimate of the spectrum for a Shot noise obtained from a time series created from the values of 100 simulated paths composed by 1800 data each one. Our final time series is composed by the mean values of the steps for each path, a procedure which ensures certain degree of robustness to the results. Simulation parameters

are displayed in the figure.

With the intention of keeping intuition of the results, let the spectrum be considered as a tool for capturing the dynamics of the time series, since a more detailed definition will be given in Chapter 3. Roughly speaking, the area under the spectrum represents the total variance of the process[9]; lower frequencies represent the data linked by yearly, quarterly periods; by contrast, higher frequencies are related to information contained on weekly, daily data. If the shape of the spectrum is higher for certain frequencies than others, it reflects that a bigger portion of the variance of the process could be explained by what has happened in that time spans (frequencies). To put it another way, the information contained in some cyclical components of the series has an important effect in explaining the total behavior of the process.

As an example of the former, consider a flat spectrum. This means that there is no special contribution to the variance of the process in any particular frequency. For instance, a white noise process has a constant spectrum[10]. By contrast, as we can see in Figure 1.2, our time series exhibits power at low frequencies. This must be intended as follows: an important part of the variability of the process is explained by larger time periods - e.g. yearly - data. Taking everything into account, the intuition behind is clear: the addition of a Shot Noise to a stochastic process could lead a source of serial dependence.

Shot noise processes have been intensively studied by Rice (1954) or Parzen (1962)[11]. Moreover, Bondesson (1988) provides a general treatment of these processes and includes extensive references.

1.2.2 A model with shot-noise effects

According to Merton (1976), the arrival of abnormal information (for example, news about a company default) produces a non-marginal change on the price of a certain asset. Merton models this component using a compound Poisson process that reflects the impact of this non-marginal information. Within the current financial literature, these abrupt changes in prices lead to the so-named *jumps*. As a result of this, Merton's model describes more realistically the behavior of stock prices.

An additional implication of Merton's model is that changes in stock

[9] For a general description of the spectral analysis, see Hamilton (1994).

[10] Indeed, the spectrum of log-returns of a GBM or a JD process is constant, too.

[11] The shot noise process is sometimes called Filtered Poisson process, *although this terminology is not standard* (Parzen, 1962).

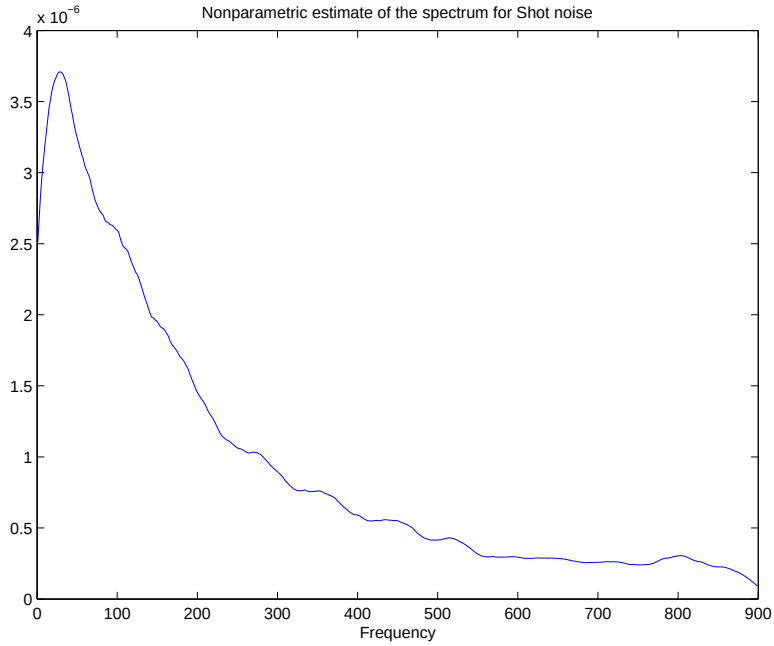


Figure 1.2: Nonparametric estimate of the population spectrum for a classical Shot noise of equation (1.1.2.1). The response function is $h(t-\tau_k) = e^{-a(t-\tau_k)}$. The random variable A_k is lognormally distributed as $A = e^{-\beta^2/2+\beta\varepsilon} - 1$, with $\varepsilon \sim N(0,1)$ and mean $E[A_k] = 0$. Simulation parameters are $\lambda = 0.04$, $\beta = 0.10$ and $a = 0.50$. Bandwidth parameter of the estimate has been fixed to 40.

prices have a long-term effect[12], that is, the effect of a change in stock prices does not vanish with time. After a upward jump, prices following a jump diffusion process (JD) like in Merton (1976) do not return to the initial (pre-jump) level. Figure 1.3 illustrates graphically this situation: two processes generated with the same seed, GBM and JD, follow equal trajectories before a jump event but, after the jump, both processes follow very different paths.

By contrast, it could be thought that these changes *fade away* on the long run. An intuitive explanation for this decay is that a positive jump is followed by a profit taking, encouraging the investors to sell their assets. Similarly, an abnormal decrease in the asset price can encourage investors to buy such asset. In terms of the stochastic literature, these impulses can produce a *shot noise* effect.

Altmann *et al* (2007) presents a model that captures some of these effects

[12]This must not be confused with long-term memory in second moments.

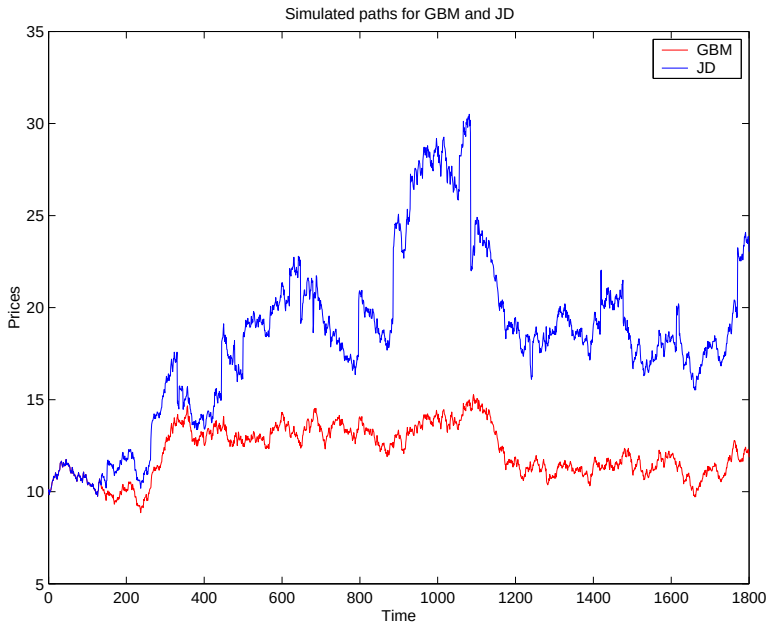


Figure 1.3: Representation of path simulations for Geometric Brownian Motion (GBM) and Jump Diffusion (JD) models. Random variable U for JD process is lognormally distributed as $U = e^{-\beta^2/2+\beta\varepsilon} - 1$, with $\varepsilon \sim N(0,1)$ and mean $E[U] = 0$. Simulation parameters (annualized) are $\mu = 0.05$, $\sigma = 0.20$, $\lambda = 5.00$ and $\beta = 0.10$

and, then, can be an adequate candidate to provide a better fit to market data than some of the models previously proposed in the literature. This model is named Shot-noise jump-diffusion (SNJD, hereafter) and assumes that the process for the stock price is given by the following expression

$$S_t = S_0 \exp\left[\left(\mu - \frac{\sigma^2}{2}\right) t + \sigma B_t\right] \prod_{j=1}^{N_t} \left[1 + U_j h\left(t - \tau_j\right)\right] \qquad (1.1.2.2)$$

where

- S_j denotes the asset price at time j
- μ, σ are constants
- $\{B_t\}_t$ is a standard Brownian motion
- $\{N_t\}_{t\geq 0}$ is a Poisson process with intensity λ

- $\{U_j\}_j$ is a sequence of i.i.d. jumps
- h is an arbitrary response function (to be determined)
- τ_j is the instant in which the j-th jump appears

Additionally, it is assumed that $\{B_t\}_t, \{N_t\}_t$, and $\{U_j\}_j$ are mutually independent.

For illustrative purposes, one example of the decay function $h(t-\tau_j)$ can be the following:

$$h(t) = \begin{cases} 0, & t < 0 \\ \exp(-at), & t \geq 0 \end{cases} \tag{1.1.2.3}$$

where $a \in \mathbb{R}^+$ is the velocity of the jump effect decay.

Figure 1.4 displays the simulation of the SNJD process in equation (1.1.2.2) with exponential response function of the form (1.1.2.3). As it will be comment later, this specification of response function is capable of generate spikes, a common feature in interest rates (Benito *et al*, 2006) or electrical price series (Escribano *et al*, 2002).

Another advantage of the SNJD model is its capability to introduce serial correlation in returns by means of the Shot noise process, as it was previously noticed in subsection 2.1. The degree of serial correlation can be modulated, for example, by modifying the parameter a in equation (1.1.2.3). In this case, high values of a (fast decaying effect) increase the level of autocorrelation for the first lag. Finally, with $a = 0$ we obtain the Merton (1976) process as a nested version of the model.

1.2.3 Main features

We provide here some results that serve to illustrate some of the main features of the model SNJD (see 1.1.2.2). Adopting a position that will be held through this dissertation, we simulate a SNJD model with exponential response function as in (1.1.2.3). For simplicity and without loss of generality, all the simulations are developed for the model in (1.1.2.2) with jump sizes $U = e^{-\beta^2/2+\beta\epsilon} - 1, \varepsilon \sim N(0,1)$. The assumption of a lognormal distribution for the jump sizes is common in the jump-diffusion literature (see, for instance, Merton, 1976)[13]. To decrease the number of parameters

[13]Some studies propose exponential distribution for the jump size. See, for instance, Duffie and Garlenau (2001) in the context of credit risk markets, or Barone-Adesi and Giggli (2002) and Villaplana (2003) for electricity markets, among others.

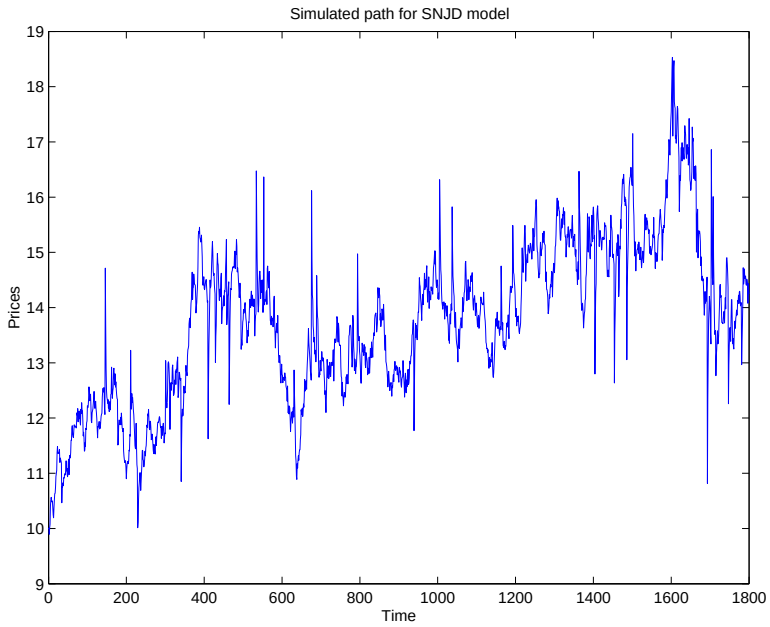


Figure 1.4: Representation of path simulations for a Shot-Noise Jump Diffusion (Altmann *et al*, 2004) process. Random variable U is lognormally distributed as $U = e^{-\beta^2/2+\beta\varepsilon} - 1$, with $\varepsilon \sim N(0,1)$ and mean $E[U] = 0$. Simulation (annualized) parameters are $\mu = 0.05$, $\sigma = 0.20$, $\lambda = 10.00$, $\beta = 0.05$ and $a = 0.05$

under estimation, we impose $E[U] = 0$, a common assumption also[14] (Ait-Sahalia, 2004). Finally, to our knowledge, the functional form of $h(t)$ has not been reported yet in the financial literature. Then, for analytical and intuition purposes, we assume an exponential decreasing function of the form $h(t) = e^{-at}$.

The main advantage of having this specification is that we can relate, in a simple form, the results obtained for our SNJD model with the two standard models for stock prices: firstly, the Geometric Brownian Motion (GBM), posited by Black and Scholes (1973) for describing the movement of stock prices; secondly, the Jump Diffusion process (JD), introduced by Merton (1976) as a more realistic description of stock prices behavior. Both models are nested by the SNJD process[15].

[14] Navas (2003) and references therein offers an interesting discussion about this assumption in the JD process literature.

[15] According to the proposed response function, if the parameter a is zero, the SNJD model leads to the JD model of Merton (1976). On the other hand, the GBM is obtained

Then, the parameter vector results $\Lambda = (\mu, \sigma, \lambda, \beta, a)' \in R^5$, where μ and σ are the drift and volatility parameters of the Geometric Brownian motion, λ is the intensity of the Poisson process, β is related to the mean and variance of the jump size, and the parameter a refers to the speed of the decaying effect after a jump event.

To show the differences between these models, we simulate the SNJD model against GBM and JD processes, separately. Basically, we compare the paths, distribution of log-returns and autocorrelation function (ACF) of log-returns, absolute log-returns and squared log-returns generated by these three models.

Some previous intuitions

To keep some intuitions in mind, Figure 1.5 exhibits the sample paths of the SNJD process against GBM and JD processes, respectively. Notice that all simulations have been carried out with the same seed, so Figure 1.5 shows differences among paths introduced by the jumps (and the Shot noise). The jump magnitude β is a little bit bigger than empirical, but it has been fixed to 0.10 to clarify the effect of the jump. Moreover, to see the persistence of the Shot noise on prices, the SNJD process has been represented when the parameter a equals to 0.5 and 1.5.

Upper graph in Figure 1.5 represents the SNJD trajectories against that of GBM. As it is possible to see, the three simulated processes share the trajectories: it is due to the the effect introduced by the Shot noise, where a change in stock prices fades away with time. By contrast, the lower graph in Figure 1.5 exhibits the SNJD paths against that of JD. As Figure 1.5 reflects, after a upward/downward jump, prices following a JD process do not return to the initial (pre-jump) level.

An inspection of the histogram of log-returns in Figure 1.6 reveals that, as expected, the SNJD model does not perfectly fit to a normal distribution. As Figure 1.6 exhibits, the SNJD model distribution has fatter tails than the normal one.

Autocorrelation function coefficients

Figure 1.7 shows the autocorrelation functions for the different processes under study. The simulation exercise comprises 100 different paths of 1800 data each one. The mean sample autocorrelation values for each trajectory have been represented in graphs, where the axis X captures the different

in absence of jumps.

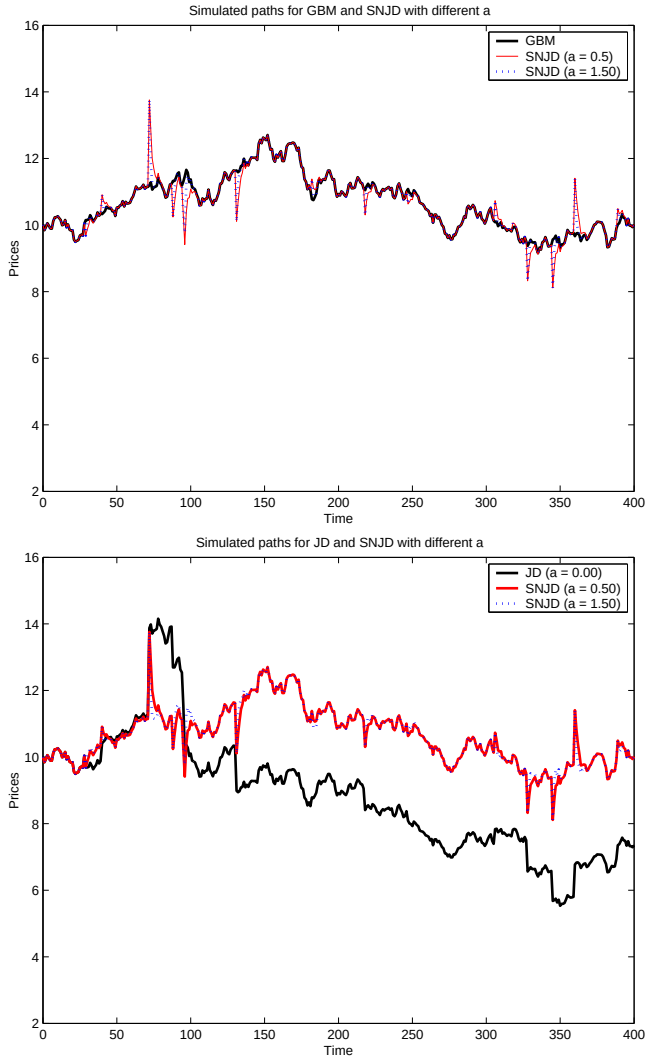


Figure 1.5: Representation of path simulations for a GBM and JD processes against different specifications of the SNJD process. Simulations parameters (annualized) are $\mu = 0.05$, $\sigma = 0.20$, $\lambda = 10.00$ and $\beta = 0.10$. Parameters a are 0.5 and 1.5, respectively. Response function for SNJD is $h(t - \tau_k) = e^{-a(t-\tau_k)}$ and the random variable U is lognormally distributed as $U = e^{-\beta^2/2+\beta\varepsilon} - 1$, with $\varepsilon \sim N(0, 1)$ and mean $E[U] = 0$.

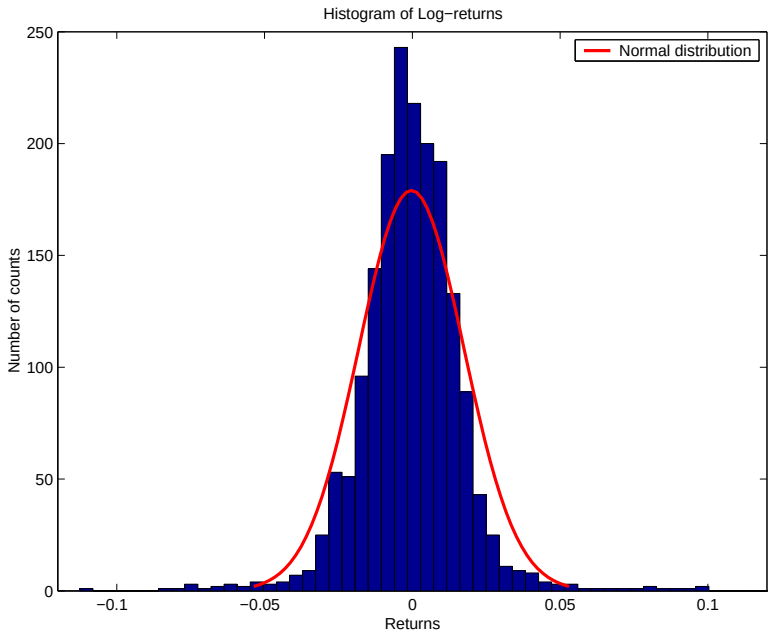


Figure 1.6: Histogram for a SNJD process. Response function for SNJD model is $h(t-\tau_k) = e^{-a(t-\tau_k)}$ and the random variable U is lognormally distributed as $U = e^{-\beta^2/2+\beta\varepsilon} - 1$, with $\varepsilon \sim N(0,1)$ and mean $E[U] = 0$. Simulation (annualized) parameters are $\mu = 0.05$, $\sigma = 0.20$, $\lambda = 10.00$, $\beta = 0.05$ and $a = 0.5$.

lags and axis Y represents their magnitude. In this figure, the first, second and third columns correspond to the autocorrelation coefficients for standard, absolute and squared log-returns, respectively. GBM, JD and SNJD processes are displayed in rows. The response function used here for the SNJD process is $h(t-\tau_k) = e^{-a(t-\tau_k)}$ and the random variable U is lognormally distributed with mean $E[U] = 0$. Finally, the simulation (annualized) parameters are $\mu = 0.05$, $\sigma = 0.20$, $\lambda = 10.00$, $\beta = 0.05$ and $a = 0.5$.

As we can see, the main feature of the SNJD model with exponential decay response function is the capability to generate a large persistence in log-returns, in contrast with GBM and JD models. Needless to say, different specifications of response function would lead to different patterns in the serial correlation behavior. Coming back to the exponential SNJD case, this process generates negative autocorrelation coefficient values, which decay exponentially with respect to the number of lags, as it will be presented later.

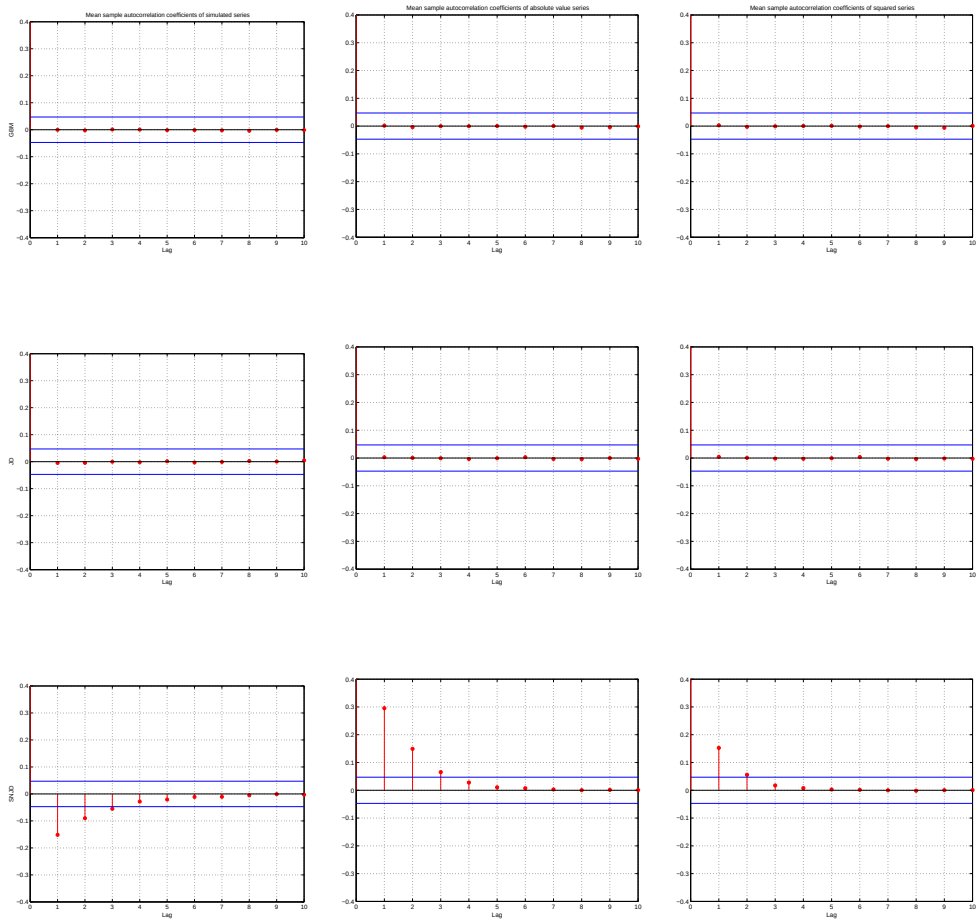


Figure 1.7: Autocorrelation functions (ACF) of log-returns, absolute log-returns and square log-returns for the GBM, JD and SNJD processes. First column correspond to log-returns. Second and third represents autocorrelations for absolute and squared log-returns. Each row contains ACFs for GBM, JD and SNJD processes, respectively. Simulation (annualized) parameters are $\mu = 0.05$, $\sigma = 0.20$, $\lambda = 10.00$, $\beta = 0.05$ and $a = 0.5$. Response function for SNJD is $h(t - \tau_k) = e^{-a(t-\tau_k)}$ and the random variable U is lognormally distributed with mean $E[U] = 0$.

1.3 The distribution of the process

This section focuses on computing the characteristic function of the log-returns of the SNJD process, as a necessary previous step to perform the empirical analysis of the model with shot noise. We also provide expressions for the first moments, and compare our results with those obtained for the JD processes in Das and Sundaram (1999) or Ait-Sahalia (2004).

1.3.1 Characteristic function

Let $(\Omega, \digamma_t, P)$ be a probability space and consider the filtration $\digamma_t$. Let $\{t \in [0,T]\}$ be the set of trading dates. Δt (> 0) denotes the length of time between two price observations. We compute the log-return of the stock between the times t and $t + \Delta t$ as $Z_t = \ln(S_{t+\Delta t}/S_t)$. Using (1.1.2.2), Z_t is given by the sum of three independent random variables, X_t, Y_t and H_t where

$$X_t = \left(\mu - \frac{\sigma^2}{2}\right)\Delta t + \sigma\left(B_{t+\Delta t} - B_t\right) \tag{1.1.3.1}$$

$$Y_t = \sum_{j=N_t+1}^{N_{t+\Delta t}} \ln\left[1 + U_j h\left(t + \Delta t - \tau_j\right)\right] \tag{1.1.3.2}$$

$$H_t = \sum_{j=1}^{N_t} \ln\left[\frac{1 + U_j h\left(t + \Delta t - \tau_j\right)}{1 + U_j h\left(t - \tau_j\right)}\right] \tag{1.1.3.3}$$

It is worth to say some words for capturing the intuition behind the former equations. The term X_t in expression (1.1.3.1) is obtained when we compute the log-returns of a GBM process. Basically, it expresses the contribution to the total log-returns of the SNJD model in (1.1.2.2) by part of the GBM term. As we see, this contribution is *continuous* in time[16] and, due to the independence property among Brownian increments, log-returns generated by this term are *not* serially correlated[17].

The term Y_t in expression (1.1.3.2) reflects the *punctual* contribution of a jump event on total log-returns. Remember that N_t is a *counter* process that takes values $0, 1, 2, ...$ and so on. Moreover, the probability of that

[16] Because of the paths of a Brownian motion are continuous through time.

[17] The mathematical concepts of continuity and independence are usually connected with those of completeness and efficiency of the markets on Economics (see, for instance, Cambpell *et al* (1997) or Delbaen and Schachermayer (1997) and references therein, among others). Our intention here is just to notice this point in the equations, but we will come back to this question more ahead.

one jump happens in between times t and $t + \Delta t$ is $P[N_{t+\Delta t} - N_t = 1] \approx \lambda \Delta t$, with λ the frequency that a jump occurs[18]. As equation (1.1.3.2) expresses, this summand term just exists when a new event arises, and it is zero otherwise. For the above mentioned reasons, it is clear that the contributions of the term Y_t to the total log-returns of the SNJD process are spread through time, that is to say, they are *discontinuous* inputs of the total log-returns. Moreover, this realizations are not connected[19], by independence of increments of the Poisson process.

If we compute the log-returns of a JD process, like Merton (1976), this just leads to the expressions obtained for X_t and Y_t. Not surprisingly, an additional term appears when it comes to calculate the log-returns of a SNJD process: the H_t term on equation (1.1.3.3). As we can see, H_t represents the contribution of a jump event with probability $\lambda \Delta t$ *through* time; to put it another way, equation (1.1.3.3) captures the addition of the value of the different shocks that occur in accordance to a Poisson process with rate λ, assuming that they decrease over time at a deterministic rate (Ross, 1996). Looking at equation (1.1.3.3), it is clear that this expression retains all the information about the past history of the process. As it will be pointed out in the next section, this term is the source of the autocovariance of the SNJD process. Finally, notice that observations are connected through time; moreover, regardless of the fact that the contribution of a jump shock is punctual, their effects are *transitory*.

The probability law of any stochastic process V_t could be expressed in terms of its characteristic function, named $\Phi_V(\xi)$ (Parzen, 1962). The following Lemma summarizes this result for the process Z_t.

Lemma 1 *The characteristic function of the process Z_t is given by*

$$\Phi_{Z_t}(\xi) = \exp\left(i\left(\mu - \frac{\sigma^2}{2}\right)\Delta t\xi - \frac{1}{2}\sigma^2 \Delta t \xi^2\right)$$

[18] The actual rate of jumps per year differs across markets (see, for instance, Jorion (1988) or Das, 2002). For indexes of stock prices, this rate is low (Ait-Sahalia (2004) assumes a rate of five jumps by year on average, *much higher than would be realistic*). This is the reason why the probability that more than one jump happens in a short period of time will be considered almost insignificant.

[19] More precisely, the observations are not *almost* connected. We clarify this point in the next paragraph.

$$\times \exp\left(\lambda \int_t^{t+\Delta t} E\left[e^{i\xi \ln[1+Uh(t+\Delta t-\tau)]} - 1\right] d\tau\right)$$
$$\times \exp\left(\lambda \int_0^{t} E\left[e^{i\xi \ln\left[\frac{1+Uh(t+\Delta t-\tau)}{1+Uh(t-\tau)}\right]} - 1\right] d\tau\right) \qquad (1.1.3.4)$$

Proof. See the Appendix ■

These results must nest those obtained for simplest processes, as jump-diffusion processes without shot noise effects. The next example links the SNJD model with the JD processes.

Example 1.1 *In the case of a pure jump process ($h = 1$), expression (1.1.3.4) becomes*

$$\Phi_{Z_t}(\xi) = \exp\left(i\left(\mu - \frac{\sigma^2}{2}\right)\Delta t\xi - \frac{1}{2}\sigma^2\Delta t\xi^2 + \lambda E\left[e^{i\xi \ln[1+U]} - 1\right]\Delta t\right)$$

As expected, this characteristic function equates that obtained by Das and Sundaram (1999) for a JD process.

1.3.2 Moments of the process

Once obtained the characteristic function of the log-returns of the stock price given by model (1.1.2.2), the non-central unconditional moments can be computed by differentiating expression (1.1.3.4), that is,

$$\mu'_n = E\left[Z_t^n\right] = \frac{1}{i^n} \left.\frac{\partial \Phi_{Z_t}}{\partial \xi}\right|_{\xi=0} \qquad (1.1.3.5)$$

where μ'_n denotes the *n-th* non-central moment. Given the relationship between central moments, denoted by μ_n, and non-central moments (see Abramowich and Stegun (1968), 26.1.14), the next lemma provides the first four moments for the log-returns of the SNJD process.

Lemma 2 *The first four moments of the log-returns for a SNJD process*

(see (1.1.2.2)) are given by

$$\begin{aligned} E[Z_t] &= \left(\mu - \frac{\sigma^2}{2}\right)\Delta t + \lambda \int_t^{t+\Delta t} E\left[\ln\left(1 + Uh\left(t + \Delta t - \tau\right)\right)\right] d\tau \\ &\quad + \lambda \int_0^t E\left[\ln\left(\frac{1 + Uh\left(t + \Delta t - \tau\right)}{1 + Uh\left(t - \tau\right)}\right)\right] d\tau \end{aligned}$$

$$\begin{aligned} Var[Z_t] &= E\left[\left(Z_t - E\left(Z_t\right)\right)^2\right] = \mu_2' - \left(\mu_1'\right)^2 \\ &= \sigma^2 \Delta t + \lambda \int_t^{t+\Delta t} E\left[\ln^2\left(1 + Uh\left(t + \Delta t - \tau\right)\right)\right] d\tau \\ &\quad + \lambda \int_0^t E\left[\ln^2\left(\frac{1 + Uh\left(t + \Delta t - \tau\right)}{1 + Uh\left(t - \tau\right)}\right)\right] d\tau \end{aligned}$$

$$Skewness = \frac{\mu_3}{\mu_2^{3/2}} = \frac{2\left(\mu_1'\right)^3 - 3\mu_1'\mu_2' + \mu_3'}{\mu_2^{3/2}} = \frac{1}{\mu_2^{3/2}}\left(C^3 + D^3\right)$$

$$Kurtosis = \frac{\mu_4}{\mu_2^2} = \frac{-3(\mu_1')^4 + 6(\mu_1')^2\mu_2' - 4\mu_1'\mu_3' + \mu_4'}{\mu_2^2} = 3 + \frac{C^4 + D^4}{\mu_2^2}$$

with

$$\begin{aligned} C^n &= \lambda \int_t^{t+\Delta t} E\left[\left\{\ln\left(1 + Uh\left(t + \Delta t - \tau\right)\right)\right\}^n\right] d\tau \\ D^n &= \lambda \int_0^t E\left[\left\{\ln\left(\frac{1 + Uh\left(t + \Delta t - \tau\right)}{1 + Uh\left(t - \tau\right)}\right)\right\}^n\right] d\tau \end{aligned}$$

Proof. See the Appendix ■

Former expressions generalize the results obtained for GBM and JD. As it was previously shown, it is also possible to connect the moments of a SNJD process with those of the JD process as the next example shows.

Example 1.2 *Let us consider a JD process as that posited by Merton (1976). In this case,* $h(t) = 1$*, and*

$$\begin{aligned} C^n &= \lambda \Delta t E\left[\left\{\ln\left(1 + U\right)\right\}^n\right] \\ D^n &= 0 \end{aligned}$$

Then, the first moments of the log-returns for a SNJD process are as follows:

$$\begin{aligned} E[Z_t] &= \left(\left(\mu - \frac{\sigma^2}{2}\right) + \lambda E[\ln(1+U)]\right)\Delta t \\ Var[Z_t] &= \left(\sigma^2 + \lambda E\left[\{\ln(1+U)\}^2\right]\right)\Delta t \\ Skewness &= \frac{1}{\sqrt{\Delta t}}\left[\frac{\lambda E\left[\{\ln(1+U)\}^3\right]}{\left(\sigma^2 + \lambda E\left[\{\ln(1+U)\}^2\right]\right)^{3/2}}\right] \\ Kurtosis &= 3 + \frac{1}{\Delta t}\left[\frac{\lambda E\left[\{\ln(1+U)\}^4\right]}{\left(\sigma^2 + \lambda E\left[\{\ln(1+U)\}^2\right]\right)^{2}}\right] \end{aligned}$$

as obtained in Das and Sundaram (1999) or Aït-Sahalia (2004)[20].

To compare the skewness and excess of kurtosis generated by the JD and SNJD models, Table 1.2 displays the values obtained in both models[21]. Jump returns are normally distributed with $ln(1+U) \sim N(\theta, \beta^2)$. As it will be our reference through the paper, the SNJD model has been simulated by using a response function of the form $h(\cdot) = e^{-a(\cdot)}$. Simulation parameters are also given in Table 1.2.

By and large, Table 1.2 exhibits parameters of skewness and excess of kurtosis for JD and SNJD process very close each other. In general, this fact is corroborated across different time periods or different set of parameters. The causes of this similarities arise in the form of the response function chosen, as an inspection of the skewness and kurtosis equations for both models reveals: it seems that an exponential decaying function response produces a behavior in the first moments equal to those generated by a JD model[22].

Advancing some results that will be provided in next sections, it is worth to mention that these similarities among the JD and SNJD model with exponential response function do not happen in cross-moments, as Figure 1.7 reflects: the autocorrelation coefficients for first lags of the SNJD model

[20] Aït-Sahalia (2004) also provides an expression for the centered, absolute moments of non-integer order.

[21] This Table is directly inspired in Das and Sundaram (1999).

[22] This point will be mathematically clarified by the numerical approximation result of the next subsection.

Table 1.2: Skewness and excess of kurtosis in the JD and SNJD processes

Parameters			Skewness						Excess of kurtosis					
			1d		1w		1m		1d		1w		1m	
σ	θ	β	JD	SNJD	JD	SNJD	JD	SNJD	JD	SNJD	JD	SNJD	JD	SNJD
0.24	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.33	0.33	0.07	0.07	0.02	0.02
0.24	-0.001	0.02	-0.01	0.00	0.00	0.00	0.00	0.00	0.32	0.33	0.07	0.07	0.02	0.02
0.24	-0.100	0.02	-2.77	-2.60	-1.26	-1.18	-0.60	-0.57	12.07	11.17	2.49	2.31	0.57	0.53
0.20	0.00	0.04	0.00	0.14	0.00	0.06	0.00	0.03	6.17	6.27	1.27	1.29	0.29	0.30
0.20	-0.001	0.04	-0.06	0.08	-0.03	0.04	-0.01	0.02	6.07	6.22	1.25	1.28	0.29	0.30
0.20	-0.100	0.04	-3.81	-3.57	-1.73	-1.62	-0.83	-0.78	19.76	19.02	4.08	3.92	0.94	0.90
0.15	0.00	0.06	0.00	0.66	0.00	0.30	0.00	0.14	28.63	29.53	5.91	6.09	1.36	1.41
0.15	-0.001	0.06	-0.12	0.55	-0.06	0.24	-0.03	0.12	28.32	29.35	5.84	6.06	1.35	1.40
0.15	-0.100	0.06	-5.23	-4.93	-2.38	-2.24	-1.14	-1.08	31.26	32.07	6.45	6.62	1.49	1.53

Skewness and excess of kurtosis for the JD and SNJD process. Jump size in returns is normally distributed with $ln(1+U) \sim N(\theta, \beta^2)$. Response function in SNJD process is $h(\cdot) = e^{-a(\cdot)}$. The annual volatility has been fixed to $\sigma_{annual} = 0.25$ for all cases. Parameter λ (the number of jumps per year) is also constant and equal to 10. Taking into account the expression for the total variance of the JD process $\sigma^2\Delta t + \lambda\Delta t E[\{\ln(1+U)\}^2]$ is possible to keep intuition about the results obtained: for example, a parameter $\beta = 0.02$ represents about 10% of contribution of the jumps to the total variance; similarly, values of $\beta = 0.04, 0.06$ constitutes a 36% and 64% over the total variance of the JD process due to jumps.

are significant and different from zero; by contrast, the GBM and JD models are not significant and close to zero value, as Figure 1.7 reveals.

Some additional conclusions arise from Table 1.2. It seems that the jumps are the source of asymmetries in returns generated by JD and SNJD processes. This is corroborated by the fact that, generally speaking, no skewness is produced in the case of symmetric jumps ($\theta = 0$). Additionally, in the case of negative values for jump means, it is possible to observe how as the contribution of jumps to the total variance of the process increases (from 10% for $\beta = 0.02$ to about 60% of $\beta = 0.06$) the magnitude of the skewness also increases, passing from -2.70 to about -5.00 in daily frequency (the same applies for weekly and monthly ones).

Similarly to the skewness case, the jumps are also responsible of the kurtosis in JD and SNJD processes. As expected, the models of jumps produces excess of kurtosis higher than usual in any set of parameters. For instance, observe the huge differences of excess of kurtosis among cases without (almost) jump variance contribution ($\sigma = 0.24$) and high jump variance ($\sigma = 0.15$).

Finally, Table 1.2 exhibits a decaying in skewness and excess of kurtosis that tends to the Normal distribution when the time frequency is reduced[23]. This fact is interpreted as a direct result of the Central Limit Theorem.

1.3.3 Numerical approximations

Expressions of moments include some integrals of the type

$$C^n = \lambda \int_t^{t+\Delta t} E\left[\left\{\ln\left(1 + Uh\left(t + \Delta t - \tau\right)\right)\right\}^n\right] d\tau \qquad (1.1.3.6)$$

$$D^n = \lambda \int_0^t E\left[\left\{\ln\left(\frac{1 + Uh\left(t + \Delta t - \tau\right)}{1 + Uh\left(t - \tau\right)}\right)\right\}^n\right] d\tau \qquad (1.1.3.7)$$

whose computation could be burdensome. To avoid this problem, the next lemma provides some useful approximations.

[23] Das and Sundaram (1999) mention that the rate of decaying of skewness and kurtosis in JD models is faster than that empirically observed. According to this, the authors notice that implied volatility smiles should not exist at long-periods (three months), as it is contrary to the market. This could be considered as a drawback of the JD models.

Lemma 3 *The expressions (1.1.3.6) and (1.1.3.7) can be approximated by*

$$C^n \simeq \lambda E\left[U^n\right] \int_t^{t+\Delta t} \left[h(t+\Delta t-\tau)\right]^n d\tau \tag{1.1.3.8}$$

$$D^n \simeq \lambda E\left[U^n\right] (\Delta t)^n \int_0^t \left[\frac{d}{dt} h(t-\tau)\right]^n d\tau \tag{1.1.3.9}$$

Proof. See the Appendix. ■

As it will be of interest, the next example provides the numerical approximations for C^n and D^n using the numerical response function in (1.1.2.3)

Example 1.3 *In case of* $h(\cdot) = e^{-a(\cdot)}$*, the terms (1.1.3.6) and (1.1.3.7) can be approximated by*

$$C^n \simeq \lambda E\left[U^n\right] \Delta t \tag{1.1.3.10}$$

$$D^n \simeq \lambda E\left[U^n\right] (-a\Delta t)^n \frac{1}{na}\left(1 - e^{-nat}\right) \tag{1.1.3.11}$$

Expressions (1.1.3.10) and (1.1.3.11) justify the results included in Table 1.2: since the term $(\Delta t)^n$ is almost zero for $n \geq 2$, equation (1.1.3.11) vanishes. On the other hand, the form of term C^n reflects that, in this special response function, the SNJD model behaves as an standard JD model.

1.4 Cross moments and spectra

As was exhibited in Figure 1.7, the term (1.1.3.3) introduces some kind of serial dependence in log-returns. Then, it is clear that some kind of information could be presented in the cross moments of the series generated by the model (1.1.2.2). With regard to the former, it could be interesting to look carefully to the dynamics of the process, with attention to its cross moments and its spectrum.

This section computes the autocovariance and population spectrum of the log-returns of the process (1.1.2.2). The last part of this section illustrates some related intuitions concerning the serial dependence by using simulations. Some of the ideas presented here will be treated in detail in Chapter 3.

1.4.1 Autocovariance

Prior to computing the autocovariance of the log-returns of process (1.1.2.2), it is important to formalize three assumptions concerning to the term Y_t and H_t (see equations (1.1.3.2) and (1.1.3.3)), respectively, that simplify enormously the computations involved in.

Assumption 1 *The increments of the Poisson process can be approximated by a Bernoulli variable. If so, we have that*[24]*,*

$$Y_t = \sum_{j=N_t+1}^{N_{t+\Delta t}} \ln\left[1 + U_j h\left(t + \Delta t - \tau_j\right)\right] \tag{1.1.4.1}$$

$$\simeq \begin{cases} 0, & \text{if } N_{t+\Delta t} - N_t = 0 \\ U_{1+N_t}\, h(0), & \text{if } N_{t+\Delta t} - N_t = 1, \text{ with } N_t = 0, 1, 2, \ldots \end{cases} \tag{1.1.4.2}$$

being Δt a small time increment.

Now, two more technical assumptions must be added.

Assumption 2 *There is no effect of the response function in the moment of event happens ($h(0)$).*

Assumption 3 *The term H_t in equation (1.1.3.3) can be approximated by*

$$H_t = \sum_{j=1}^{N_t} U_j (h\left(t + \Delta t - \tau_j\right) - h\left(t - \tau_j\right)) \tag{1.1.4.3}$$

Basically, assumption 2 states that the only contribution in the time that jump event happens is (just) due to the random variable U_j[25]. Additionally,

[24] Basically, we assume that the probability of one jump happened within the interval $[t,\, t + \Delta t]$ is

$$\begin{aligned} P[N_{t+\Delta t} - N_t = 0] &\simeq 1 - \lambda \Delta t \\ P[N_{t+\Delta t} - N_t = 1] &\simeq \lambda \Delta t \end{aligned}$$

and zero otherwise. For a detailed discussion of this point see Ball and Torous (1983). Ait-Sahalia (2004) also provides additional insights about this issue.

[25] Our intention here is to avoid additional contributions to the jump shock in the same moment that jump triggers. As we will see, by imposing this restriction on the SNJD process, we will be able to consider as independent random variables the elements in Y_t of equation (1.1.3.2), for small time intervals Δt.]

Assumption 3 just introduces a more suitable representation of expression H_t based on the power expansion of the logarithm. As we will see, this result will become useful for computing the autocovariance of the process Z_t.

The independence between B_t, N_t and U_j simplifies enormously the computation of the autocovariance of a SNJD process. As we refer in the appendix, we address to the reader to Parzen (1962) for a formal proof of the autocovariance of the process H_t (see expression (1.1.3.3)).

The next lemma provides the autocovariance for log-returns of the process (1.1.2.2).

Lemma 4 *The autocovariance of the SNJD process is given by*

$$\begin{aligned} Cov\left[Z_{t+n\Delta t}, Z_t\right] &= \sigma^2 \Delta t \delta(n) + \lambda E[U^2] \Delta t \delta(n) \\ &+ \lambda \int_0^t E\left[\ln\left(\frac{1+Uh\left(t+\Delta t-\tau\right)}{1+Uh\left(t-\tau\right)}\right)\ln\left(\frac{1+Uh\left(t+(n+1)\Delta t-\tau\right)}{1+Uh\left(t+n\Delta t-\tau\right)}\right)\right] d\tau \\ &+ Cov\left[\sum_{j=N_t+1}^{N_{t+\Delta t}} U_j, \sum_{j=N_t+1}^{N_{t+n\Delta t}} U_j(h(t+n\Delta t-\tau_j)-h(t-\tau_j))\right] \end{aligned} \tag{1.1.4.4}$$

where $n = 0, 1, 2, \ldots$ *denotes the lag and* $\delta(\cdot)$ *is the usual delta Dirac function.*

Proof. See the Appendix ■

Notice the fact that the Shot noise is the responsible of the autocovariance of the SNJD process. Again, as expected, this result generalizes those obtained for standard diffusion process as GBM or JD.

1.4.2 Population spectrum

As it will be developed in Chapter 3, due to the dynamical properties of process (1.1.2.2) could be interesting to explore the estimation of the Shot noise process in the frequency domain. According to this, the computation of an expression for the population spectrum of log-returns is necessary.

Let $s_Z(w)$ denote the population spectrum of a covariance-stationary process Z, where w denotes the set of different *frequencies*[26]. As Hamilton (1994) points out, this spectrum is computed by applying the Fourier

[26]These frequencies w will be defined more precisely in Chapter 3.

transform to the autocovariance function of the process:

$$s_Z(w) = \int_{-\infty}^{+\infty} \gamma(\varepsilon)\, e^{-iw\varepsilon} d\varepsilon \tag{1.1.4.5}$$

where $\gamma(\varepsilon)$ denotes the autocovariance function $Cov[Z_{t+\varepsilon}, Z_t]$ of the process Z_t and ε the time lag[27].

A previous result (see Hamilton (1994), p. 172) will be useful to estimate the population spectrum of processes of type (1.1.2.2). This result establishes that the population spectrum of the sum of uncorrelated processes is equal to the sum of the corresponding spectra. Then, given the stochastic process $Z_t = X_t + Y_t + H_t$ (with no correlation between them), it is possible to demonstrate that

$$s_Z(w) = s_X(w) + s_Y(w) + s_H(w) \tag{1.1.4.6}$$

Therefore, the expression of the log-returns for the model (1.1.2.2) can be written as the sum of two white noises (from the Brownian and JD parts, respectively) plus a Shot noise component.

The following lemma provides the population spectrum of the SNJD process by means of applying the Fourier transform to (1.1.4.4).

Lemma 5 *The population spectrum of the log-returns of the process (1.1.2.2) is given by*

$$\begin{aligned} s_Z(w) &= \frac{\sigma^2}{2\pi}\Delta t + \frac{\lambda}{2\pi}E[U^2]\Delta t \\ &\quad + \frac{\lambda}{2\pi}\int_{-\infty}^{+\infty} e^{-iw\varepsilon} d\varepsilon \\ &\quad \times \int_0^t E\left[\ln\left(\frac{1+Uh(t+\Delta t-\tau)}{1+Uh(t-\tau)}\right)\ln\left(\frac{1+Uh(t+\Delta t+\varepsilon-\tau)}{1+Uh(t+\varepsilon-\tau)}\right)\right] d\tau \\ &\quad + \frac{\lambda}{2\pi}\int_{-\infty}^{+\infty} e^{-iw\varepsilon} d\varepsilon \\ &\quad \times Cov\left[\sum_{j=N_t+1}^{N_{t+\Delta t}} U_j, \sum_{j=N_t+1}^{N_{t+\varepsilon}} U_j(h(t+\varepsilon-\tau_j) - h(t-\tau_j))\right] \end{aligned} \tag{1.1.4.7}$$

[27]For the ease of explanation, we will introduce here a new notation for time lags. Without loss of generality, we denote $\varepsilon = n\Delta t$, $\delta(\varepsilon) = \delta(n)$ where $n = 0, 1, 2, \ldots$

Proof. See the Appendix ■

The last expression seems too general and must be computed individually for any specification of response function $h(\cdot)$. Anyway, this result shows that the addition of a shot noise process in standard diffusion processes - as GBM or JD - may lead to different distribution of power along frequencies[28].

1.5 The exponential decaying function case

The SNJD with exponential response function constitutes the paradigm of the model posited by Altmann *et al* (2007). It could be considered as the simplest, most intuitive form of introducing a deterministic behavior after the jump.

This section provides some results about the SNJD model with exponential decaying function. Most of them have already been mentioned, and they are just applications of former sections. To the best of our knowledge, there are no references about other candidates for this function. Anyway, the results given here will be used in next chapters to check the performance of the different methodologies proposed for estimating the parameters of a SNJD model.

1.5.1 Moments and autocovariance function

We present here the main first moments and correlation expressions for the SNJD model with exponential decaying function.

By replacing the function $h(\cdot) = e^{-a(\cdot)}$ in expression of moments in Lemma 2, we obtain

$$E\left[Z_t\right] = \left(\mu - \frac{\sigma^2}{2}\right)\Delta t \tag{1.1.5.1}$$

$$Var\left[Z_t\right] = \left(\sigma^2 + \lambda E[U^2]\right)\Delta t \tag{1.1.5.2}$$

[28] As we will detail later, this is an aspect that can be relevant to explain the power at low frequencies that appears in some economic variables (Lo, 1991). Of course, this must not be intended as the Brownian or jump parts do not contribute to the spectrum of the process.

$$Skewness = \frac{1}{\sqrt{\Delta t}}\left[\frac{\lambda E[U^3]}{(Var[Z_t])^{3/2}}\right] \quad (1.1.5.3)$$

$$Kurtosis = 3+\frac{1}{\Delta t}\left[\frac{\lambda E[U^4]}{(Var[Z_t])^{2}}\right] \quad (1.1.5.4)$$

where we have used the approximations for exponential decaying function of the example 3.

Notice that expressions (1.1.5.1)–(1.1.5.4) reduce to those of the JD process in Example 2. At light of the last results, it seems that the SNJD process with $h(\cdot)$ function of the exponential form is similar to a JD process, at least in their first four moments[29]. This last fact could explain the results previously obtained in Table 1.2, where results for JD and SNJD process are close each other.

Although the JD and SNJD with exponential response function processes are similar in their first four moments, their main differences arise in their cross moments, as it will be presented in the next lines. For illustrative purposes, consider the following example where we provide an expression for the autocovariance of the SNJD process,

Example 1.4 *In the case that $h(t) = e^{-at}$, the autocovariance for the log-returns of the process (1.1.2.2) is given by*

$$Cov\left[Z_{t+n\Delta t}, Z_t\right] = \sigma^2\Delta t\delta(n) + \lambda E\left[U^2\right]\Delta t\delta(n) + \lambda(e^{-a\Delta t}-1)\,E[U^2]\,\Delta t\,e^{-a\,n\Delta t} \quad (1.1.5.5)$$

where $n = 0, 1, 2, \ldots$ denotes the lag and $\delta(\cdot)$ is the usual delta Dirac function.

Proof. See the Appendix ■

The first two terms in equation (1.1.5.5) reflect the contribution of the GBM and jump parts to the autocovariance of the process. As expression (1.1.5.5) reflects, these two terms are null for non-zero lags. This means

[29] More precisely, the SNJD with exponential decaying function and the JD processes are *almost* equal in their first four moments. Notice that terms into expectations differ from $E[ln(1+U)^n]$ for a JD process, with those of $E[U^n]$ for a SNJD process. Again, the origin of these differences is related to using the numerical approximations involved in the computation of SNJD process.

that there is no additional contribution in the total variance of the process through lagged observations. This fact differs from the last term in (1.1.5.5), where it shows that the covariance has an exponential decaying pattern due to the shot noise term: in other words, an additional source of variability of the SNJD process is contained in past information.

Once provided the variance and autocovariance of the SNJD process with the exponential response function (see expressions (1.1.5.2) and (1.1.5.5)), the next example gives an expression for the autocorrelation.

Example 1.5 *The autocorrelation function for the log-returns of the exponential decaying case is*

$$\begin{aligned} Corr\left[Z_{t+n\Delta t}, Z_t\right] &= \frac{Cov\left[Z_{t+n\Delta t}, Z_t\right]}{\sqrt{Var(Z_{t+n\Delta t})}\sqrt{Var(Z_t)}} \\ &= \begin{cases} 1, & \text{if } n = 0 \\ (e^{-a\Delta t} - 1)\frac{\lambda E[U^2]}{\sigma^2 + \lambda E[U^2]}\, e^{-a\, n\Delta t}, & \text{if } n \geq 1 \end{cases} \end{aligned} \tag{1.1.5.6}$$

An inspection of the former result reveals two main results:

- The autocorrelation is always *negative* for positive values of a.
- The factor $\frac{\lambda E[U^2]}{\sigma^2 + \lambda E[U^2]}$, which represents the variance of the jump divided by that of the whole process, *modulates* the degree of correlation of the process.

To illustrate numerically the former equation, Table 1.3 includes the sample ($\widehat{\rho}_n$) and theoretical autocorrelation values (ρ_n^{theo}) of a SNJD model. The first column contains the number of lags. Each successive pair of columns exhibit, respectively, the theoretical ACF's obtained from equation (1.1.5.6) and the mean of a sample ACF coefficients of 100 simulation paths with 1,800 steps each one. Finally, this experiment is repeated for different values of the parameter a.

Some conclusions arise from Table 1.3:

1. As expected, the parameter a controls the persistence of the process over time.

Table 1.3: Autocorrelation function for different values of the parameter a

The theoretical expression for the autocorrelation functions for the SNJD process with exponential decaying function is

$$\rho_n(H^{rst}) = \begin{cases} 1, & \text{if } n = 0 \\ (e^{-a\Delta t} - 1)\,\frac{\lambda E[U^2]}{\sigma^2 + \lambda E[U^2]}\, e^{-a\, n\Delta t}, & \text{if } n \geq 1 \end{cases}$$

SNJD process

	$a = 0.2$		$a = 0.6$		$a = 1.0$		$a = 10.0$	
Lag	ρ_n^{theo}	$\widehat{\rho}_n$	ρ_n^{theo}	$\widehat{\rho}_n$	ρ_n^{theo}	$\widehat{\rho}_n$	ρ_n^{theo}	$\widehat{\rho}_n$
1	-0.144	-0.068	-0.240	-0.172	-0.225	-0.250	0.000	-0.416
2	-0.118	-0.057	-0.132	-0.096	-0.083	-0.091	0.000	0.001
3	-0.096	-0.046	-0.072	-0.047	-0.031	-0.031	0.000	-0.000
4	-0.079	-0.038	-0.040	-0.029	-0.011	-0.013	0.000	-0.005
5	-0.065	-0.027	-0.022	-0.016	-0.004	-0.003	0.000	0.005
10	-0.024	-0.013	-0.001	0.005	0.000	-0.001	0.000	0.002
25	-0.001	0.001	0.000	-0.001	0.000	0.005	0.000	-0.003
50	0.000	0.001	0.000	-0.001	0.000	-0.001	0.000	-0.001
100	0.000	-0.003	0.000	-0.001	0.000	-0.001	0.000	-0.003

Comparison of autocorrelation coefficients for different values of the parameter a. Each column a corresponds to the means of sample autocorrelation coefficients of 100 simulation paths with 1,800 steps each one of a SNJD model with response function $h(t-\tau_k) = e^{-a(t-\tau_k)}$ and the random variable U lognormally distributed as $U = e^{-\beta^2/2+\beta\varepsilon} - 1$, with $\varepsilon \sim N(0,1)$ and mean $E[U] = 0$. Simulated (annualized) parameters are $\mu = 0.05$, $\sigma = 0.20$, $\lambda = 10.00$, $\beta = 0.10$.

2. The theoretical autocorrelation values for the first few lags of $a = 0.6$ and 1.0 seem to be closer to those obtained for $a = 0.2$ and 10.0, which could be a result of the numerical approximations carried out.

3. Any autocorrelation value for the SNJD process under study exhibits *negative* values, in contrast to the classic Shot noise results reported in Table 1.1.

Lastly, we provide an additional evidence about the degree of persistence generated *just* by the contribution of the Shot noise part in the SNJD model: Figure 1.8 displays the mean sample autocorrelation coefficients of 100 simulated paths composed by 1,800 data each one. The autocorrelation coefficients correspond to the values, absolute values and squared values of log-returns in the process (1.1.2.2), respectively.

As we can see, the correlation coefficients are negative and significant up to third order for the case of simulated values of a SNJD with exponential decay function. Furthermore, (positive) autocorrelation coefficients go even to order six in case of absolute log-returns. Again, these results seem to confirm that the SNJD model is capable of generate autocorrelation due to the Shot noise process.

1.5.2 The spectrum

Traditionally, the statistical tools of Spectral Analysis have not been widely used in the context of financial econometrics in continuous time. We provide here some results that seems to evidence the usefulness of considering the estimation of the SNJD process with exponential decaying function in the frequency domain[30]. Some additional insights about the process and the estimation procedure in the frequency domain will be treated in a larger detail in Chapter 3.

To start, consider the population spectrum for a SNJD process with exponential decaying response function,

Example 1.6 *If $h(\cdot) = e^{-a(\cdot)}$, the population spectrum of the log-returns for the model (1.1.2.2) is given by*

$$s_Z(w) = \frac{1}{2\pi}\left(\sigma^2 + \lambda E[U^2] \; + \lambda(e^{-a\Delta t} - 1)E\left[U^2\right]\frac{2a}{a^2 + w^2}\right)\Delta t \quad (1.1.5.7)$$

[30] We are aware that Spectral Analysis tools are not widely known among financial economists, and its intuition and usefulness must been carefully explained. For the sake of brevity, we devote the Chapter 3 to this task, and this section just provides some results that will be used in the future.

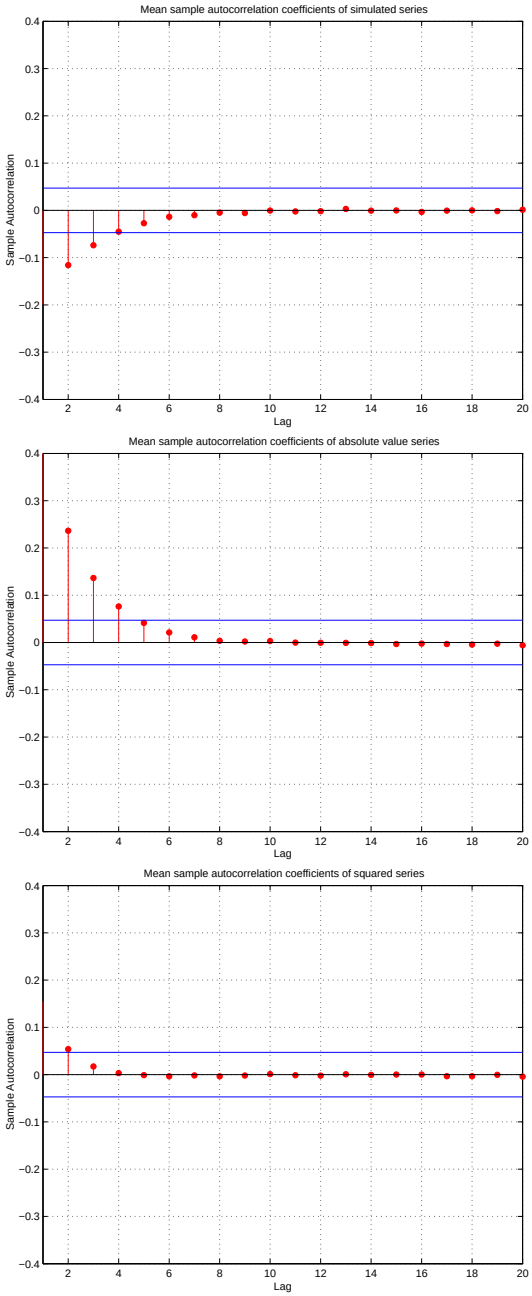


Figure 1.8: Sample autocorrelation coefficients for a SNJD process with response function $h(t-\tau_k) = e^{-a(t-\tau_k)}$ and random variable U_j is lognormally distributed as $U = e^{-\beta^2/2+\beta\varepsilon} - 1$, with $\varepsilon \sim N(0,1)$ and mean $E[U] = 0$. Simulated parameters are $\sigma = 0.00$, $\lambda = 10$, $\beta = 0.10$ and $a = 0.50$.

Proof. See the Appendix ■

The expression of spectrum in (1.1.5.7) splits the total spectrum of the SNJD process in three different sources of variability or *noise*:

1. The variance associated to the GBM part, $\sigma^2/2\pi$, which is constant and it does not depend on frequency w.
2. The part linked to the jump event, which is also constant and equal to $\lambda E[U^2]/2\pi$
3. The Shot noise part, which depends on the frequency w.

As it will be presented in Chapter 3, the noise whose spectrum is constant for any frequency is called a *white noise* process. As a proof of consistency of the expression (1.1.5.7), we observe that, as expected, the spectrum for a JD process (the previous one for $a = 0$) is the sum of two white noises:

$$s_Z(w) = \frac{1}{2\pi}\left(\sigma^2 + \lambda E[U^2]\right)\Delta t$$

Finally, to get a graphical representation of the last issues, Figure 1.9 exhibits the sample and theoretical spectrum for a SNJD and JD process. Some details about the estimation are given in the figure.

As Figure 1.9 shows, the main difference across the spectra are the positive slope that we observe in the case of a SNJD model: the spectrum for the JD process remains constant, as predicted by the theory (blue line); by contrast, the SNJD process spectrum seems to present a increasing pattern along different frequencies[31].

1.6 The economic significance of the SNJD model

This section briefly addresses some of the possible implications of using a SNJD model as posited by Altmann *et al* (2007). Basically, we outline the possible applications of this type of models.

1.6.1 Long range dependence

Campbell *et al* (1997) refer as long-range memory as observations in the remote past that are nontrivially correlated with observations in the distant

[31] The waved shape observed in both figures results from the sample estimate used. What we are interested on the figures is the general slope of the curve, and not in the behavior of local regions.

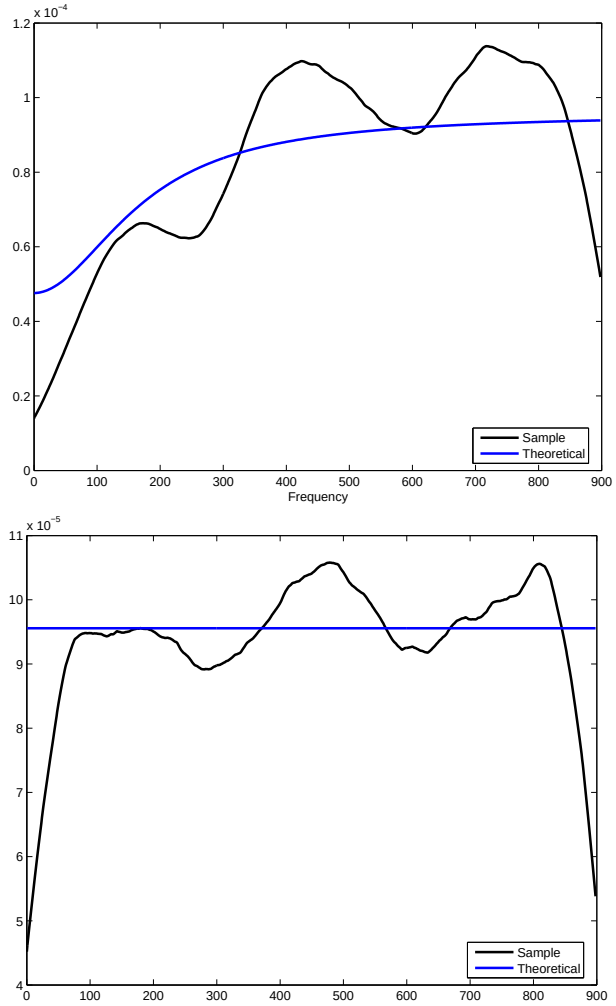


Figure 1.9: Sample and theoretical spectrum for a SNJD and JD process. Upper figure correspond to those spectrum obtained for a SNJD process, and lower one to those for a JD process. The sample spectrum has been made by using a non-parametric estimate posited in Hamilton (1994) with bandwidth fixed to $h = 100$. SNJD model response function is $h(t - \tau_k) = e^{-a(t-\tau_k)}$ with random variable U lognormally distributed as $U = e^{-\beta^2/2+\beta\varepsilon} - 1$, $\varepsilon \sim N(0, 1)$ and mean $E[U] = 0$. Simulated parameters (annualized) are $\sigma = 0.05$, $\lambda = 10$, $\beta = 0.10$ and $a = 0.60$.

future, even as the time span between the two observations increases. As it is pointed out by Singleton (2006), an important amount of studies indicate the presence of some kind of serial dependence in stock returns.

With regard to this issue, Lo (1991) enumerates some possibles implications of long-range dependence on financial economics. He also emphasized the inconsistency with long-memory of the most common continuous-time models. Moreover, Lo (1991) also concludes that stochastic models for short-range dependence could capture correctly the behavior of asset returns.

Although the previous evidences of the SNJD process suggest the capability of the model to generate short-range persistence in log-returns (see expression (1.1.5.6) and figure 1.8), we are inclined to believe that a suitable specification of function $h(\cdot)$ of equation (1.1.2.2) could lead to produce an pattern of dependence in long term.

1.6.2 The spikes

Some economic series exhibit abrupt changes in their value in short periods of time. Visually, we can observe how this type of series exhibits a jump followed by a instantaneous drop back to the previous level (Weron, 2005). This kind of effects is often named *spikes* in the financial literature[32]. Economically, its origin could be due to changes on the monetary policy (Das (2002) or Benito *et al*, 2007) or extreme fluctuations among supply and demand (Barone-Adesi and Gigli (2002) or Lucía and Schwartz, 2002).

As an example, Figure 1.10 displays two graphs: the upper one exhibits the sample path of the European Overnight Index Average (EONIA) rate, r_t, and its increments, $r_t - r_{t-1}$; the lower one exhibits the sample spectrum using the EONIA increments. The sample period ranges from 05/02/2004 to 23/08/2005.

The upper graph in Figure 1.10 remembers the path of the classic Shot noise process displayed in Figure 1.1. It is possible to observe in this figure the cited spikes spread on the whole path. It seems that these spikes are responsible of the higher contribution in the sample variance produced in higher frequencies observed in the sample spectrum of figure 1.10. The shape of the curve recalls the theoretical pattern described by a SNJD model with exponential decaying function exhibited in the upper graph of Figure 1.8.

Taking previous considerations into account, the SNJD model could be considered as an adequate candidate for dealing with such empirical patterns.

[32] We reserve the name of *jumps* for those events that, once triggered, return -or not- to their previous position at a lower rate than those for the spikes.

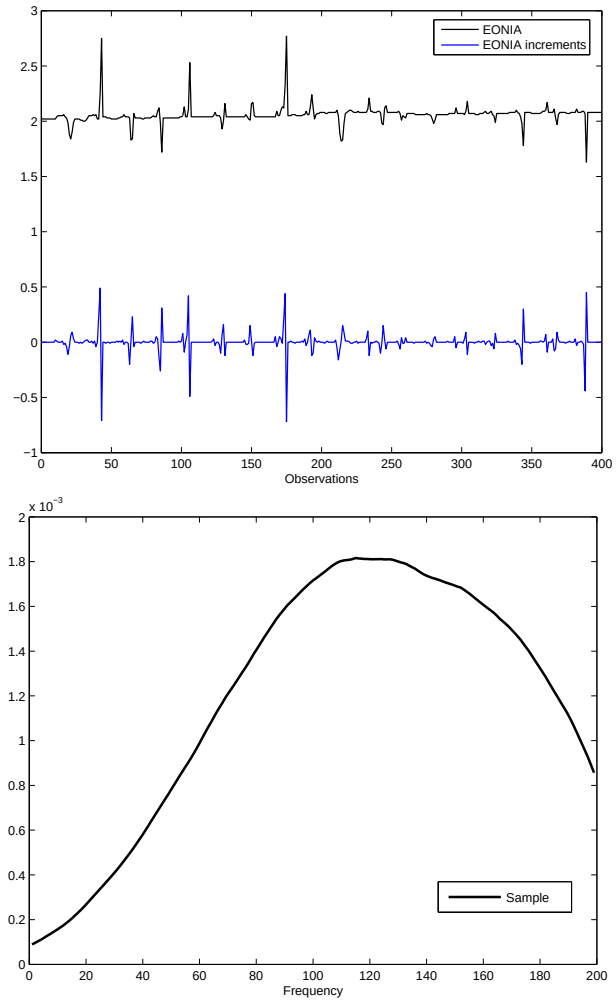


Figure 1.10: Sample path and spectrum for the EONIA rate. The upper graph displays the sample path (black line) and the increments (blue line) of the EONIA. The lower graph shows the sample spectrum obtained by using a non-parametric estimate posited in Hamilton (1994) with bandwidth fixed to $h = 50$.

1.6.3 Risk management

On the one hand, the estimation of value at risk (VaR) and other tail statistics depends dramatically on the large price changes. As pointed out by Ait-Sahalia (2004), the foundations of many of these techniques are oriented to distinguish between the eventual jumps from the daily brownian noise. The presence of a pattern of serial dependence in asset returns could leads to a miscalculation in the components of the total volatility of the data generating process, leading to possible model risk effects.

On the other hand, some standard risk tools in option pricing, as the delta hedging, are designed for small changes in the prices (Duffie and Singleton, 2003). To the best of our knowledge, we have no evidence, empirical or simulated, about the consequences for risk management of the persistence over the time of jumps effects in the underlying. Concerning this question, we are inclined to believe that the SNJD model could provide additional insights on this issue.

1.7 Conclusions

To a certain extent, the formulation of modern finance from Black and Scholes (1973) to our days is based on continuous-time stochastic processes. From the whole amount of available processes, the Affine-jump diffusion model family studied by Duffie *et al* (2000) seems to offer an optimum balance between empirical adequacy and analytical tractability. However, several recent empirical studies raise the question about considering new models that should be capable of capturing some of the drawbacks present in the existing models by, of course, keeping the usual requirements about parsimony on the number of variables and economic intuition.

This chapter has presented a model that adjusts to former requests. We focuses on the model posited by Altmann *et al* (2007), the SNJD model. Basically, it is an extension of the jump-diffusion models proposed by Merton (1976) where an additional term (named shot-noise function) is added to the Poisson process. The intuition behind is that after a jump triggers, its effect can fade away on the long run. As a result of this, the SNJD model is capable of generate serial persistence in asset returns. As far as we know, this is the first study that analyzes in detail the main features of the SNJD model.

We provide a general treatment of the Shot noise processes in the literature. To the best of our knowledge, no similar references have been found in the financial literature. We give an expression for log-returns of the SNJD

model, and we identify the source of autocovariance in the process. Not surprisingly, we find that the SNJD model is capable of nesting former versions of continuous-time processes as GBM or JD.

We provide the Characteristic function of the SNJD process and derive some of its main moments, also providing some useful numerical approximations for computing purposes. Regarding the dynamics of the model, we provide the covariance function and the spectrum of the SNJD process. As we have pointed out, we observe that a considerable source of information is embedded in past information. Moreover, the analysis in the frequency domain appears to be revealed as a powerful tool when estimating the SNJD process.

Finally, an extension to the paradigm of this models, the SNJD model with exponential decaying function, is also studied. Our results seems to conclude that the SNJD model with exponential decaying function is close to the JD version if we focus on their central moments, differing in their autocorrelation function specifications. To sum up, some economic implications of the usage of the SNJD models are also provided.

Appendix

Proof of Lemma 1

A previous step to compute the characteristic function is to calculate the log-return process. Dividing two realizations of the model (1.1.2.2) in times t and $t+\Delta t$, we obtain

$$\begin{aligned}
\frac{S_{t+\Delta t}}{S_t} &= \exp\left[\left(\mu-\frac{\sigma^2}{2}\right)\Delta t+\sigma\left(B_{t+\Delta t}-B_t\right)\right] \\
&\quad\times\frac{\prod_{j=1}^{N_{t+\Delta t}}\left[1+U_j h\left(t+\Delta t-\tau_j\right)\right]}{\prod_{j=1}^{N_t}\left[1+U_j h\left(t-\tau_j\right)\right]}\frac{\prod_{j=1}^{N_t}\left[1+U_j h\left(t+\Delta t-\tau_j\right)\right]}{\prod_{j=1}^{N_t}\left[1+U_j h\left(t+\Delta t-\tau_j\right)\right]} \\
&= \exp\left[\left(\mu-\frac{\sigma^2}{2}\right)\Delta t+\sigma\left(B_{t+\Delta t}-B_t\right)\right] \\
&\quad\times\prod_{j=N_t+1}^{N_{t+\Delta t}}\left[1+U_j h\left(t+\Delta t-\tau_j\right)\right]\times\prod_{j=1}^{N_t}\frac{\left[1+U_j h\left(t+\Delta t-\tau_j\right)\right]}{\left[1+U_j h\left(t-\tau_j\right)\right]}
\end{aligned}$$

Following this, the process for the log-returns ($Z_t \equiv \ln\left(S_{t+\Delta t}/S_t\right)$) can be written as the sum of three independent contributions: a white noise process X_t, the marked point part Y_t and a shot noise contribution H_t, with

$$X_t = \left(\mu-\frac{\sigma^2}{2}\right)\Delta t+\sigma\left(B_{t+\Delta t}-B_t\right) \tag{1.1.7.1}$$

$$Y_t = \sum_{j=N_t+1}^{N_{t+\Delta t}}\ln\left[1+U_j h\left(t+\Delta t-\tau_j\right)\right] \tag{1.1.7.2}$$

$$H_t = \sum_{j=1}^{N_t}\ln\left[\frac{1+U_j h\left(t+\Delta t-\tau_j\right)}{1+U_j h\left(t-\tau_j\right)}\right] \tag{1.1.7.3}$$

We will compute the joint characteristic function (CF) of the process (1.1.2.2) using (1.1.7.1)-(1.1.7.3). Using standard probability theory (see Feller, 1971), the CF of a sum of independent random variables is equal to the product of each individual CF's,

$$\begin{aligned}
Z_t &= X_t+Y_t+H_t \rightarrow \\
\Phi_{Z_t}(\xi) &= \Phi_{X_t}(\xi)\times\Phi_{Y_t}(\xi)\times\Phi_{H_t}(\xi)
\end{aligned} \tag{1.1.7.4}$$

where

- $\Phi_R(\xi)$ is the CF of the random variable R, defined by the Fourier transform

$$\Phi_R(\xi) = E\left[e^{i\xi R}\right] = \int e^{i\xi Z} dF(r)$$

- ξ is the transform variable and $F(r)$ denotes the cumulative distribution function (c.d.f.) of the random variable R.

As (1.1.7.1) is the sum of two contributions (deterministic and random, respectively) and $B_{t+\Delta t} - B_t \sim N(0, \Delta t)$, we obtain that

$$\Phi_{X_t}(\xi) = \exp\left(\left[i\left(\mu - \frac{\sigma^2}{2}\right)\xi - \frac{1}{2}\sigma^2\xi^2\right]\Delta t\right) \tag{1.1.7.5}$$

We will use the properties of the Poisson process to derive the CF of the components (1.1.7.2) and (1.1.7.3).[33] CF for both process are solved in the same manner, so we present here the calculations for (1.1.7.3). This procedure is analogous for (1.1.7.2). Consider a fixed time interval $[0, T]$. Let N_t be an homogeneous Poisson process with intensity λ and let m be the number of jumps in $[0, t]$, $N_t - N_0 \equiv N_t = m$, and τ_m denotes the jump times. As pointed out in Parzen (1962) or Klüppelberg and Mikosch (1995), the random vector $[\tau_1, \cdots, \tau_m]$ has the same distribution as the order statistics of a sample of m i.i.d. random variables uniformly distributed on $[0, t]$. Additionally, each of the $m!$ possible orderings has equal probability because of the independence between the random variables.

[33] The characteristic function for shot noise processes has been calculated in Parzen (1962) and Klüppelberg and Mikosch (1995). The main part of the proof used here has been borrowed from these two references.

Using H_t as given in (1.1.7.3), we obtain

$$\begin{aligned}
&\Phi_{H_t}(\xi) = E\left[e^{i\xi H_t}\right] = \\
&= E\left[\exp\left\{i\xi\sum_{j=1}^{N_t}\ln\left[\frac{1+U_j h(t+\Delta t-\tau_j)}{1+U_j h(t-\tau_j)}\right]\right\}|N_t=m\right]\times P(N_t=m) \\
&= \sum_{m=0}^{+\infty}\frac{(\lambda t)^m}{m!}e^{-\lambda t} \\
&\quad\times\frac{m!}{(t)^m}\int_0^t d\tau_1\int_{\tau_1}^t d\tau_2 ...\int_{\tau_{m-1}}^t d\tau_m \\
&\quad\times\prod_{j=1}^m E\left[\exp\left\{i\xi\ln\left(\frac{1+U_j h(t+\Delta t-\tau_j)}{1+U_j h(t-\tau_j)}\right)\right\}\right] \\
&= e^{-\lambda t}\sum_{m=0}^{\infty}\frac{(\lambda t)^m}{m!}\times\left[\frac{1}{t}\int_0^t E\left[\exp\left\{i\xi\ln\left(\frac{1+Uh(t+\Delta t-\tau)}{1+Uh(t-\tau)}\right)\right\}\right]d\tau\right]^m \\
&= e^{-\lambda t}e^{\lambda t\left[\frac{1}{t}\int_0^t E\left[\exp\left\{i\xi\ln\left(\frac{1+Uh(t+\Delta t-\tau)}{1+Uh(t-\tau)}\right)\right\}\right]d\tau\right]} \\
&= \exp\left\{\lambda\int_0^t E\left[\exp\left\{i\xi\ln\left(\frac{1+Uh(t+\Delta t-\tau)}{1+Uh(t-\tau)}\right)\right\}-1\right]d\tau\right\} \qquad (1.1.7.6)
\end{aligned}$$

Similarly, using (1.1.7.2) we obtain

$$\Phi_{Y_t}(\xi) = \exp\left\{\lambda\int_t^{t+\Delta t} E\left[\exp\left\{i\xi\ln\left(1+Uh(t+\Delta t-\tau)\right)\right\}-1\right]d\tau\right\} \qquad (1.1.7.7)$$

Finally, substituting (1.1.7.5)-(1.1.7.7) into (1.1.7.4) leads to the desired result.

Proof of Lemma 2

The first two non-central moments of the process (1.1.2.2) are computed using (1.1.3.4) and applying (1.1.3.5) for $n = 1, 2$.

$$\begin{aligned}
\mu_1' &= \left(\mu - \frac{\sigma^2}{2}\right)\Delta t + \lambda \int_t^{t+\Delta t} E\left[\ln\left(1 + Uh\left(t + \Delta t - \tau\right)\right)\right] d\tau \\
&\quad + \lambda \int_0^t E\left[\ln\left(\frac{1 + Uh\left(t + \Delta t - \tau\right)}{1 + Uh\left(t - \tau\right)}\right)\right] d\tau \\
\mu_2' &= \sigma^2 \Delta t + \lambda \int_t^{t+\Delta t} E\left[\ln^2\left(1 + Uh\left(t + \Delta t - \tau\right)\right)\right] d\tau \\
&\quad + \lambda \int_0^t E\left[\ln^2\left(\frac{1 + Uh\left(t + \Delta t - \tau\right)}{1 + Uh\left(t - \tau\right)}\right)\right] d\tau + \left(\mu_1'\right)^2
\end{aligned}$$

Computation of the third and fourth non-central moments could be cumbersome. For the ease of explanation, we adopt a more suitable notation.

The third and fourth non-central moments of the process (1.1.2.2) are

$$\begin{aligned}
\mu_3' &= C^3 + D^3 - \left[3\left(A + C + D\right)\left(2B - C^2 - D^2\right)\right] + \left(A + C + D\right)^3 \\
\mu_4' &= C^4 + D^4 + 4\left(A + C + D\right)\left(C^3 + D^3\right) \\
&\quad + 3\left(2B - C^2 - D^2\right)^2 - 5\left(2B - C^2 - D^2\right)\left(A + C + D\right) \\
&\quad - \left(2B - C^2 - D^2\right)\left(A + C + D\right)^2 + \left(A + C + D\right)^4
\end{aligned}$$

with

$$\begin{aligned}
A &= \left(\mu - \frac{\sigma^2}{2}\right)\Delta t \\
B &= -\frac{\sigma^2}{2}\Delta t \\
C^n &= \lambda \int_t^{t+\Delta t} E\left[\left\{\ln\left(1 + Uh\left(t + \Delta t - \tau\right)\right)\right\}^n\right] d\tau \\
D^n &= \lambda \int_0^t E\left[\left\{\ln\left(\frac{1 + Uh\left(t + \Delta t - \tau\right)}{1 + Uh\left(t - \tau\right)}\right)\right\}^n\right] d\tau
\end{aligned}$$

Using the relationship between central and non-central moments (see Abramowich and Stegun, 26.1.14), we can compute the first four non-central

moments of the process (1.1.2.2):

$$\begin{aligned} Mean &= \mu_1' = A + C + D \\ Variance &= \mu_2' - \left(\mu_1'\right)^2 = -2B + C^2 + D^2 \\ Skewness &= \frac{\mu_3}{\mu_2^{3/2}} = \frac{2\left(\mu_1'\right)^3 - 3\mu_1'\mu_2' + \mu_3'}{\mu_2^{3/2}} = \frac{1}{\mu_2^{3/2}}\left(C^3 + D^3\right) \\ Kurtosis &= \frac{\mu_4}{\mu_2^2} = \frac{-3\mu_1'^4 + 6\mu_1'^2\mu_2' - 4\mu_1'\mu_3' + \mu_4'}{\mu_2^2} = 3 + \frac{C^4 + D^4}{\mu_2^2} \end{aligned}$$

Proof of Lemma 3 and Example 3

To obtain the different approximations for expressions (1.1.3.6) and (1.1.3.7) we will focus on the logarithm terms of the integrals. By using the approximation of the logarithmic function $\ln(1+x) \simeq x + O\left(x^2\right)$, for $|x| < 1$ we have

$$\begin{aligned} C^n &= \lambda \int_t^{t+\Delta t} E[\{ln(1 + Uh(t + \Delta t - \tau))\}^n]d\tau \\ &\simeq \lambda E[U^n] \int_t^{t+\Delta t} [h(t + \Delta t - \tau)]^n d\tau \\ D^n &= \lambda \int_0^t E\left[\left\{ln\left(\frac{1 + Uh(t + \Delta t - \tau)}{1 + Uh(t - \tau)}\right)\right\}^n\right] d\tau \\ &\simeq \lambda E[U^n] \int_0^t [h(t + \Delta t - \tau) - h(t - \tau)]^n d\tau \\ &\simeq \lambda E[U^n](\Delta t)^n \int_0^t \left[\frac{d}{dt}h(t - \tau)\right]^n d\tau \end{aligned}$$

In the case of the exponential response function of example 3, applying

the power series for the exponential function $e^x \simeq 1 + x + O\left(x^2\right)$, we obtain

$$\begin{aligned} C^n &\simeq \lambda E\left[U^n\right] \int_t^{t+\Delta t} e^{-na(t+\Delta t-\tau)} d\tau = \lambda E\left[U^n\right] \frac{1}{na}\left(1 - e^{-na\Delta t}\right) \\ &\simeq \lambda E\left[U^n\right] \Delta t \\ D^n &\simeq \lambda E\left[U^n\right] (\Delta t)^n \int_0^t (-a)^n e^{-na(t-\tau)} d\tau \\ &= \lambda E\left[U^n\right] (-a\Delta t)^n \frac{1}{na}\left(1 - e^{-nat}\right) \end{aligned}$$

Proof of Lemma 4

Expanding the expression for log-returns Z_t of the model (1.1.2.2) in its three components X_t, Y_t and H_t (see (1.1.3.1)-(1.1.3.3)) helps to identify the possible sources of autocovariance of the process Z_t,

$$Cov\left[Z_{t+n\Delta t}, Z_t\right] = Cov[X_{t+n\Delta t} + Y_{t+n\Delta t} + H_{t+n\Delta t},\, X_t + Y_t + H_t]$$

with $n = 0, 1, 2, \ldots$ and so on.

As $\{B_t\}_t, \{N_t\}_t$, and $\{U_j\}_j$ are mutually independent, it is possible to isolate the covariance of the elements X_t of the remaining terms, due to the independence of the increments of the Brownian motion. By contrast, this turns different when it comes to study the linear relationship between the jump part Y_t and the shot noise part H_t. Following this, the autocovariance of Z_t can be written as

$$\begin{aligned} Cov\left[Z_{t+n\Delta t}, Z_t\right] &= Cov\left[X_{t+n\Delta t}, X_t\right] + Cov\left[Y_{t+n\Delta t}, Y_t\right] \\ &+ Cov\left[H_{t+n\Delta t}, H_t\right] + Cov\left[H_{t+n\Delta t}, Y_t\right] \quad (1.1.7.8) \end{aligned}$$

where $n = 0, 1, 2, \ldots$. We provide now the autocovariance of each term in (1.1.7.8) separately:

1. **Brownian part.** Due to the independence between the increments of the Brownian motion, the expression (1.1.3.1) has an autocovariance function of the form,

$$Cov\left[X_{t+n\Delta t}, X_t\right] = \sigma^2 \Delta t \delta\left(n\right) \quad (1.1.7.9)$$

where $\delta(x)$ is the usual Dirac function, that is

$$\delta(x) = \begin{cases} 1, & \text{if } x = 0 \\ 0, & \text{otherwise} \end{cases}$$

2. **Jump part.** Considering there is no effect of the response function in the moment of event happens (assumptions 1 and 2), and noting that the random variables U_j are i.i.d., the autocovariance function of Y_t turns easily computable,

$$Cov\left[Y_{t+n\Delta t}, Y_t\right] = \lambda E[U^2]\Delta t\delta(n) \qquad (1.1.7.10)$$

3. **Shot noise part.** For a formal proof of the covariance of the Shot noise, we refer the reader[34] to Parzen (1962), pp.147. Using the expression for H_t as given in (1.1.3.3), we have

$$Cov\left[H_{t+n\Delta t}, H_t\right] = \lambda \int_0^t E\left[\ln\left(\frac{1+Uh\left(t+\Delta t-\tau\right)}{1+Uh\left(t-\tau\right)}\right) \times \ln\left(\frac{1+Uh\left(t+\Delta t+n\Delta t-\tau\right)}{1+Uh\left(t+n\Delta t-\tau\right)}\right)\right] d\tau \qquad (1.1.7.11)$$

4. **Jump / Shot-noise part.** The persistence in time of the realization of the random variable U_j produces a cross effect among the Y_t and H_t processes, which contributes to the appearance of a new source of autocovariance in the process Z_t.

   Using assumptions 1 and 3, we display here a general expression for the covariance between the Jump and Shot-noise parts[35],

   $Cov\left[H_{t+n\Delta t}, Y_t\right] =$

$$= Cov\left[\sum_{j=N_t+1}^{N_{t+\Delta t}} U_j, \sum_{j=1}^{N_{t+n\Delta t}} U_j(h(t+n\Delta t-\tau_j)-h(t-\tau_j))\right]$$

$$= Cov\left[\sum_{j=N_t+1}^{N_{t+\Delta t}} U_j, \sum_{j=N_t+1}^{N_{t+n\Delta t}} U_j(h(t+n\Delta t-\tau_j)-h(t-\tau_j))\right] \qquad (1.1.7.12)$$

[34] Other alternative sources for computing the autocovariance of the Shot noise are Ross (1996) or Kluppelberg and Mikosch (1995).

[35] Weare not able to compute the general expression of the autocovariance in the Jump-Shot noise part as it differs from one to another specifications of the response function.

Finally, replacing equations (1.1.7.9)-(1.1.7.12) into (1.1.7.8) we obtain the desired expression.

Proof of Lemma 5

As we have previously mentioned, the expression of the log-returns for the model (1.1.2.2) can be written as the sum of two white noise terms plus a shot noise component. Again, we apply the Fourier transform to each term in (1.1.4.4):

1. **Brownian part.** Here, the Fourier transform of a delta function leads to a spectrum of the form

$$s_X(w) = \frac{\sigma^2}{2\pi}\Delta t$$

which is the usual spectrum of a white noise, constant for all the frequencies w.

2. **Jump part.** Former result equally applies to the jump part. For this case, we also have

$$s_Y(w) = \frac{\lambda}{2\pi}E[U^2]\Delta t$$

an spectrum where there is no dependence of frequencies w.

3. **Shot noise part.** Alternative definitions of function $h(\cdot)$ lead to different population spectrum. We just indicate the general form for the population spectrum of the shot noise,

$$\begin{aligned} s_H(w) &= \frac{\lambda}{2\pi}\int_{-\infty}^{+\infty} e^{-iw\varepsilon}d\varepsilon \times \\ &\int_0^t E\left[\ln\left(\frac{1+Uh(t+\Delta t-\tau)}{1+Uh(t-\tau)}\right)\ln\left(\frac{1+Uh(t+\Delta t+\varepsilon-\tau)}{1+Uh(t+\varepsilon-\tau)}\right)\right]d\tau \end{aligned}$$

4. **Jump-Shot noise part.** Similarly to the Shot noise part, we present the procedure to compute the spectrum of the Jump / Shot-noise part, since this must be calculated for any different specification of the response function $h(\cdot)$.

$$\begin{aligned} s_{HY}(w) &= \frac{\lambda}{2\pi}\int_{-\infty}^{+\infty} e^{-iw\varepsilon}d\varepsilon \\ &\quad \times Cov\left[\sum_{j=N_t+1}^{N_{t+\Delta t}} U_j, \sum_{j=N_t+1}^{N_{t+\varepsilon}} U_j(h(t+\varepsilon-\tau_j)-h(t-\tau_j))\right] \end{aligned}$$

Finally, the population spectrum of the log-returns $s_Z(w)$ is given by the sum of the four contributions, that is

$$s_Z(w) = s_X(w) + s_Y(w) + s_H(w) + s_{HY}(w)$$

Proof of Example 4

As the Brownian and jump parts of the autocovariance are immediately computed, we just focus on the Shot noise and Jump / Shot-noise terms.

1. **Shot noise part.** Replacing $h(t) = e^{-at}$ into (1.1.7.11), we obtain

$$\begin{aligned} &Cov\left[H_{t+n\Delta t}, H_t\right] \\ &= \lambda\int_0^t E\left[\ln\left(\frac{1+Ue^{-a(t+\Delta t-\tau)}}{1+Ue^{-a(t-\tau)}}\right)\ln\left(\frac{1+Ue^{-a(t+(n+1)\Delta t-\tau)}}{1+Ue^{-a(t+n\Delta t-\tau)}}\right)\right]d\tau \end{aligned}$$

Applying the Taylor expansion for the logarithmic function, this autocovariance can be approximated as follows:

$$Cov\left[H_{t+n\Delta t}, H_t\right] \simeq \lambda E\left[U^2\right](e^{-a\Delta t}-1)^2\frac{1}{2a}\left(1-e^{-2at}\right)e^{-an\Delta t}$$

Recognizing that the terms Δt is "close enough" to zero, we get that

$$Cov\left[H_{t+n\Delta t}, H_t\right] \simeq \lambda E\left[U^2\right]\frac{a}{2}\left(1-e^{-2at}\right)(\Delta t)^2 e^{-an\Delta t}$$

Note that this autocovariance vanishes when $t \to \infty$ as the term $(\Delta t)^2$ is almost zero.

2. **Jump / Shot-noise part.** As noted previously, the persistence in time of the realizations of the random variable U_j produce a cross effect between the Y_t and H_t processes, which contributes to the appearance of a new source of autocovariance in the process Z_t.

   Replacing $h(t) = e^{-at}$ into (1.1.7.12), we obtain

$$Cov\left[H_{t+n\Delta t}, Y_t\right] = \left(e^{-a\Delta t} - 1\right) Cov\left[\sum_{j=N_t+1}^{N_{t+\Delta t}} U_j, \sum_{j=N_t+1}^{N_{t+n\Delta t}} U_j \left(e^{-a(t+n\Delta t-\tau_j)}\right)\right]$$

   For $n = 1$, we get

$$\begin{aligned} Cov\left[H_{t+\Delta t}, Y_t\right] &= \left(e^{-a\Delta t} - 1\right) Cov\left[\sum_{j=N_t+1}^{N_{t+\Delta t}} U_j, \sum_{j=N_t+1}^{N_{t+\Delta t}} U_j \left(e^{-a(t+\Delta t-\tau_j)}\right)\right] \\ &= \left(e^{-a\Delta t} - 1\right) \lambda E\left[U^2\right] \Delta t e^{-a\Delta t} \\ &= \left(e^{-a\Delta t} - 1\right) Var(Y_t) e^{-a\Delta t} \end{aligned}$$

   Similarly, for $n = 2$, we obtain

$$Cov\left[H_{t+2\Delta t}, Y_t\right] = \left(e^{-a\Delta t} - 1\right) Cov\left[\sum_{j=N_t+1}^{N_{t+\Delta t}} U_j, \sum_{j=N_t+1}^{N_{t+2\Delta t}} U_j \left(e^{-a(t+2\Delta t-\tau_j)}\right)\right]$$

   Applying the linearity of the covariance function, the covariance term on the right hand side of the equation can be split as

$$\underbrace{Cov\left[\sum_{j=N_t+1}^{N_{t+\Delta t}} U_j, \sum_{j=N_t+1}^{N_{t+\Delta t}} U_j \left(e^{-a(t+2\Delta t-\tau_j)}\right)\right]}_{=Var(Y_t)\, e^{-2a\Delta t}}$$
$$+ \underbrace{Cov\left[\sum_{j=N_t+1}^{N_{t+\Delta t}} U_j, \sum_{j=N_t+\Delta t+1}^{N_{t+2\Delta t}} U_j \left(e^{-a(t+2\Delta t-\tau_j)}\right)\right]}_{=0}$$

   where the second equality is a consequence of the independence property among Poisson increments.

So, finally,

$$Cov\left[H_{t+2\Delta t}, Y_t\right] = (e^{-a\Delta t} - 1)Var(Y_t)\, e^{-2a\Delta t}$$

In a similar way to that for $n = 2$, we iterate until we get

$$Cov\left[H_{t+n\Delta t}, Y_t\right] = (e^{-a\Delta t} - 1)\, Var(Y_t)\, e^{-a\, n\Delta t}$$

Proof of Example 6

Changing the notation for the autocovariance in Example 4 leads to

$$\begin{aligned} Cov\left[Z_{t+\varepsilon}, Z_t\right] &= \sigma^2 \Delta t \delta(\varepsilon) + \lambda E\left[U^2\right] \Delta t \delta(\varepsilon) \\ &+ (e^{-a\Delta t} - 1)\, \lambda E\left[U^2\right] \Delta t\, e^{-a\,\varepsilon} \end{aligned}$$

Now, applying the Fourier transform on the previous equation we get

$$s_Z(w) = \frac{1}{2\pi}\left(\sigma^2 + \lambda E[U^2] \; + \lambda(e^{-a\Delta t} - 1)E\left[U^2\right] \frac{2a}{a^2 + w^2}\right)\Delta t$$

which completes the proof.

Chapter 2

Estimation of the SNJD process in the Time Domain

This chapter[1] is devoted to the estimation in the time domain of the Shot-Noise Jump-Diffusion process. We discuss several proposals that have already been done in the literature for estimating jump-diffusion process in the line of that posited by Merton (1976). Then, we analyze the estimation of the parameters in the process (1.1.2.2) by means of a Generalized Method of Moments (GMM) estimate proposed in Hansen (1982)[2].

2.1 Introduction

In 1976, Merton (1976) proposed a jump-diffusion process where the returns are approximately distributed as a mixture of normal distributions conditional on the increments of a random variable that follows a Poisson distribution. This class of models allows us to capture the high kurtosis observed in the distribution of daily asset returns.

[1]I am sincerely grateful to W. Stute and M. Moreno for their contributions on this chapter. I gratefully acknowledge helpful comments and suggestions from A. Novales and A. León. Finally, I also acknowledge financial support from the Plan Nacional de I+D+I (project BEC2003-02084) from the Spanish Government and project GIU 06/53 of the University of the Basque Country and Basque Government, and specially to José M. Usategui. Previous drafts of this chapter have been presented under the name "Calibrating Shot Noise Processes" in the 2004 EuroWorking Group on Financial Modelling and the 8th Italian-Spanish Meeting on Financial Mathematics. I have also benefited from the comments of participants in seminars at University of Basque Country, Universidad Complutense de Madrid, Universitat de les Illes Balears, and Universidad Carlos III de Madrid.

[2]Notation used in this chapter has been taken from Hamilton (1994).

In addition, the specification of this jump-diffusion process keeps certain economic intuition: according to this model, the total change in the asset price is due to the joint effect of two types of changes: a) the *normal* ones, due to the bid-ask crossing of common information, and b) the *abnormal* ones, where the impact of the information on the price produces a non-marginal change in the asset price.

As a result of the Merton (1976) model, an increasing interest in the estimation of this type of diffusion process in the financial literature has been observed. We refer here some different approaches for estimating the parameters of a jump diffusion process. Basically, to our knowledge, the main attempts can be summarized as follows:

1. The method of cumulants, used by Press (1967), Beckers (1981) or Ball and Torous (1983)

2. Maximum Likelihood (ML), as those papers of Ball and Torous (1983, 1985), Jorion (1988) or Honoré (1998).

3. Generalized Method of Moments (GMM), as the works of Das and Sundaram (1999) or Aït-Sahalia (2004).

4. Empirical Characteristic Function (ECF), posited in Duffie, Pan and Singleton (2000), Singleton (2001) or Carrasco and Florens (2000), among others.

5. The simulation-based techniques, as the Efficient Moment Method (EMM) of Gallant and Tauchen (1996) or the Simulated Method of Moments (SMM) of Duffie and Singleton (1993) or Gorieroux and Monfort (1996).

Press (1967) and Beckers (1981) consider the parameters of a jump diffusion process using the method of the cumulants, a power series expansion of the characteristic function logarithm of the process (See Abramowitz and Stegun (1972) for details). However, both papers obtain negative estimations of the variance of the process and the jump for some assets.

Maybe, the ML procedure constitutes the most direct approach to estimate a JD process. Conditioning on the number of jumps that can happen, the transition probability density of the process can be easily derived, as shown in Ball and Torous (1983) or Aït-Sahalia (2004), among others.

Under some assumptions, Ball and Torous (1983, 1985) demonstrate that a JD process could be approximated by a mixture of normally distributed variables: Ball and Torous (1983) propose a Bernoulli distribution as an

approach to the Poisson process that models the arrival of information, estimating the parameters of a JD process by using the method of the cumulants and ML techniques.

Additionally, Ball and Torous (1985) and Jorion (1988) obtain the ML estimates of a JD process by maximizing the value of a ML truncated function. However, the estimation of Ball and Torous (1985) is not consistent with some of the assumptions indicated in their paper, as was pointed out by Navas (2003). On the other hand, Honoré (1998) also encounters some numerical problems in the ML estimations of Merton (1976) type processes mainly because of the likelihood function is unbounded.

This problem is partially solved in Honoré (1998) by restricting the variance of the normal distribution with a constant. An alternative solution[3] is also provided by Hamilton (1994) by means of using the EM algorithm of Dempster *et al* (1977). Finally, the estimation by ML of more complex JD process that incorporate ARCH-type effects could be found on Jorion (1988) in a foreign exchange rates context, or Das (2002) or Benito *et al* (2007) within an interest rate environment.

Even though the GMM technique proposed by Hansen (1982) has received considerably attention in the estimation of continuous-path diffusion processes[4], the application to the jump-diffusion processes area is not as considerably as in the diffusion cases.

Anyway, it is clear that the GMM approach can provide an alternative way to the ML estimation of the parameters of a JD process. A classical reference on this matter is Das and Sundaram (1999), who compute the characteristic function of a JD process and provide the expressions for the first few moments of the process.

More recently, Aït-Sahalia (2004) has extended the expression obtained in Das and Sundaram (1999) to the moments of non-integer order. He also proves that it is possible to disentangle the jumps from the underlying Brownian noise, without addressing what happens with the remaining parameters of the jump. For an empirical application of the GMM approach we refer the reader to Moreno and Peña (1996), who use a GMM estimate inspired on Aase and Guttorp (1987) for detecting jumps in the Spanish short-term interbanking interest rates.

Closely linked to the GMM approach, a recent procedure is based on the Empirical Characteristic Function (ECF) has been used to estimate diffusion

[3]Hamilton (1994) formulates the problem of the Gaussian mixture estimation in a context of switching regime models. We refer the reader to Hamilton (1994) and references therein for a general discussion of the analysis of i.i.d. mixture distributions.

[4]See Chan *et al* (1992), Zhou (2001) or Duffee and Stanton (2004), among many others.

processes[5]. This method does not require the discretization of the process and it uses, in an efficient sense, all the information included in the sample. The basic idea for the estimation of the ECF is the minimization of various distance measures between the ECF and the Characteristic Function (CF) of the process. Unfortunately, as discussed in Yu (2004), this methodology is computationally burdensome in the presence of serial dependence.

During recent years, an increasing amount of literature has applied the simulation-based econometric methods to the estimation of JD processes. For instance, Andersen *et al* (2002) use the EMM approach of Gallant and Tauchen (1996) for estimating the behavior of equity-index returns under some different diffusion processes, including a JD process with time-varying intensity. A second interesting reference is Jiang (1998) who estimates a JD model for exchange rates using the Indirect Inference method of Gouriéroux *et al* (1993).

As many of the above studies subordinate the estimation procedure to the option pricing performance of the JD models, it might be convenient to mention a general treatment of this point in the jump-diffusion literature. Basically, the first reference to this point is Merton (1976), who prices a European call option assuming a jump-diffusion process for the underlying asset. Considering this process, the market is incomplete and, therefore, the payments of the derivative asset can not be replicated. In this set up, Merton obtains a closed-form expression for the price of the European call option assuming that the jump risk can be diversified.

Likewise, Ball and Torous (1985), Jorion (1988) and Amin (1993) use the Merton model for pricing derivatives, although the variance used has been recently put in doubt when one compares the prices of those options with those obtained from the Black-Sholes model[6]. Additionally, Amin (1993) develops a discretized option pricing model inspired in the model posited by Cox *et al* (1979). Finally, greater degrees of sophistication in derivative pricing with jump processes can be found in Scott (1997), Andersen and Andreasen (2000) or Andersen *et al* (2002), among many others.

Other studies have extended the analysis of jump models to other markets, as foreign exchange or electricity ones. A classical reference of the former is Jorion (1988), where the empirical distribution of the returns for the stock exchange market is analyzed using a variant without restrictions on the density function used by Press (1967). Regarding electricity markets,

[5]See, for instance, Feuerverger and McDunnough (1981a), Singleton (2001), Jiang and Knight (2002), Chacko and Viceira (2003) or Yu (2004), among others.

[6]See, for instance, Das and Sundaram (1999) and Navas (2003).

Escribano *et al* (2002) provide an empirical study across different countries for a time-dependent JD process.

The structure of this chapter is as follows. Section 2 reviews the GMM methodology and its properties. Section 3 provides some details about the estimation and Section 4 develops a preliminary study to estimate the SNJD parameters by using restricted versions of the model. Sections 5 and 6 carry out an exhaustive Monte Carlo analysis of the SNJD process. Finally, Section 7 summarizes the main conclusions obtained.

2.2 The GMM estimate

This section surveys the GMM approach and is mainly based on Hansen (1982) and Hamilton (1994). We explain the intuition behind the GMM estimate, setting up the mathematical notation employed during this chapter.

2.2.1 General setting

The GMM methodology is based on the asymptotic relationship between the population moments (μ_i) and their sample counterparts ($\hat{\mu}_i$), where the term i-denotes the moment order. The intuition behind is simple: given a large sample size, it is expected that sample moments tend to be close to population moments, it is said $\widehat{\mu}_i \xrightarrow{p} \mu_i$. Basically, the main idea of this method consists on finding the parameters that match, as close as possible, the population moments with the sample moments. This matching is obtained by minimizing the quadratic form

$$J_T = \min_{\theta} \left[g\left(\theta\right)\right]' W_T \left[g\left(\theta\right)\right] \tag{2.2.2.1}$$

where $g\left(\theta\right)$ is a vector function of dimension $r \times 1$ which contains the set of sample and population moments, and consequently, it also contains the unknown parameters vector, θ, that we are looking for. The number r is known as the number of *moment conditions* and plays a fundamental role in the GMM estimate, as we will refer. Finally, W_T is called the *weighting matrix*. The formal definitions of $g\left(\theta\right)$ and W_T are given below[7].

The mathematical setup of the GMM framework is as follows: Let y_t denote a random variable, and consider $(y_1, y_2, ..., y_T)$, a set of i.i.d. observations of y_t. Assuming that y_t has a known distribution $f_{y_t}\left(\theta\right)$ and moments up to order n, define θ as the parameter vector that determines

[7] The subindex T denotes the sample size. As we will see, it is included in the definition of the estimate J_T.

completely the distribution f_{y_t}. The i-th population moments μ_i of y_t are given as

$$E\left[y_t^i\right] = \mu_i(\theta), \text{ for } i = 1, 2, \cdots, n$$

On the other hand, the sample moment estimate, $\hat{\mu}_i$, is given by

$$\widehat{\mu}_i\left(\widehat{\theta}_T\right) = \frac{1}{T}\sum_{t=1}^{T} y_t^i$$

As was previously cited, the idea consists on minimizing the distance between the population and sample moments, given by

$$\mu_i(\theta) - \widehat{\mu}_i\left(\widehat{\theta}_T\right) = 0$$

or, to put it another way more conveniently, minimize the error term[8] $h(\theta_0)$

$$h(\theta_0) = \mu_i(\theta) - y_t^i$$

where θ_0 denotes the true value of θ, whose mean should be zero (see Cochrane, 2005). This last point is obtained by defining $g(\theta)$ as the sample mean of the errors, that is,

$$g(\theta) \equiv \frac{1}{T}\sum_{t=1}^{T} h(\theta_0) = E\left[h(\theta_0)\right] = 0 \tag{2.2.2.2}$$

To keep intuition in equation (2.2.2.1), consider the weighting matrix $W_T = I$, with I the identity matrix. In this case, we are minimizing the sum of quadratic errors between the observed sample moments and their population counterparts. Here, the estimate J_T is known as the *classical method of moments*(Hamilton, 1994), and is consistent and asymptotically normal distributed (Cochrane, 2005). Similarly to this author, J_T will be called *first stage* estimate.

2.2.2 The weighting matrix

Hansen (1982) provides an expression for the weighting matrix W_T in (2.2.2.1) that is efficient, in the sense that the minimum asymptotic variance for the GMM estimate $\widehat{\theta}_T$ is obtained when

$$J_T = \min_{\theta}\left[g(\theta)\right]' S_T^{-1}\left[g(\theta)\right] \tag{2.2.2.3}$$

[8]More precisely, the error term is defined as $h(\theta, w_t)$ (Hamilton, 1994), where w_t is an $(h \times 1)$ vector of variables that are observed at date t . We have focused here on the univariate extension of the problem, so the term w_t has been dropped out for the ease of explanation.

This is obtained replacing W_T by S_T^{-1}, with S_T a variance-covariance matrix of the form

$$S_T = \frac{1}{T}\sum_{t=1}^{T}\left[h\left(\theta_0\right)\right]\left[h\left(\theta_0\right)\right]' \tag{2.2.2.4}$$

Now, the estimate J_T in equation (2.2.2.3) is a consistent, asymptotically normal, and asymptotically efficient estimate (Cochrane, 2005) of the parameter vector (θ), as it was demonstrated in Hansen (1982).

In the presence of serial correlation, Hamilton (1994) suggests to use the Newey-West (1987) estimate as an alternative to S_T,

$$\widehat{S}_T = \widehat{\Gamma}_{0,T} + \sum_{v=1}^{q}\left[1-\left(\frac{v}{(q+1)}\right)\right]\left(\widehat{\Gamma}_{v,T}+\widehat{\Gamma}'_{v,T}\right) \tag{2.2.2.5}$$

with

$$\widehat{\Gamma}_{v,T} = \frac{1}{T}\sum_{t=v+1}^{T}\left[h\left(\widehat{\theta},w_t\right)\right]\left[h\left(\widehat{\theta},w_{t-v}\right)\right]'$$

but other alternatives for the estimates are also possible[9].

2.3 General setting

This section provides some details about the estimation. Firstly, we recall the model under study and its parameters of interest. Additionally, we also provide the main moments used in the GMM study. Finally, we describe some simulation features.

2.3.1 Model description

The model under study is a SNJD model of the form specified in equation (1.1.2.2) in Chapter 1. For analytical and intuition purposes, we have considered an exponential decreasing function[10] of the form $h\left(t\right)=e^{-at}$ in our estimations. The advantages of this specification are the following:

1. The JD process can be nested easily within a SNJD one, just by assuming that $a=0$.

2. The results obtained in this chapter can be related to the existing literature on JD processes.

[9] See Hamilton (1994) or Cliff (2003) and references therein.

[10] To the best of our knowledge, a functional form for the response function $h(t)$ has not yet been reported in the financial literature.

3. The performance of SNJD and JD processes can be compared easily to that of simulated samples[11].

Thus, our model under study results

$$S_t = S_0 \exp\left[\left(\mu - \frac{\sigma^2}{2}\right)t + \sigma B_t\right] \prod_{j=1}^{N_t}\left[1 + U_j e^{-a(t-\tau_j)}\right] \tag{2.2.3.1}$$

where it is assumed that $\{B_t\}_t, \{N_t\}_t$, and $\{U_j\}_j$ are mutually independent.

For simplicity, we have assumed here that the jump size corresponds to a realization of a lognormal random variable U of the form $U = e^{-\beta^2/2+\beta\varepsilon} - 1$, with $\varepsilon \sim N(0,1)$. The assumption of a lognormal distribution for the jump size is common in the jump-diffusion literature (see, for instance, Merton (1976)).

Without loss of generality, we have imposed that jumps have zero mean ($E[U] = 0$), which simplifies the number of parameters to estimate. This is a common assumption also done by Merton (1976) or Ait-Sahalia (2004), among others.

What it follows is the set of main moment conditions used through this dissertation:

$$E[Z_t] = \left(\mu - \frac{\sigma^2}{2}\right)\Delta t \tag{2.2.3.2}$$

$$Var[Z_t] = \left(\sigma^2 + \lambda E[U^2]\right)\Delta t \tag{2.2.3.3}$$

$$Skewness = \lambda \frac{E[U^3]}{(Var[Z_t])^{3/2}}\Delta t \tag{2.2.3.4}$$

$$Kurtosis = \lambda \frac{E[U^4]}{(Var[Z_t])^{2}}\Delta t \tag{2.2.3.5}$$

$$Corr[Z_{t+n\Delta t}, Z_t] = (e^{-a\Delta t} - 1)\lambda \frac{E[U^2]}{\sigma^2 + \lambda E[U^2]}\, e^{-a\, n\Delta t}, \text{ if } n \geq 1 \tag{2.2.3.6}$$

Equations (2.2.3.2)-(2.2.3.5) are usually standard in any GMM estimation[12]. Moreover, due to the capability of SNJD process to generate serial dependence, it can be interesting to include the information contained in cross moments by means of the correlation function included in equation (2.2.3.6).

[11] A SNJD model with exponential decaying function as that assumed here leads to the Merton (1976) model by doing the parameter a equal to zero.

[12] Notice the fact that equations (2.2.3.2) to (2.2.3.5) are *equal* for JD and SNJD process under the exponential response specification of the SNJD process, which underline the closeness among the JD and exponential response SNJD process, as it was previously mentioned in Chapter 1.

Taking everything into account, the parameter vector of log-returns in (2.2.3.1) results $\Lambda = (\mu, \sigma, \lambda, \beta, a)' \in R^5$, where μ and σ are the drift and the volatility of the Brownian motion, respectively; λ is the intensity of the Poisson process and measures the probability of an extreme event happens; β is related to the mean and variance of the jump size - its amplitude -, and the parameter a refers to the speed of the decaying effect after a jump event.

2.3.2 Simulation details

We detail here some aspects of the simulations that will be carried out. Mainly, each experiment involves the simulations of 3,000 sample paths. Each path is composed by 1,800 data. Assuming that the frequency of observation is daily, these results correspond to about seven years of observations. The assumption of daily returns is common in the jumps literature[13]. All the parameters are computed on a daily basis.

We are interested in comparing the performance of the GMM methodology for SNJD and JD models so two sets of simulations have been done. Since the four first moments in equations (2.2.3.2)-(2.2.3.5) are equal in both models under study, they have been taken as the first few orthogonality conditions.

Concerning to the JD case, the number of moment conditions has also been fixed to four, matching the set of parameters to estimate, $\Lambda_{\text{JD}} = (\mu, \sigma, \lambda, \beta)' \in R^4$, which results on that the equation system (2.2.2.2) is completely specified. With respect to the SNJD model, the number of moment conditions has been fixed to six: the four previous ones plus two additional conditions referring the first and second autocorrelation coefficients. In this case, the number of moment conditions exceeds the number of parameters, $\Lambda_{\text{SNJD}} = (\mu, \sigma, \lambda, \beta, a)' \in R^5$, so the system (2.2.2.2) is overidentified.

Some objections could be done to the number of moment conditions used. Our procedure follows the ideas of Jiang and Knight (2002), who suggest to select lower order moments, mainly based on two reasons:

1. The trade-off between the estimation of the weighting matrix and the deterioration of this estimate in finite samples due to the inclusion of new moments, and

2. The erratic finite sample behave of very high moments because the

[13] See, for instance, Ball and Torous (1983, 1985) or Jorion (1988), among others. Additionally, Ait-Sahalia (2004) provides interesting intuitions about the data frequency in jump-diffusion estimation.

presence of fat tails in the distribution[14].

We are inclined to believe that increasing the number of conditions on the JD case falls within those circumstances specified by Jiang and Knight (2002). By contrast, the addition of the cross moments in the SNJD model could result on an addition of useful information to the GMM estimate that must be considered, as it was previously noticed on Chapter 1.

We provide now some details about the optimization procedure. The GMM estimate has been computed by using a non-least squared routine, the LSQNONLIN function, of the MatLab Optimization Toolboox library[15]. The iteration procedure of GMM follows the methodology proposed in Hamilton (1994). We decide to stop this iteration procedure once the difference between two successive estimated sets of parameter values is smaller than 10^{-5} or convergence is achieved. The initial parameters have been chosen by realizations of a uniform distributed random variable vector whose bounds are those posited for the optimizer. Table 2.1 presents an standard estimation case and some additional information of the options fixed.

Regarding the set of parameter values used, our simulations are based on keeping constant some variables μ, σ, λ and β, and just modifying the parameter a, which controls the degree of serial dependence. Basically, the values fixed - on a daily basis - correspond to the following:

- Yearly return of 5% ($\mu = 0.0002$).
- Volatility of the Brownian part of 20% ($\sigma = 0.0126$).
- (Average) arrival rate of jumps per year is fixed to 10 ($\lambda = 0.0400$).
- Finally, a parameter $\beta = 0.1000$ makes possible - about 10 times per year, on average - changes on prices less or equal than $\pm 7\%$ the 65% of the times that a jump occurs.

Taking all these parameters into consideration, this leads to a total volatility of the simulated processes of about 30% per year.

Additionally, it is important to emphasize that it could be interesting to analyze the performance of our different estimates by means of changes in the simulation parameters μ, σ, λ and β. However, we have not performed this "robustness" analysis because our *main* intention is checking the availability

[14] Hansen (1982) provides a way of testing the overidentifying restrictions. However, this test can fail to detect a misspecified model, as it is pointed out in Hamilton (1994).

[15] All the computations have been carried on a PC 2.0 MHz Core 2 Duo Intel Centrino processor, with 1Gb of RAM.

of estimating the parameter a, not yet reported in the financial literature. As a second reason, for the sake of brevity, we prefer to extract some few conclusions from a bounded set of parameters better than being lost in a indistinguishable set of numbers. Anyway, we are aware of leaving unanswered a really important issue and it must be tackled in further research.

2.4 Preliminary check

Prior to the estimation of the parameters of a SNJD process (the central objective of this chapter) it can be of interest to check the ability of the GMM procedure to estimate the parameters of a restricted version of the log-returns in the model (2.2.3.1). Basically, the purpose of this previous study is twofold: firstly, it presents us a first contact with the estimation problem; secondly, it serves to check the availability of the proposed procedure in the simplest case.

Then, this section is devoted to analyze the behavior of the GMM estimates under (only) Shot noise samples, that is, samples that have not been perturbed by a Brownian part. Moreover, we will also present an additional question concerning the numerical approximations involved in deriving the moment conditions of the model (2.2.3.1).

2.4.1 The restricted SNJD model

We are interested on estimating the parameters of log-returns in a restricted version of the SNJD model in equation (2.2.3.1) where the parameters μ and σ are equal to zero. By direct substitution of expressions (1.1.3.1-1.1.3.3) in Chapter 1, our model under study results

$$H_t^r = \sum_{j=N_t+1}^{N_{t+\Delta t}} \ln[1 + U_j e^{-a(t+\Delta t-\tau_j)}] + \sum_{j=1}^{N_t} \ln\left[\frac{1 + U_j e^{-a(t+\Delta t-\tau_j)}}{1 + U_j e^{-a(t-\tau_j)}}\right] \quad (2.2.4.1)$$

where, H_t^r denotes the restricted version of the SNJD model ($\mu = \sigma = 0$) and, as usual, we assume that $\{N_t\}_t$ and $\{U_j\}_j$ are mutually independent.

One important feature of expression (2.2.4.1) is that some numerical approximations have been introduced to compute easily their moment expressions[16]. The effect of these approximations on the estimations is an unavoidable interrogate. To deal with this issue, our strategy is based on

[16] A complete discussion of this point can be found in Chapter 1.

Table 2.1: Optimization details for a standard estimation case.

I.- Optimization bounds

Parameters	μ $(\times 10^{-2})$	σ $(\times 10^{-2})$	λ $(\times 10^{-2})$	β $(\times 10^{-2})$	a
Lower	-0.24	0.13	0.04	1.00	-0.50
Upper	0.20	2.50	40.00	20.00	2.00

II.- Optimization options

Objective function tolerance	10^{-5}	Iterations	2000
Parameters tolerance	10^{-5}	MFE[a]	4000

[a] Maximum Function Evaluations

Optimization details for an standard estimation case of the JD and SNJD models. Panel I refers to the simulation, upper and lower bounds parameters. Panel II displays the tolerances, iterations and the maximum number of function evaluations allowed. The initial points of the optimization routine are realizations from an uniform random variable within the fixed lower and upper bounds. All experiments have been carried out assuming a jump size U lognormally distributed of the form $U = e^{-\beta^2/2+\beta\varepsilon} - 1$, with $\varepsilon \sim N(0,1)$ and mean $E[U] = 0$. The response function of the SNJD model is $h(t-\tau_k) = e^{-a(t-\tau_k)}$.

looking for a simple, manageable stochastic process whose statistical features - moments, autocorrelation function, etc - are as close as possible to those of equation (2.2.4.1). Then, we use the same econometric approach for estimating the parameters involved in both processes. Finally, we look at the results of this manageable stochastic process obtained from both models to analyze the differences between them.

Of course, analyzing this idea, it could be objected that the processes under study are not *directly* comparable as they are not strictly equal. However, we are looking for a general intuition about the estimation procedure.

Basically, we are wondering about the adequacy of the GMM estimates in the Shot noise samples, and try to understand how long is the influence on the approximations taken on the estimations: whether we find substantial differences in the behavior of the estimations can indicate that "something is not going well".

To clarify the former ideas, let us propose a stochastic process of the form

$$H_t^c = \sum_{i=1}^{N_t} U_i e^{-a(t-\tau)} \tag{2.2.4.2}$$

where, H_t^c denotes the *classical* Shot noise model[17] and, again, $\{N_t\}_t$ and $\{U_j\}_j$ are mutually independent.

The process in equation (2.2.4.2) will be denoted as the *classical* Shot noise process, although this terminology is not standard. This process is studied in Rice (1954) or Ross (1996), and some of its statistical properties are also provided there.

Regardless of the fact that equations (2.2.4.1) and (2.2.4.2) are different, a careful study reveals that they are closer each other than seems at a first moment.

Computing the Taylor expansion of the logarithm in expression (2.2.4.1),

[17]This process is defined in Ross (1996). See Chapter 1 for more details.

we get

$$\begin{aligned} H_t^r &= \sum_{j=N_t+1}^{N_{t+\Delta t}} \ln[1+U_j e^{-a(t+\Delta t-\tau_j)}] + \sum_{j=1}^{N_t} \ln\left[\frac{1+U_j e^{-a(t+\Delta t-\tau_j)}}{1+U_j e^{-a(t-\tau_j)}}\right] \\ &\simeq \sum_{j=N_t+1}^{N_{t+\Delta t}} U_j\, e^{-a(t+\Delta t-\tau_j)}] + (e^{-a\Delta t}-1)\sum_{j=1}^{N_t} U_j\, e^{-a(t-\tau_j)} \\ &\simeq \sum_{j=N_t+1}^{N_{t+\Delta t}} U_j + (e^{-a\Delta t}-1)\sum_{j=1}^{N_t} U_j\, e^{-a(t-\tau_j)} \end{aligned} \tag{2.2.4.3}$$

where the last expression is obtained assuming that just one jump can appear with probability $\lambda\Delta t$ within the time interval $[t, t+\Delta t]$, for a "small enough" increment Δt.

As deduced from equation (2.2.4.3), the expressions H_t^c and H_t^r seem to share an equal term[18]. At light of the last expression, we are inclined to think that the processes H_t^c and H_t^r could share some features. Two main reasons can be highlighted[19]:

1. The contribution of the random variable U_j to the H_t^r process just appears in certain moments of the life of the process.
2. The second term $\sum_{j=1}^{N_t} U_j\, e^{-a(t-\tau_j)}$ extends their influence over the entire time span.

Additionally, the variances for both processes are

$$Var(H_t^r) = \lambda E\left[U^2\right]\Delta t \tag{2.2.4.4}$$

$$Var(H_t^c) = \lambda E\left[U^2\right]\frac{1}{2a}(1-e^{-2at}) \tag{2.2.4.5}$$

and the expressions for the autocorrelation coefficients ρ_n result

$$\rho_n(H_t^r) = \begin{cases} 1, & \text{if } n=0 \\ (e^{-a\Delta t}-1)\, e^{-a\, n\Delta t}, & \text{if } n \geq 1 \end{cases} \tag{2.2.4.6}$$

$$\rho_n(H_t^c) = e^{-an\Delta t} \tag{2.2.4.7}$$

Roughly speaking, the main difference between these two process seems to be the form that their respective moment conditions have been derived.

[18] Under the assumptions taken.

[19] Both arguments support our idea that the joint study of the processes H_t^c and H_t^r could be relevant when it comes to study the performance of the GMM methodology.

In the case of the SNJD process, equations (2.2.4.4) and (2.2.4.6) have been obtained by means of approximations (see Chapter 1). For the case of the classical Shot noise, equations (2.2.4.5) and (2.2.4.7) can be derived directly (see, for instance, Ross, 1996).

At light of the former results, it could be thought that the classic Shot noise process of equation (2.2.4.2) can provide us a naive tool for comparing the performance of the methodology under approximations. Again, we realize that the moment expressions of the processes (2.2.4.1) and (2.2.4.2) are not strictly equal, so these processes can not be comparable.

According to the specification that we are adopting in equations (2.2.4.1)-(2.2.4.2), in both cases, our parameter vector results $\Lambda = (\lambda, \beta, a)' \in R^3$, where λ is the intensity of the Poisson processes, β is related to the amplitude of the jump and the parameter a refers to the speed of decaying after a jump[20].

2.4.2 GMM estimates and results

As previously mentioned, our experiment consists on the simulation of 3,000 sample paths with 1,800 data each one for the restricted SNJD process ($\mu = \sigma = 0$) of equation (2.2.4.1) and the classical Shot noise of expression (2.2.4.2). To analyze the degree of serial dependence in data, this procedure has been repeated four times, just changing the value of the parameter a to 0.2, 0.6,1.0 and 10.0. This generates four subsets of simulated data, attending to the different a parameters' value.

Secondly, we have estimated the parameters of each data subset by using the identity matrix as weighting matrix (our first stage estimate). Moreover, we also provide the results by taking into account some effects in the sample as serial correlation which, obviously, affects to the efficiency of the estimations, by means of the Newey-West (1987) estimate (our third stage estimate).

Basically, the usage of these two estimates responds to our intention of examining their behavior in two main cases: in the first place, the identity weighting matrix, which can be intended as the rudest case because the estimated parameters are not the best ones possible in terms of efficiency[21].

After this, the inclusion of a weighting matrix - Newey-West (1987) - that takes into account the possible serial correlation effects in sample can

[20] Again, we have assumed that the jump size corresponds to a realization of a lognormal random variable U of the form $U = e^{-\beta^2/2+\beta\varepsilon} - 1$, with $\varepsilon \sim N(0,1)$.

[21] Cochrane (2005) also suggests to look at the first stage estimates although it is not so efficient as possible

reveal additional insights about the estimation (see Hamilton, 1994). Again, it is worth to notice that this is a first approach to the methodology as a more detailed study will be presented in the next sections.

Our procedure involves three moment conditions, corresponding to the variance and the two first autocorrelation coefficients of both restricted SNJD and classical Shot noise processes. We have just taken two moment conditions referring to the correlation terms as we have the intention of limiting as much as possible the set of information that arrives at our estimate[22].

Autocorrelation coefficients

To start, Table 2.2 exhibits the sample and theoretical autocorrelation coefficients of the series under study. The sample autocorrelation coefficients have been computed from the mean values of the first few lag autocorrelation coefficients for the different paths. Theoretical expressions appear in this table.

Some conclusions arise from the inspection of Table 2.2. In general, the sample autocorrelation values for the restricted SNJD process are not so close to that predicted from the theory, specially if we compare with those obtained for the classical Shot noise. Moreover, looking at the sample and theoretical values of the coefficients for the SNJD model, it seems that they are closer in certain values ($a = 0.6, 1.0$) than in other ones ($a = 0.2, 10.0$). These last results can be drawing some features about the adequacy of the approximations used.

GMM estimates

Results in Tables 2.3 and 2.4 seem to go in the same direction as those obtained in Table 2.2. Table 2.3 refers the GMM estimates obtained for a classical Shot noise of equation (2.2.4.2), and Table 2.4 exhibits those referred to the restricted SNJD model. Both tables display the mean, median, minimum, maximum, mean squared error (MSE) and root mean squared error (RMSE) of the GMM estimated values for the parameters λ, β and a of the restricted SNJD and classic Shot noise processes. For each parameter estimate we report four columns corresponding to each different cases under analysis, i.e., for $a = 0.2, 0.6, 1.0$ and 10.0. Moreover, each table is

[22]The number of parameters to estimate (three) indicates the minimum number of orthogonality conditions that must be included to have correctly determined our problem. See Hamilton (1994) or Cochrane (2005) for additional insights on the GMM procedure.

Table 2.2: Autocorrelation values for first few lags.

Theoretical expressions for the autocorrelation functions of the restricted SNJD and classic Shot noise processes are, respectively,

$$\rho_n(H^r) = \begin{cases} 1, & \text{if } n = 0 \\ (e^{-a\Delta t} - 1)\, e^{-a\, n\Delta t}, & \text{if } n \geq 1 \end{cases}$$

and

$$\rho_n(H^c) = e^{-an\Delta t}$$

Restricted SNJD process ($\mu = \sigma = 0$)

	$a = 0.2$		$a = 0.6$		$a = 1.0$		$a = 10.0$	
Lag	ρ_{teo}	$\hat{\rho}$	ρ_{teo}	$\hat{\rho}$	ρ_{teo}	$\hat{\rho}$	ρ_{teo}	$\hat{\rho}$
1	-0.148	-0.092	-0.247	-0.227	-0.233	-0.316	0.000	-0.500
2	-0.122	-0.075	-0.136	-0.123	-0.086	-0.117	0.000	0.000
3	-0.100	-0.061	-0.075	-0.067	-0.032	-0.042	0.000	0.000
4	-0.081	-0.050	-0.041	-0.037	-0.012	-0.016	0.000	0.000
5	-0.067	-0.041	-0.023	-0.021	-0.004	-0.006	0.000	-0.001

Classical shot noise (Exact moments)

	$a = 0.2$		$a = 0.6$		$a = 1.0$		$a = 10.0$	
Lag	ρ_{teo}	$\hat{\rho}$	ρ_{teo}	$\hat{\rho}$	ρ_{teo}	$\hat{\rho}$	ρ_{teo}	$\hat{\rho}$
1	0.819	0.816	0.549	0.548	0.368	0.367	1.000	1.000
2	0.670	0.665	0.301	0.300	0.135	0.134	0.000	0.000
3	0.549	0.542	0.165	0.163	0.050	0.049	0.000	0.001
4	0.449	0.442	0.091	0.089	0.018	0.018	0.000	-0.001
5	0.368	0.360	0.050	0.048	0.007	0.006	0.000	-0.001

This table reports autocorrelation coefficients for first few lags and different values of the parameter a. First column displays the number of lags. Each successive pair of columns corresponds to the theoretical (ρ_{teo}) and sample ($\hat{\rho}$) autocorrelation values, respectively. The sample autocorrelation values are obtained from the computed mean of the first few autocorrelation values for 3,000 simulation paths with 1,800 data each one. The response function of the restricted SNJD model is $h(t - \tau_k) = e^{-a(t-\tau_k)}$ and the random variable U_j is lognormally distributed with mean $E[U] = 0$. Simulated parameters are $\lambda = 0.04$ and $\beta = 0.10$ for both models.

also divided in two main blocks, attending to the weighting matrix (identity or Newey-West (1987) estimate) used. Finally, the true values used in the simulations for the parameters are also displayed in parenthesis.

In general, in terms of RMSE, the best estimates correspond to those of the parameter a across models and weighting matrix. This is not unexpected as, in general, two out of three moment conditions are referred to this parameter. It is particularly singular that the addition of the Newey-West (1987) matrix results in a increasing of efficiency in a estimates of any order (0.2, 0.6,...) for the classical Shot noise; however, this is just observable in case of $a = 0.6, 1.0$ for the restricted SNJD model, if we exclude the case of $a = 10.00$ due to their high RMSE value.

On the whole, in terms of RMSE, it seems that the values estimated for the parameter a in the classical Shot noise are closer to their simulated values than those obtained for the restricted SNJD model. This result could be directly related to those obtained in Table 2.2 to the extent that the parameter a i) affects dramatically to the poorness of the results for restricted SNJD model for $a = 0.2$ and 10.0 values and, ii)the former could be a consequence of the numerical approximations done in the restricted SNJD. This idea is also reinforced by the behavior of the estimated a's values in the SNJD model, because they follows the pattern detected in Table 2.2 where better a's estimates are obtained for $a = 0.6, 1.0$ than for $a = 0.2, 10.0$.

Concerning the parameter λ, it seems to exist a bias with respect to their simulated values, a feature that is repeated across models and weighting matrix estimates. This bias tends to increase with the parameter a, with the only exception of $a = 10.0$ for the SNJD model.

The same conclusions arise with respect to the parameter β: the more we increase the a, the more bias we find. These patterns could be also confirmed by the value of their RMSEs for parameters λ and β, respectively. Finally, in both models, the observed differences between Mean and Median parameter with reference to λ and β parameters could be reflecting certain asymmetries in the estimate distribution.

Can we avoid this bias behavior in the estimates of the parameters λ and β? Up to a point, whether we have detected a problem with the estimation methodology is an open question that we must confirm in the next sections. It is worth to say that this is a first contact with the methodology, so these results are not definitive mainly because of the following reasons:

1. The efficiency of the estimates can be improved by increasing the set of information of our estimate, that is, increasing the number of orthogonal conditions, specially in the case of the Newey-West (1987)

weighting matrix for the restricted SNJD[23].

2. We must check whether this bias in the parameters λ and β are just present in the estimations for Shot noise samples or if the bias is also extensive to more general JD estimations[24].

Partially, it seems that the methodology is capable of facing the estimation of the parameters of the restricted SNJD model for a certain range of values, an affirmation that could be extended to the classical Shot noise in a wider range. Nevertheless, this conclusion must be still exhaustively proved.

2.5 Monte Carlo study

Once slightly checked the ability of the proposed methodology to capture the parameters of the process, we proceed now to analyze the performance of the estimations of the JD and SNJD models using the GMM estimate. Due to the strong differences between the JD and SNJD models, it could be interesting to develop two main experiments:

1. To analyze the behavior of the GMM estimate of JD and SNJD models *under* JD and SNJD samples, respectively.

2. To study the case of cross samples, that is, data generated by a model that differs from that one used in the estimation (e.g., to estimate a JD model by using a simulated SNJD sample).

Since the second study will be performed in the next section, this section is devoted to the former.

On the whole, this section is developed by considering three versions of the GMM weighting matrix: to start, we use the *classical* moment estimate; this will be our *first-stage* estimate, following notation used in Cochrane (2003). Afterwards, we check the procedure using the efficient weighting matrix of equation (2.2.2.4) posited by Hansen (1982). This estimate will be named the *second-stage* estimate. Lastly, due to the presence of serial correlation in some series under study, we analyze the results obtained by using the Newey-West (1987) estimate of expression (2.2.2.5).

[23]Two estimation exercises using 6 moment conditions for a restricted SNJD model seems to confirm this point. These results are available upon request.

[24]To our knowledge, there is no study analyzing how the presence of serial correlation affects the estimation of JD parameters. This point will be discussed in detail later.

Table 2.3: GMM estimates for the classic Shot noise process.

Parameters	$\lambda \times 10^{-2}$				$\beta \times 10^{-2}$				a			
(True value)	(4.00)				(10.00)				(below)			
Parameter a	0.20	0.60	1.00	10.00	0.20	0.60	1.00	10.00	0.20	0.60	1.00	10.00
Identity matrix												
Mean	7.20	7.70	8.57	15.81	11.36	12.33	13.48	23.19	0.19	0.57	0.96	7.51
Median	7.45	8.04	8.89	16.79	11.94	13.11	14.24	24.37	0.20	0.60	1.00	7.02
Min	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	-0.01	0.00	0.00	0.00
Max	14.20	14.76	19.98	20.00	22.22	19.53	30.00	30.00	0.29	0.83	1.44	15.60
MSE	0.17	0.21	0.29	1.58	0.13	0.20	0.28	2.09	0.00	0.02	0.05	20.72
RMSE	4.11	4.59	5.40	12.55	3.64	4.44	5.32	14.46	0.06	0.15	0.22	4.55
Newey-West (1987)												
Mean	13.80	15.05	16.30	30.47	8.55	9.73	10.80	17.93	0.20	0.60	0.99	6.77
Median	14.84	16.21	17.61	30.77	8.24	9.29	10.38	18.66	0.20	0.60	1.00	6.98
Min	0.85	0.94	0.82	1.62	1.70	1.87	1.68	1.12	0.01	0.01	0.02	0.02
Max	22.97	27.89	28.42	40.00	15.93	18.94	19.21	20.00	0.30	0.82	1.45	12.00
MSE	1.08	1.40	1.72	7.47	0.04	0.03	0.05	0.68	0.00	0.00	0.03	15.19
RMSE	10.41	11.84	13.13	27.34	2.06	1.86	2.15	8.24	0.03	0.09	0.16	3.90

This table reports Monte Carlo study to analyze GMM estimates of a classical Shot noise process (see equation (2.2.4.2)) by using the Identity and Newey-West (1987) estimates as weighting matrices. This table shows mean, median, minimum, maximum, mean squared error (MSE) and root mean squared error (RMSE) of the estimated parameters, respectively. Simulation parameters are in brackets. The random variable U_j is lognormally distributed $U = e^{-\beta^2/2+\beta\varepsilon} - 1$, with $\varepsilon \sim N(0,1)$ and mean $E[U] = 0$.

Table 2.4: GMM estimates for the restricted SNJD process ($\mu = \sigma = 0$).

Parameters	$\lambda \times 10^{-2}$				$\beta \times 10^{-2}$				a			
(True value)	(4.00)				(10.00)				(below)			
Parameter a	0.20	0.60	1.00	10.00	0.20	0.60	1.00	10.00	0.20	0.60	1.00	10.00
Identity matrix												
Mean	6.86	7.23	8.09	5.16	11.66	12.22	12.86	9.07	0.07	0.64	0.71	0.83
Median	6.77	7.12	7.55	5.03	11.55	12.05	13.42	9.26	0.08	0.70	0.72	0.83
Min	0.82	2.52	0.80	0.80	1.22	6.48	1.02	1.00	0.00	0.22	0.49	0.65
Max	12.29	13.97	18.10	19.03	19.80	17.38	21.44	24.56	0.21	1.41	0.93	0.93
MSE	0.12	0.14	0.29	0.09	0.07	0.08	0.24	0.22	0.02	0.06	0.09	84.12
RMSE	3.48	3.79	5.43	3.06	2.61	2.88	4.90	4.69	0.14	0.25	0.30	9.17
Newey-West (1987)												
Mean	8.35	19.39	8.33	8.33	8.06	8.90	8.19	8.19	2.16	0.61	0.88	1.62
Median	8.31	12.43	8.25	8.25	8.10	10.50	8.27	8.27	2.61	0.61	0.91	1.56
Min	0.81	0.82	0.81	0.81	1.01	1.00	1.00	1.00	0.00	0.00	0.00	0.00
Max	16.00	40.02	16.00	16.00	15.00	14.99	14.99	14.99	4.00	1.52	1.32	11.18
MSE	0.39	4.73	0.38	0.38	0.20	0.14	0.20	0.20	6.12	0.06	0.06	70.50
RMSE	6.23	21.76	6.20	6.20	4.50	3.68	4.44	4.44	2.47	0.23	0.25	8.40

This table includes Monte Carlo study to analyze GMM estimates of a restricted SNJD process (see equation (2.2.4.1)) by using the Identity and Newey-West (1987) estimates as weighting matrix. Parameters μ and σ for general SNJD model of equation (2.2.3.1) have been fixed to zero under the restricted SNJD model. This table shows mean, median, minimum, maximum, mean squared error (MSE) and root mean squared error (RMSE) of the estimated parameters, respectively. Simulation parameters are in brackets. The response function of the restricted SNJD model is $h(t-\tau_k) = e^{-a(t-\tau_k)}$, and the random variable U_j is lognormally distributed $U = e^{-\beta^2/2+\beta\varepsilon} - 1$, with $\varepsilon \sim N(0,1)$ and mean $E[U] = 0$.

2.5.1 GMM estimates for the JD processes

Results

We report here the results obtained from Monte Carlo simulations for the JD process. As there is no correlation in the sample[25] there is no sense in using the Newey-West (1987) weighting matrix. Then, we only present the results for the first and second-stage estimates. In general terms, we expect that the inclusion of the second stage estimate must lead, *a priori*, in a bias reduction of the estimates obtained in the first stage.

To clarify the exposition, just a few words about the process: our study comprises the simulations of the log-returns in model (2.2.3.1). The vector of parameters under study is $\Lambda = (\mu, \sigma, \lambda, \beta)' \in R^4$, where μ and σ refer to the drift and volatility of the Brownian part of the process, λ captures the (instantaneously) probability that a jump happens and β refers to the jump amplitude when a jump occurs. The parameter a has been fixed to zero. As we are using through this chapter, the jump size U is lognormally distributed of the form $U = e^{-\beta^2/2+\beta\varepsilon} - 1$, with $\varepsilon \sim N(0,1)$. Finally, the simulated JD process can be seen as a version of the Merton (1976) process with a null jump size mean.

Table 2.5 exhibits the GMM estimates for a JD process in a sample of 3,000 paths with 1,800 sample data each one. True simulation values appear in parenthesis. Each parameter block is divided in two columns, attending to the weighting matrix used (first or second-stage estimate). Rows in this table also display the mean, median, minimum, maximum, mean squared error (MSE) and root mean squared error (RMSE) for the estimated parameters.

Moreover, we provide the P percentiles $[2.5\%, 97.5\%]$ of each set of estimates obtained. Finally, we also include the Pearson's variation coefficient (CV), which indicates the relationship among the standard deviation and the mean of a distribution, and serves us to compare two distributions with different means[26]. Of course, we expect that the inclusion of the second stage estimate results in a reduction of the CV values computed for any parameter.

Looking at the results in Table 2.5, the mean values of estimates obtained are, in general, close to their true values attending to their respective standard deviation values. Of course, standard deviations decline substantially when it comes to the estimates results of the second stage estimates. In line

[25] As previously developed in Chapter 1, a JD process is not capable of generating serial persistence.

[26] The higher the CV, the higher the dispersion in the data

Table 2.5: GMM estimates for the JD process

	$\widehat{\mu} \times 10^{-2}$ (0.02)		$\widehat{\sigma} \times 10^{-2}$ (1.26)		$\widehat{\lambda} \times 10^{-2}$ (4.00)		$\widehat{\beta} \times 10^{-2}$ (10.00)	
Statistics	1st	2nd	1st	2nd	1st	2nd	1st	2nd
Mean	0.02	0.02	1.33	0.92	7.41	6.49	7.23	8.15
Median	0.02	0.02	1.39	0.96	7.17	6.50	7.19	8.18
Std	0.06	0.04	0.60	0.23	3.64	0.40	2.63	1.02
Min	-0.18	-0.14	0.06	0.00	0.50	2.72	1.01	1.02
Max	0.19	0.16	2.66	1.90	17.19	7.96	15.18	12.06
MSE	0.00	0.00	0.00	0.00	0.25	0.06	0.15	0.04
RMSE	0.06	0.04	0.60	0.41	4.99	2.52	3.82	2.11
$P_{[2.5,97.5]}$	[-0.09, 0.13]	[-0.06,0.10]	[0.26,2.32]	[0.37,1.27]	[1.59,14.94]	[5.71,7.21]	[2.22,12.57]	[6.12, 10.03]
CV	298.67	203.93	44.80	24.48	49.18	6.20	36.41	12.53

This table shows the GMM first and second stage estimates of the JD process. We assume a jump size U lognormally distributed of the form $U = e^{-\beta^2/2+\beta\varepsilon} - 1$, with $\varepsilon \sim N(0,1)$. First row contains the parameters of JD process (true simulation values in brackets). This table shows mean, median, minimum, maximum, mean square error (MSE) and root mean square error (RMSE) of the estimated parameters using the GMM methodology, respectively. We also provide the P percentiles $[2.5\%, 97.5\%]$ and the Pearson's variation coefficient (CV) of each set of estimates obtained. CV is defined as the ratio of the standard deviation to the mean, and it serves us to compare the dispersion of distributions with different means. Finally, the first stage column correspond to a GMM with the identity as weighting matrix. Second stage column uses the efficient version of weighting matrix of Hansen (1982).

with this point, it is possible to observe that the efficient weighting matrix of Hansen (1982) - our second-stage estimate - improves the efficiency of the estimates, a point that is also confirmed by the reduction in terms of RMSE and CV for each pair of parameter estimates. This is particularly notable in the case of the parameter λ, which initially goes from a RMSE of 4.99×10^{-2} to 2.52×10^{-2}.

Anyway, the estimated values of the parameter λ are bigger (in mean) than their true simulated values, and this is also evident once taking into account the small value of the standard deviations registered for the second stage estimate; consequently, this results in a increasing of the rate of jump events, going from 10 jumps per year (simulated) to about 16 per year (estimated). We will go back exhaustively to this point later.

If we attend to the differences among mean and median parameter estimates, the distribution for σ and λ parameters seems to be asymmetrical with respect to its first stage estimate, and only σ with respect to the second one. This point is also confirmed by the histograms for the first and second stage estimates for the parameter σ (see Figure 2.1) and the parameters λ and β (see Figure 2.2).

Moreover, Figures 2.1 and 2.2 show that the variance of the sample distributions decreases considerably when the efficient weighting matrix of Hansen (1982) is included. Letting this issue apart, we can consider, as a general rule, that all the parameter distributions seem to be normally distributed[27].

With respect to the estimation of the parameter σ, Figure 2.1 shows that the histogram of their estimated values does not seem to be normally distributed for the first stage estimate and, for the second stage estimate, it also exhibits a left tail of the distribution that it fatter than that from a normal distribution. We have repeated the experiment by using other samples, and results do not change qualitatively. We think that the estimation of this parameter can be affected of numerical problems, although the next paragraphs present a general discussion about this point.

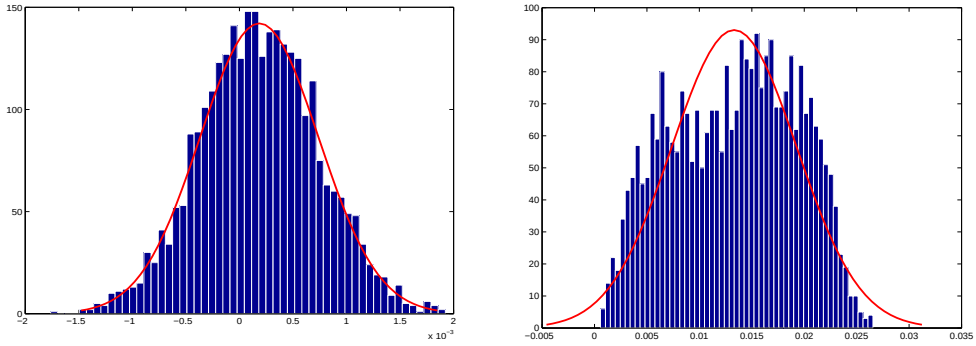
Comments on the JD results

Two issues have not been analyzed in the precedent paragraphs: the first one refers to the observed bias detected in the estimation of the parameter λ; the second one treats about problems in the asymptotic distribution of the parameter σ. At a certain extent, both questions could be connected.

[27] This feature is predicted by the theory. See, for instance, Hansen (1982) or Cochrane (2005).

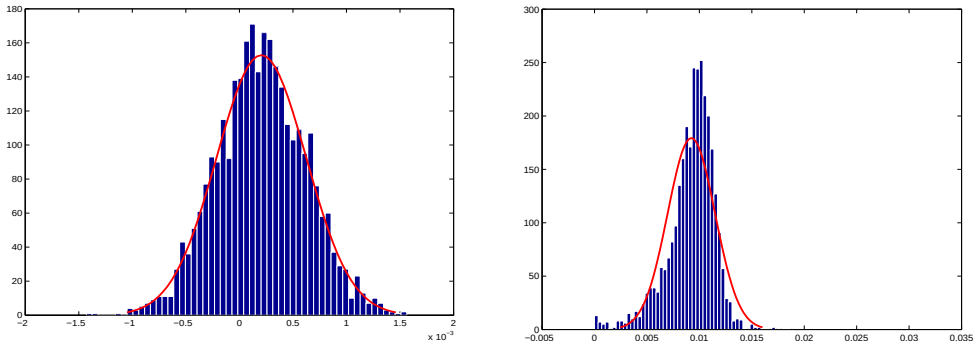
Figure 2.1: Histogram for the JD estimate parameters μ and σ

Identity weighting matrix (first stage estimate)



$\hat{\mu}$ $\hat{\sigma}$

Hansen (1982) weighting matrix (second stage estimate)

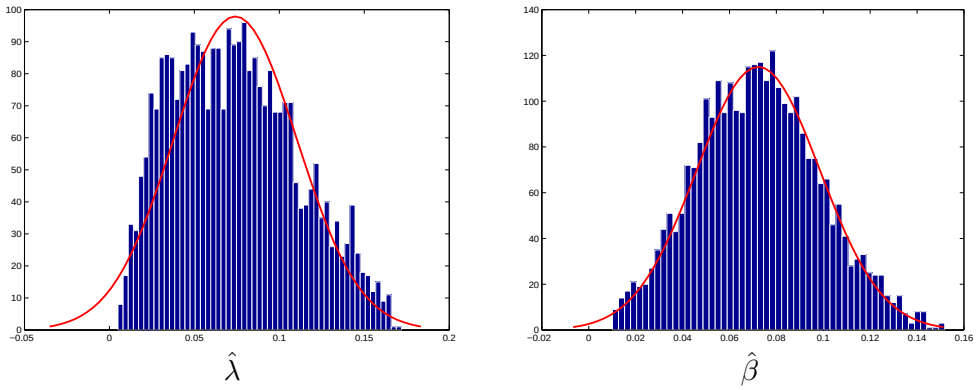


$\hat{\mu}$ $\hat{\sigma}$

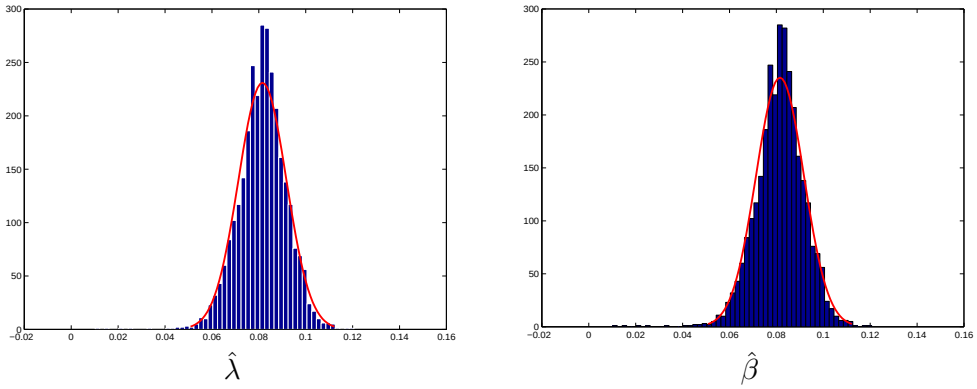
Sample and asymptotic distributions for the GMM estimates obtained with Identity and Hansen (1982) weighting matrix for the parameters μ and σ. The sample is composed by simulations of 3,000 paths with 1,800 data each one, corresponding to daily frequency. Simulations parameters are $\mu = 0.0002$, $\sigma = 0.0126$, $\lambda = 0.0400$ and $\beta = 0.1000$ (parameter a has been fixed to zero). The jump size U is lognormally distributed of the form $U = e^{-\beta^2/2+\beta\varepsilon} - 1$, with $\varepsilon \sim N(0,1)$ and mean $E[U] = 0$. Notice that figures are equally scaled in the X axis for each pair of estimates corresponding to the same parameter.

Figure 2.2: Histogram for the JD estimate parameters λ and β

Identity weighting matrix (first stage estimate)



Hansen (1982) weighting matrix (second stage estimate)



Sample and asymptotic distributions for the GMM estimates obtained with Identity and Hansen (1982) weighting matrix for the parameters λ and β. The sample is composed by simulations of 3,000 paths with 1,800 data each one, corresponding to daily frequency. Simulations parameters are $\mu = 0.0002$, $\sigma = 0.0126$, $\lambda = 0.0400$ and $\beta = 0.1000$ (parameter a has been fixed to zero). The jump size U is lognormally distributed of the form $U = e^{-\beta^2/2+\beta\varepsilon} - 1$, with $\varepsilon \sim N(0,1)$ and mean $E[U] = 0$. Note that figures are equally scaled in the X axis for each pair of estimates corresponding to the same parameter.

Regarding the first issue, a possible explanation to the bias in the parameter λ can be given by the observed reduction in the mean value of estimated σ's (second-stage column): this might be a consequence of a *transference* of volatility from the Brownian to the jump part of the process.

To clarify this idea, let us put some numbers to our discussion: from the expression of the total variance[28] of a JD process, we have that the theoretical (daily) variance[29] is about $\sigma^2_{theo} = 5.59 \times 10^{-4}$, and the sample (daily) variance[30] obtained is $\sigma^2_{sample} = 5.16 \times 10^{-4}$. These values imply an annualized standard deviation of about $\sigma_{theo} = 37\%$ and $\sigma_{sample} = 36\%$ for the theoretical and sample model parameters, respectively. By contrast, the annual rate of jumps implied from theoretical and sample results are 10 and 16, respectively.

Some additional arguments in favour of this effect on the total JD variance behavior could be found by looking at the joint behavior of the parameters under study. Firstly, we have estimated the correlation matrix of the results obtained for the second stage estimates[31]. Table 2.6 reports these correlation coefficients.

As Table 2.6 exhibits, only the correlation coefficients for the pairs (σ, λ) and (σ, β) remains significant for a 95%-confidence level. The correlation values for (σ, λ) and (σ, β) are about 0.45 and -0.73, respectively. Indeed, it seems to be confirmed the existence of a linear relationship in between these two terms.

By contrast, this decreasing effect on the total variance of the process is compensated by an increase on the jump amplitude: there is a negative correlation between σ and β while the correlation between λ and β remains insignificant.

Additionally, to have a visual idea of this issue, Figures 2.3 and 2.4 include the point clouds for some pair of the parameters under study (σ, λ, and β).

As expected, the point clouds in Figures 2.3 and 2.4 confirm the sign and dependence existing between the pairs (σ, λ) and (σ, β): we observe

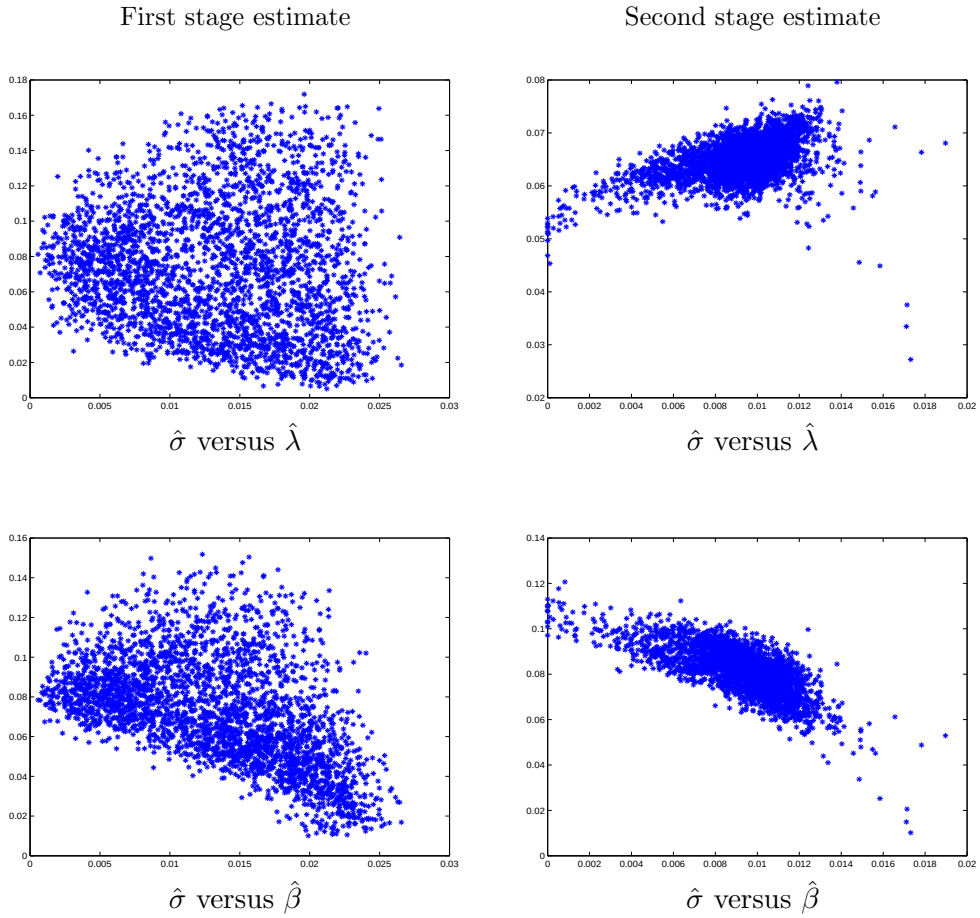
[28]The total variance of a JD process is $Var(Z_t) = \sigma^2 \Delta t + \lambda E[U^2] \Delta t$, where we have assumed a jump size U lognormally distributed of the form $U = e^{-\beta^2/2 + \beta\varepsilon} - 1$, with $\varepsilon \sim N(0,1)$, and $E[U^2] = \beta^2 + \beta^4/4$. See Chapter 1 for a detailed development of this formula.

[29]This variance is computed from the true simulation parameters.

[30]This variance is computed from the mean values for the parameters σ, β and λ of the second stage estimates in Table 2.5.

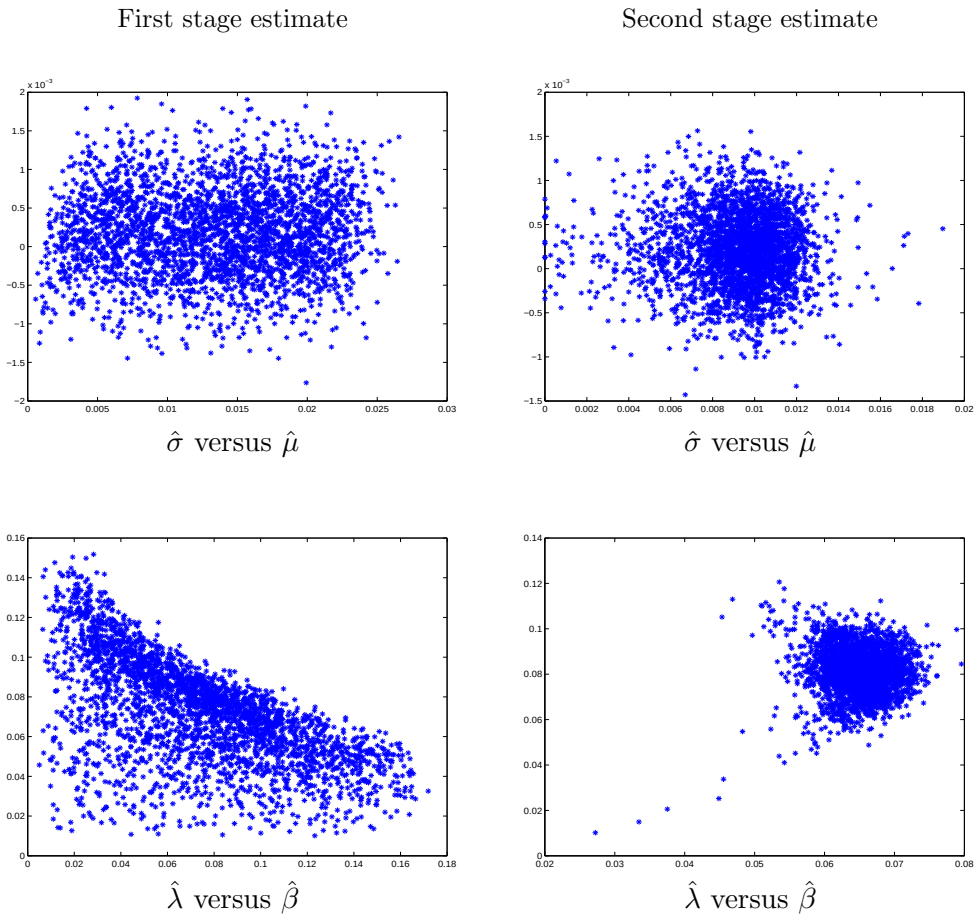
[31]The correlation matrix for the first stage estimates has also been computed and is available upon request.

Figure 2.3: Points clouds for different GMM estimates of JD process



Different point clouds for the GMM estimates of the JD process parameters. The first (second) column represents the estimates obtained by using the identity (Hansen, 1982) matrix as weighting matrix in the GMM estimates (first and second stage estimates, respectively). Simulation parameters are $\mu = 0.0002$, $\sigma = 0.0126$, $\lambda = 0.0400$ and $\beta = 0.1000$ (parameter a has been fixed to zero). The jump size U is lognormally distributed of the form $U = e^{-\beta^2/2+\beta\varepsilon} - 1$, with $\varepsilon \sim N(0,1)$ and mean $E[U] = 0$. Finally, "$\hat{\sigma}$ versus $\hat{\lambda}$ ($\hat{\beta}$)" refers to parameter $\hat{\sigma}$ on the axis OX and parameter $\hat{\lambda}$ ($\hat{\beta}$) on the axis OY.

Figure 2.4: Points clouds for different GMM estimates of JD process



Different point clouds for the GMM estimates of the JD process parameters. The first (second) column represents the estimates obtained by using the identity (Hansen, 1982) matrix as weighting matrix in the GMM estimates (first and second stage estimates, respectively). Simulation parameters are $\mu = 0.0002$, $\sigma = 0.0126$, $\lambda = 0.0400$ and $\beta = 0.1000$ (parameter a has been fixed to zero). The jump size U is lognormally distributed of the form $U = e^{-\beta^2/2+\beta\varepsilon} - 1$, with $\varepsilon \sim N(0,1)$ and mean $E[U] = 0$. Finally, "$\hat{\sigma}$ versus $\hat{\mu}$" refers to parameter $\hat{\sigma}$ on the axis OX and parameter $\hat{\mu}$ on the axis OY (and similarly for "$\hat{\lambda}$ versus $\hat{\beta}$").

Table 2.6: Correlation matrix for the GMM second stage JD results

	μ	σ	λ	β
μ	1.0000			
σ	**-0.0261**	1.0000		
λ	**0.0218**	0.4489	1.0000	
β	**0.0193**	-0.7305	**-0.0062**	1.0000

Correlation matrix of the GMM estimates for the JD process using Newey-West (1987) as weighting matrix. The parameter vector is $\Lambda = (\mu, \sigma, \lambda, \beta)' \in R^4$, and the simulation parameter values are $\mu = 0.0002$, $\sigma = 0.0126$, $\lambda = 0.0400$ and $\beta = 0.1000$. Parameters in bold are not significant for a 95%-confidence level.

elliptical forms with a clear tendency between variables. By contrast, the pairs included in Figure 2.4 remain in a circular form, a clear sign of linear independence.

To a certain degree, these differences on allocating the total variance of the process seem to explain this bias. To put it another way, the GMM estimate does not appear to assign completely the correct variance due to those different contributions, the Brownian and jump ones. Whether this behave it is due to numerical problems on estimating σ or, by contrast, it is caused by a possible cross-effect among the remaining parameters in the total variance expression of the JD process, it is a question that remains open.

2.5.2 GMM estimates of SNJD processes

We display now the results obtained for the GMM estimations of the SNJD model under different simulation samples generated by the SNJD process.

Results

As it was extensively shown in Chapter 1, one of the main features of the SNJD models is their capability to generate serial correlation in returns. With the purpose of checking whether the GMM approach is able to capture this dependence in data, we have developed three experiments which differ in the degree of autocorrelation generated in the sample. Basically, we

have modified the parameter a among experiments, keeping the remaining components of the parameter vector (μ, σ, λ, and β) constant across experiments.

Table 2.7 exhibits the GMM estimated parameters for the SNJD model under study. The two first columns of this table display the parameters under study and their true values, respectively. Afterwards, three main column blocks contain the results for Identity (first stage), Hansen (1982) (second stage) and Newey-West (1987) (third stage) weighting matrix of the GMM estimates[32].

To keep manageable the main results of the process, each block reports the mean, root mean squared error (RMSE) and the P percentiles $[2.5\%, 97.5\%]$ of the proposed estimates (percentiles are in brackets below the mean statistic). Finally, to compare across samples of the different experiments, the Pearson's coefficient of variation (CV) is also included.

We discuss now in more detail some features derived from Table 2.7:

- **Parameter** σ On the whole, we observe for equal weighting matrix a decay pattern in the reported mean value of σ when the parameter a increases (e.g., the second stage estimate passes from 1.75×10^{-2} to 1.16×10^{-2} for $a = 0.20$ and 1.00, respectively).

  When the parameter a increases, the bias increases for the first stage estimate and decreases for the second and third stage estimates. In all the cases, the third stage estimate provides a slightly smaller bias than the second stage one.

  As a general rule, the inclusion of the Newey-West (1987) weighting matrix increases the efficiency of the estimates, as reflected in the following facts:

  - The CV coefficients decay across different estimates for equal set of parameter values, mainly due to the fact that the dispersion in the estimated data decreases.
  - The 95%-confidence interval for the third stage estimates are narrower than considering the remaining estimates[33].

[32] For the sake of brevity, we have omitted the mean sample autocorrelation coefficients for the three developed experiments. Just for keeping an intuition, the mean sample autocorrelation results are close to those obtained in Table 2.2 for the restricted version of the SNJD model. On the whole, it can be said that these series exhibit significant sample autocorrelation values for first and second lags. This justifies the usage of the Newey-West (1987) weighting matrix estimate. Anyway, the results concerning the sample autocorrelation of simulated series are available upon request.

[33] The values obtained for the 2.5%-percentile of the parameter σ from the first and

Table 2.7: GMM estimates for the SNJD process

Parameter	True value	First stage			Second stage			Third stage		
		Mean	RMSE	CV	Mean	RMSE	CV	Mean	RMSE	CV
$\mu \times 10^{-2}$	0.02	0.02 [−0.12,0.12]	0.06	348.85	0.06 [−0.03,0.14]	0.06	73.06	0.05 [−0.03,0.13]	0.05	72.92
$\sigma \times 10^{-2}$	1.26	1.54 [0.06,2.93]	0.76	45.59	1.75 [1.38,2.10]	0.54	12.80	1.72 [1.36,2.08]	0.49	10.63
$\lambda \times 10^{-2}$	4.00	4.99 [1.61,9.10]	2.17	38.69	8.30 [2.17,16.88]	5.94	49.28	8.87 [2.80,16.53]	6.01	39.74
$\beta \times 10^{-2}$	10.00	11.31 [4.53,19.06]	4.08	34.13	8.01 [5.59,10.59]	2.37	16.02	7.72 [5.53,10.44]	2.60	16.16
a	0.20	0.51 [0.06,2.38]	0.67	1.18	0.38 [0.02,1.45]	0.43	1.00	0.48 [0.04,1.75]	0.60	1.09
$\mu \times 10^{-2}$	0.02	0.01 [−0.06,0.09]	0.04	253.35	0.06 [−0.03,0.14]	0.06	73.26	0.05 [−0.03,0.11]	0.04	78.00
$\sigma \times 10^{-2}$	1.26	0.92 [−0.01,1.65]	0.54	45.34	1.59 [−0.63,2.09]	0.57	29.26	1.55 [1.29,1.81]	0.32	8.42
$\lambda \times 10^{-2}$	4.00	7.14 [2.94,11.59]	3.75	28.70	15.91 [3.62,36.39]	13.83	44.26	17.15 [10.23,24.45]	13.65	21.26
$\beta \times 10^{-2}$	10.00	10.49 [5.78,15.69]	2.61	24.39	7.02 [5.20,9.43]	3.16	15.17	6.57 [4.98,8.19]	3.53	12.51
a	0.60	0.63 [0.16,1.31]	0.31	0.49	0.60 [0.03,1.13]	0.26	0.43	0.61 [0.25,1.08]	0.22	0.37
$\mu \times 10^{-2}$	0.02	0.01 [−0.05,0.07]	0.03	237.38	0.06 [−0.03,0.14]	0.06	70.86	0.04 [−0.03,0.11]	0.04	80.68
$\sigma \times 10^{-2}$	1.26	0.30 [−0.03,1.36]	1.05	141.88	1.16 [−0.32,2.00]	0.80	68.84	1.37 [1.04,1.62]	0.18	10.89
$\lambda \times 10^{-2}$	4.00	9.15 [4.19,13.00]	5.60	24.04	25.78 [8.97,47.99]	24.41	42.71	23.76 [15.07,31.86]	20.28	19.15
$\beta \times 10^{-2}$	10.00	19.30 [9.30,26.61]	10.32	23.27	6.50 [4.90,8.51]	3.62	14.14	6.18 [4.82,7.65]	3.89	11.57
a	1.00	0.84 [0.04,1.13]	0.27	0.26	0.84 [0.03,1.32]	0.33	0.35	0.83 [0.45,1.20]	0.27	0.25

This Table includes the GMM (first, second and third stage) estimates of the SNJD process for different values of a parameter. The first stage estimate corresponds to a GMM with the identity as weighting matrix. The second stages estimate uses the efficient version of the weighting matrix of Hansen (1982). Due to the presence of autocorrelation in the sample, we use the weighting matrix version of Newey-West (1987) as third stage estimate. We assume a jump size U lognormally distributed of the form $U = e^{-\beta^2/2+\beta\varepsilon} - 1$, with $\varepsilon \sim N(0,1)$, and response function of the form $h(t) = e^{-at}$. Each row refers the true simulation values, mean, RMSE and CV of first to third stage estimates for each parameter of the SNJD process, respectively.

Finally, the decay of the RMSE values across different estimates for equal set of parameter values might reflect a trade-off between precision and bias: the Newey-West (1987) estimates present bias but these estimates are (a little) more precise than the remaining estimates.

- **Parameter** λ In general, the first stage estimates of λ exhibit substantially smaller bias than the second and third stage ones, even for different values of the parameter a. Moreover, the mean λ values seem to be biased to the right, a pattern that is also repeated across estimates and different degrees of serial correlation.

  It is generally remarkable that the more we try to improve the efficiency of our estimations, the higher is the bias obtained (see, for instance, the case of $a = 0.6$ where the λ bias goes from 11.91 to 13.15). Note also that the Identity and Newey-West (1987) matrices seem to provide more efficient results than the Hansen (1982) matrix[34].

  Finally, at light of the results obtained, as the smallest RMSE is obtained for the first stage estimates, it seems that something is happening in the λ estimates when we include the Hansen (1982) or Newey-West (1987) for the cases under study.

Before analyzing the remaining parameters in Table 2.7, it can be interesting to have a visual idea of the former results. With this aim, Figure 2.5 displays the sample distribution for the first, second and third estimates of the parameters σ and λ when $a = 0.6$. The true simulation values are $\mu = 0.02 \times 10^{-2}$, $\sigma = 1.26 \times 10^{-2}$, $\lambda = 4.00 \times 10^{-2}$ and $\beta = 10.00 \times 10^{-2}$. The graphs in the same column are scaled on the axis OX with the intention of observing possible displacements of the sample distribution along this axis.

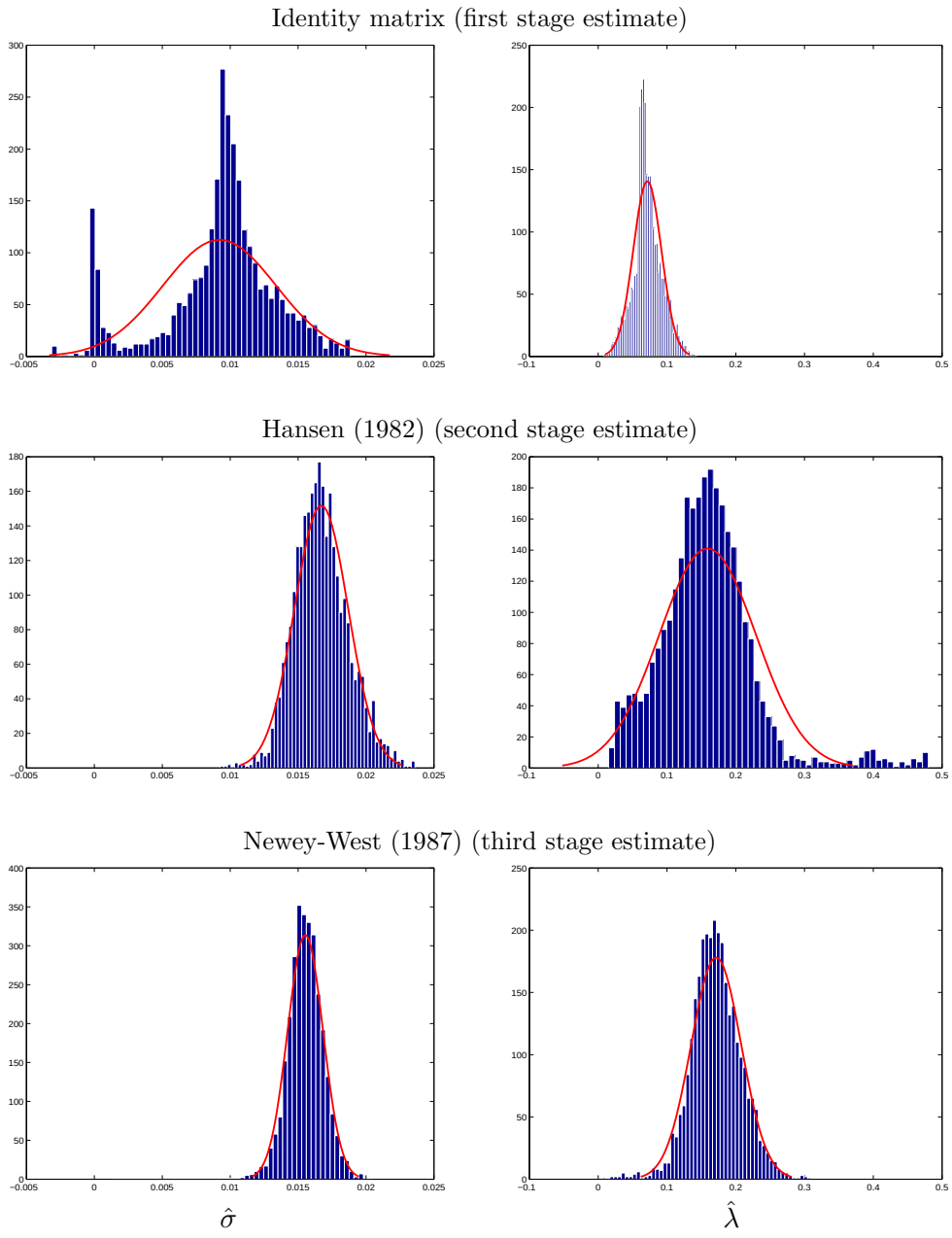
As Figure 2.5 reveals, the Newey-West (1987) matrix for the parameter σ seems to offer more efficient results than those obtained with the Identity or Hansen (1982) matrices. It is also appreciated a bias with respect to its true value (the simulated value for σ is 0.0126) and this bias is smaller for the case of the Identity matrix. Regarding this matrix, we can observe that part of the sample (less than 10%) takes values around zero or even negative, which can be revealing some numerical problems in the methodology.

---

second stage estimates for $a = 0.6$ and $a = 1.0$ are negative. We put these results down to numerical problems in the optimization procedure. Anyway, our experiments seem to indicate a negligible probability of having a negative σ value in the estimations.

[34] For instance, the coefficients of variation for λ in the case $a = 1.0$ are 24.04, 42.71 and 19.15 for the first to third stage estimates, respectively.

Figure 2.5: Sample distribution of the parameters σ and λ for $a = 0.6$



Sample distribution of estimated σ and λ SNJD parameters for different weighting matrices. Simulation values are $\mu = 0.02 \times 10^{-2}$, $\sigma = 1.26 \times 10^{-2}$, $\lambda = 4.00 \times 10^{-2}$, $\beta = 10.00 \times 10^{-2}$ and $a = 0.60$. The random variable U is lognormally distributed with mean $E[U] = 0$. Note that the axis OX are equally scaled.

The conclusions regarding the estimates of the parameter λ in Figure 2.5 are almost similar: the results with higher bias correspond to the more efficient matrices employed (see graphs including the second and third stage estimates).

Moreover, the sample distributions obtained for the first and second stage estimates for the parameter λ offer different outlines than the normal distribution. This point may be due to the presence of autocorrelation in the sample, which has not been adequately corrected. As in the case of the parameter σ, it can be seen that the smaller dispersion is obtained for the sample distribution of the first stage estimate.

After this digression, we continue with the analysis of the remaining parameters:

- **Parameter** β Leaving the first estimate results aside, the β estimates obtained seem to be generally slightly biased to the left. This could be a sign of negative correlation between the parameters λ and β, a point that will be discussed later.

  On the whole, the efficiency of the estimates is improved by the inclusion of Hansen (1982) and Newey-West (1987) estimates for all cases, as CV and percentile coefficients seem to reflect. Finally, the mean values of the parameters β tends to decrease at a small rate as the parameter a increases.

- **Parameter a** This parameter is directly connected to the autocorrelation of the SNJD model. The worst performance in terms of RMSE is obtained in the case of small serial dependence ($a = 0.2$), where the methodology seems to be slightly biased[35].

  The mean values for $a = 0.6$ and $a = 1.0$ keep almost constant across estimates, but their efficiency is improved[36] as well as their precision[37].

Once again, it may be interesting to display the sample distribution obtained for the parameter a. Figure 2.6 exhibits the sample distributions obtained for $a = 0.6$ under different GMM weighting matrices. Basically, the results for the estimates are similar each other: their distributions are not normal and all of them present problems in the left tail.

[35]The computed p-values for $a = 0.20$ are 0.40, 0.46 and 0.50 from the first to the third estimates, respectively. The remaining p-values are available upon request

[36]For instance, the variation coefficients are 0.49, 0.43 and 0.37 for the first, second and third estimates, respectively.

[37]For $a = 0.6$, the RMSE values are 0.31, 0.26 and 0.22 for the first to the third estimates, respectively.

Thus, it is possible to appreciate that some observations (less than 5% of the total amount in the worst case) are around zero. Apparently, there is no bias in the obtained results. Finally, it seems that the estimates obtained with the Newey-West (1987) matrix present less dispersion than those obtained with the alternative matrices.

To complete this study, it can be interesting to compare the results included in Table 2.7 with those displayed for the restricted SNJD model in Table 2.4. Basically, we believe that the main connecting points are as follows:

1. In mean, an increase in the parameter a results in larger λ coefficients.

2. The bias of the λ estimates increases when the Newey-West (1987) estimate is included in the restricted SNJD model, as the restricted SNJD model one does. This increase in the bias is not accompanied by a larger efficiency for the λ estimate of the non-restricted SNJD model.

3. The parameter β tends to increase with the parameter a in the restricted SNJD models. By contrast, the opposite behavior appears in the non-restricted SNJD models. This fact might result in a negative correlation pattern between the parameters β and λ.

4. On the whole, the inclusion of the Newey-West (1987) weighting matrix leads to more precise estimates of the parameter β in both restricted and non-restricted SNJD models.

5. The Newey-West (1987) matrix provides the least biased and most precise estimates of the parameter a in both restricted and non-restricted SNJD models. Moreover, these results are specially remarkable when $a = 0.6$ and $a = 1.00$.

Finally, to provide some insights about the robustness of the estimates we have obtained, Figure 2.7 displays the final values of the GMM estimate at points of convergence under different specifications of the weighting matrix used. The upper plot of this figure exhibits the objective function value of the estimate J_T on expression (2.2.2.3) as a function of a and σ, while the lower plot graphs the function value depending on a and λ. Both plots are related to the parameters for the SNJD model corresponding to the case $a = 1.00$ in Table 2.7. We have chosen these variables as our intention is to

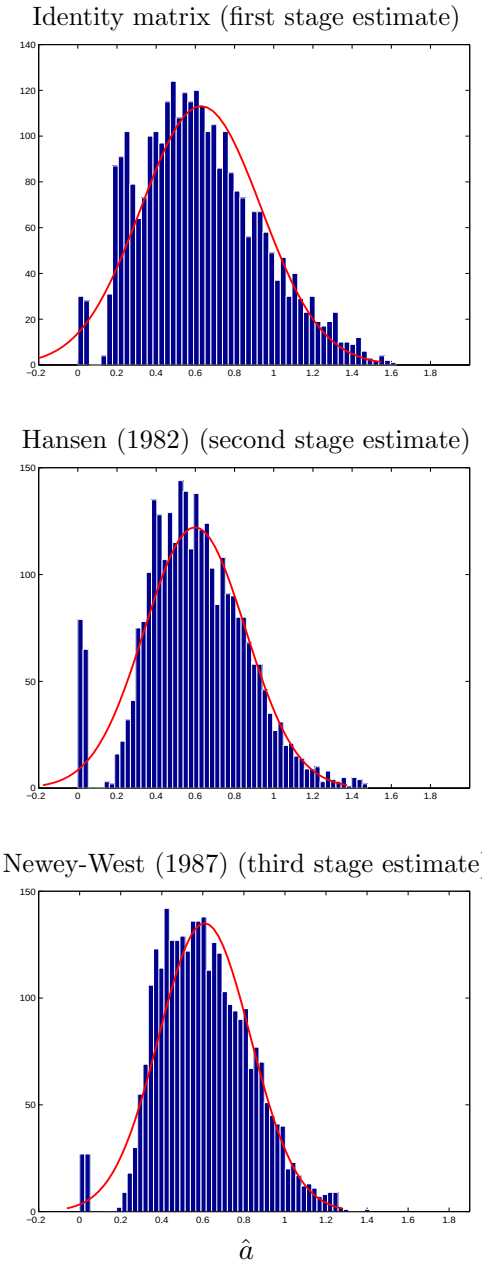


Figure 2.6: Sample distribution of estimated a SNJD parameter for different weighting matrices. Simulation values are $\mu = 0.02 \times 10^{-2}$, $\sigma = 1.26 \times 10^{-2}$, $\lambda = 4.00 \times 10^{-2}$, $\beta = 10.00 \times 10^{-2}$ and $a = 0.60$. The random variable U is lognormally distributed with mean $E[U] = 0$. Note that the axes OX are equally scaled in all the graphs.

understand, as much as possible, the behavior of the total variance involved in the process[38].

As Figure 2.7 reports, the majority of the observations are clustered on the lower function values for both pairs of variables in both plots. This may be intended as a sign of convergence of the numerical procedure used. It is possible to observe as the Identity matrix produces some clusters of observations in a function value area (around 0.05) bigger than the minimum reached (almost zero).

Finally, we can see that the Newey-West (1987) estimate observations tend to be grouped in a smaller region than the other estimates. This result can be understood in terms of a) independence between pairs of parameters and b) efficiency of the estimates employed.

Comments on the SNJD results

Similarly to the equally-named section on JD process, we proceed now to analyze the results obtained for the SNJD process.

Basically, some questions arise from the results in Table 2.7: firstly, the observed bias in the λ and σ results; secondly, the lost of efficiency in the parameter λ when the Hansen (1982) and Newey-West (1987) weighting matrices are introduced in the GMM estimates.

To provide some additional evidence regarding the bias question, Table 2.8 provides the different percentiles obtained for the SNJD estimates by using the Newey-West (1987) weighting matrix. Our intention is to offer a complete description of the behavior of the most suitable weighting matrix estimate for these cases[39].

The first column in this Table displays the parameter and its true value (in parenthesis) while the second column shows the different values of the parameter a we are considering in our analysis. Third and fourth columns report the mean and root mean squared error (RMSE) values, respectively. Finally, the percentile block displays these values for certain percentage values.

Some results in Table 2.8 seem to corroborate those included in Table 2.7. For example, the bias in the parameter σ is reduced when the coefficient a increases. This result is also accompanied by a substantial increase in the λ

[38] We do not find substantial differences by using another argument for the function value as, for instance, β. Anyway, the entire results used to build these and previous graphs are available upon request.

[39] As mentioned through the paper, the Newey-West (1987) weighting matrix takes into account the presence of serial correlation, as happens in these cases

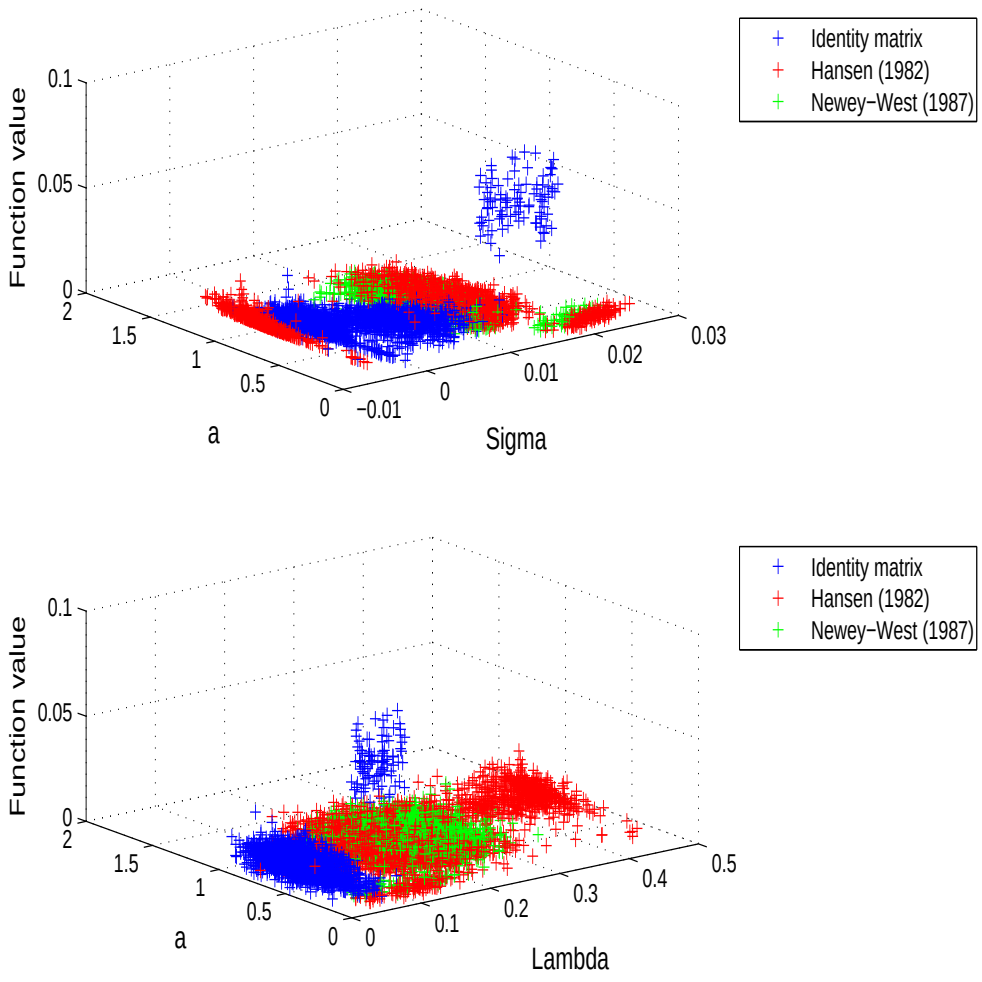


Figure 2.7: Function values of the GMM estimate at points of convergence for different pairs of variables. Figure represents the results for a SNJD model estimation under different specifications of the weighting matrix: Identity (blue colour), Hansen (1982) (red colour) and Newey-West (1987) (green colour). Upper and lower plots display the function values of the estimate J_T on expression (2.2.2.3) as a function of a, σ and a, λ, respectively. Simulated parameters are $\mu = 0.0002$, $\sigma = 0.0126$, $\lambda = 0.0400$, $\beta = 0.1000$ and $a = 1.00$. The jump size U is lognormally distributed of the form $U = e^{-\beta^2/2+\beta\varepsilon} - 1$, with $\varepsilon \sim N(0,1)$ and mean $E[U] = 0$.

Table 2.8: Percentiles for the third stage GMM estimates of the SNJD process

Parameter ($\times 10^{-2}$)	a	Mean ($\times 10^{-2}$)	RMSE ($\times 10^{-2}$)	Percentile ($\times 10^{-2}$)						
				2.5%	5%	10%	50%	90%	95%	97.5%
μ	0.2	0.05	0.05	-0.03	-0.01	0.00	0.05	0.10	0.12	0.13
(0.02)	0.6	0.05	0.04	-0.03	-0.01	0.00	0.05	0.09	0.10	0.11
	1.0	0.04	0.04	-0.03	-0.01	0.00	0.04	0.08	0.09	0.11
σ	0.2	1.72	0.49	1.36	1.43	1.50	1.72	1.95	2.03	2.08
(1.26)	0.6	1.55	0.32	1.29	1.34	1.40	1.55	1.71	1.76	1.81
	1.0	1.37	0.18	1.04	1.12	1.18	1.38	1.53	1.58	1.62
λ	0.2	8.87	6.01	2.80	3.42	4.34	8.79	13.42	14.89	16.53
(4.00)	0.6	17.15	13.65	10.23	11.62	12.98	17.03	21.77	23.13	24.45
	1.0	23.76	20.78	15.07	16.90	18.52	23.81	29.36	30.65	31.86
β	0.2	7.72	2.60	5.53	5.88	6.28	7.63	9.35	9.95	10.44
(10.00)	0.6	6.57	3.53	4.98	5.24	5.55	6.56	7.60	7.92	8.19
	1.0	6.18	3.89	4.82	5.07	5.32	6.16	7.06	7.36	7.65
a	0.2	0.48	0.59	0.04	0.06	0.10	0.21	1.39	1.59	1.75
(right)	0.6	0.61	0.22	0.25	0.30	0.35	0.59	0.90	0.99	1.08
	1.0	0.83	0.27	0.45	0.52	0.59	0.84	1.07	1.14	1.20

GMM estimates for the SNJD process using Newey-West (1987) as weighting matrix. We assume a jump size U lognormally distributed of the form $U = e^{-\beta^2/2+\beta\varepsilon} - 1$, with $\varepsilon \sim N(0,1)$. The first column contains the parameters of the SNJD process (true simulation values in parenthesis). Second to fourth columns show the a parameter case and the mean and root mean squared error (RMSE) of the estimated parameters using the GMM methodology, respectively. Finally, the columns five to the end exhibit the different percentiles from the sample distribution of the estimates obtained.

bias; roughly speaking, the whole distribution of the parameter λ is moved to the right as a increases. It seems that the parameters σ and λ go in opposite directions.

Regarding the parameter β, Table 2.8 shows that its bias increases with the parameter a although this increase tends to be small. Finally, the dispersion of the parameters σ and β decreases with the parameter a; nevertheless, an stable pattern for the parameter λ is not detected.

According to Table 2.8, the results about the parameter a do not differ from those in Table 2.7: the most precise and unbiased estimate is obtained for $a = 0.6$ and the smallest dispersion is achieved for $a = 1.0$. for $a = 0.20$, the methodology seems to capture the value of this parameter (the median is equal to 0.21), but a non-negligible amount of a estimates that are far from the true value tends to corrupt the results[40].

What about the correlation between estimates? Table 2.9 displays the correlation matrices obtained for the estimates used in Table 2.8, where the simulation parameters are $\mu = 0.0002$, $\sigma = 0.0126$, $\lambda = 0.0400$ and $\beta = 0.1000$. The results included in these correlation matrices correspond to the cases $a = 0.20, 0.60, 1.00$.

The main features of the results included in this Table are the following:

- A negative correlation between the parameters σ and λ is observed in the three cases.
- This negative pattern is also observed between the parameters λ and β whose correlation coefficients are significant at the 95%-confidence level although they decrease with a.
- The correlation between the parameters σ and β has not a constant sign and, as previously, it also decreases with the value of a.
- Finally, an increase in the value of the parameter a implies a more negative correlation with σ in direct contrast to the correlation with λ that moves from negative (-0.2246) to positive (0.3455) values.

Just to provide a visual idea of former results, Figure 2.8 exhibits the point clouds of different pairs of parameters. The graphs in Figure 2.8 correspond to the GMM results of the SNJD process using the Newey-West (1987) weighting matrix.

As can be seen, the negative correlation between σ and λ and between λ and β is clear as we can appreciate a negative slope in both cases. Regarding

[40] This fact seems to be confirmed by a mean value of a that is bigger than the median one.

Table 2.9: Correlation matrices for the different SNJD estimates

$a = 0.20$	μ	σ	λ	β	a
μ	1.0000				
σ	0.0968	1.0000			
λ	**-0.0112**	-0.6748	1.0000		
β	-0.0832	0.6603	-0.6031	1.0000	
a	**0.0333**	0.1993	-0.2246	0.1422	1.0000

$a = 0.60$	μ	σ	λ	β	a
μ	1.0000				
σ	0.1426	1.0000			
λ	0.0396	-0.3521	1.0000		
β	-0.1314	0.2185	-0.3255	1.0000	
a	**0.0034**	-0.1643	0.2248	-0.0763	1.0000

$a = 1.00$	μ	σ	λ	β	a
μ	1.0000				
σ	0.1438	1.0000			
λ	**0.0094**	-0.3619	1.0000		
β	-0.0728	-0.0714	-0.2386	1.0000	
a	**-0.0323**	-0.4971	0.3476	-0.0474	1.0000

Correlation matrices of the GMM estimates for the SNJD process using Newey-West (1987) as weighting matrix. The parameter vector is $\Lambda = (\mu, \sigma, \lambda, \beta, a)' \in R^5$, and the simulation parameter values are $\mu = 0.0002$, $\sigma = 0.0126$, $\lambda = 0.0400$ and $\beta = 0.1000$. Results correspond to the cases $a = 0.20, 0.60, 1.00$. Parameters in bold are not significant at a 95%-confidence level.

the correlation between a and σ (λ), we can see a negative (positive) slope although this pattern is not so clear as in the previous cases. Finally, it seems that the correlation of a with any of these parameters presents a higher dispersion than that in the remaining pairs.

Up to this point, it seems that there exists enough evidences about the joint behavior of the parameters σ and λ in the sample under study. This may reveal a similar effect to that previously cited in the JD process: a volatility transference pattern appears to be detected from the Brownian part of the process to its jump part. One possible explanation to this fact can be that the higher the increase in the parameter a, the higher the total variance of the SNJD process.

Then, in some sense, the GMM estimate tends to allocate the variance to the jump part at cost of the Brownian component, an argument that was found in the results for the first stage estimate in Table 2.7. Then, in absence of an efficient weighting matrix, the GMM estimate tends to assign the variance of the process to the jump part (the value of λ increases with the parameter a).

This last result may seem paradoxical: as equation (2.2.3.3) reveals, the total variance of a SNJD process does not depend on the parameter a. However, it should be intuitive that a higher value of the parameter a introduces more variance to the process[41].

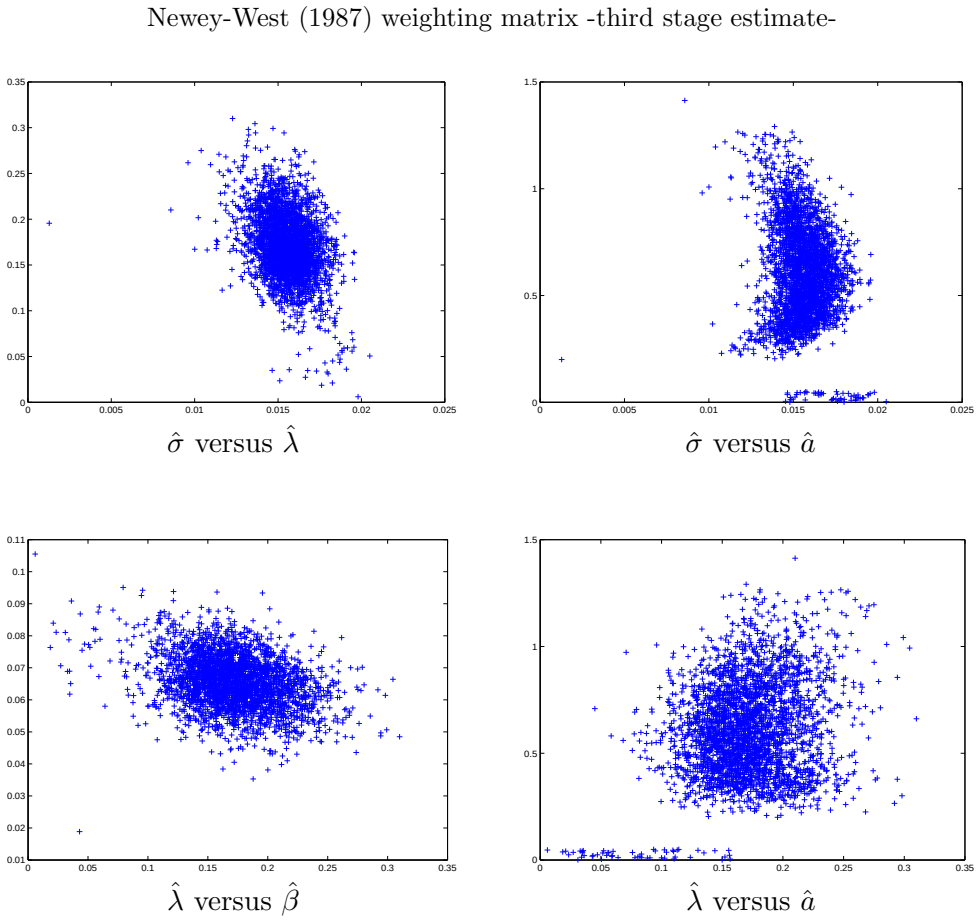
But remember that the moment expressions (2.2.3.2)-(2.2.3.5) have been obtained by using numerical approximations[42]: the higher a, the worst results. Table 2.10 illustrates this effect: we are comparing the sample annual volatility of the SNJD process simulated with its theoretical values for different values of a. As this Table exhibits, the sample volatility tends to increase with the parameter a. Additionally, the difference between estimated and theoretical volatility increases with the parameter a.

It seems that the key point to understand this puzzle comes from different sources: firstly, the numerical approximations done to derive the moments, as previously deduced from Tables 2.3 and 2.4; secondly, the GMM itself.

[41] A possible explanation can be that a higher value of the parameter a leads to a higher decay. In the extreme case of a huge value for a, our sample is mainly constituted by Brownian data with infrequent, *successive* (and with opposite sign) jump observations spread in. To put it another way, a SNJD sample histogram with a high value of a ($a = 1.0$, for instance) will include many observations clustered around the mean of the process and many infrequent observations far away this mean (almost the double than in a JD process). For example, an upward jump will be followed by a downward response with almost the same value. As a result, more anomalous data is introduced in the sample, leading to an increase in the total variance of the process.

[42] See Chapter 1 for details.

Figure 2.8: Points clouds for different GMM estimates of the SNJD process



Different point clouds for GMM estimates of SNJD process parameters. Graphs correspond to the estimates obtained by using the Newey-West (1987) matrix as weighting one in the GMM estimates (third stage estimate). Simulation parameters are $\mu = 0.0002$, $\sigma = 0.0126$, $\lambda = 0.0400$, $\beta = 0.1000$ and $a = 0.60$. The jump size U is lognormally distributed of the form $U = e^{-\beta^2/2+\beta\varepsilon} - 1$, with $\varepsilon \sim N(0,1)$ and mean $E[U] = 0$. Finally, "$\hat{\sigma}$ versus $\hat{\lambda}$" refers to the parameter $\hat{\sigma}$ on the axis OX and the parameter $\hat{\lambda}$ on the axis OY (and similarly in the remaining graphs).

Table 2.10: Sample and theoretical standard deviation for SNJD process

Std deviation	a		
(%)	0.20	0.60	1.00
σ_{theo}	45.44	49.55	52.37
σ_{sample}	38.63	41.10	43.08
Diff	6.81	8.45	9.29

Sample and theoretical annualized standard deviation for the SNJD process. Daily variance has been computed as basis to yearly standard deviation. σ_{theo} has been computed by substituting the mean values of the GMM (third stage) estimates of the SNJD process(see Table 2.7) in expression (2.2.3.3). σ_{sample} has been computed from the mean daily variance of the SNJD sample used. Simulation parameters are $\mu = 0.0002$, $\sigma = 0.0126$, $\lambda = 0.0400$ and $\beta = 0.1000$. The jump size U is lognormally distributed of the form $U = e^{-\beta^2/2+\beta\varepsilon} - 1$, with $\varepsilon \sim N(0,1)$ and mean $E[U] = 0$.

Regarding this second issue, we can mention the following potential sources of problems:

- The choice of the moment conditions set (Zhou, 2001).
- The efficiency of the estimates under the joint estimation of data (Ait-Sahalia, 2004).
- The weighting matrices (Zhou, 2001): it seems evident that the inclusion of these matrices leads to re-allocate the total variance, but this is not efficient, as deduced from Table 2.7.

2.6 Cross samples estimation

Once studied the ability of the GMM methodology to estimate the parameters of process in expression (2.2.3.1), it can be interesting to analyze the case where the simulated sample differs from the model that we try to estimate. For example, imagine a sample generated by a SNJD model, what will be the parameter values obtained if we estimate a JD model?

To answer this question, we have carried out two main experiments:

1. We have simulated a SNJD sample and we have looked for the values of a JD model.

2. We have proceeded in a inverse manner, generating a JD sample and estimating the parameters of a SNJD model.

Our intention here is to understand the behavior of the estimate in two alternative cases of what is commonly known as model error:

1. Misspecification of the theoretical model (SNJD sample - JD model estimation).

2. Overparametrization of the theoretical model (JD sample - SNJD model estimation).

The purpose of this section is to check the ability of the GMM estimate to deal with samples that differ slightly from the model under estimation. Although it is not directly tackled here, the question addressed here is connected with the consequences, in terms of the estimation parameters, of a *myopic* investor who firmly believes that the market follows a specific model when actually it does not happen.

2.6.1 Results

As it was previously noticed, we have carried out two main experiments: in first place, a JD model has been estimated from a simulated SNJD sample. Secondly, a simulated JD sample has been used for estimating the parameters of a SNJD model. Table 2.11 presents the results obtained for those experiments.

For both experiments, the first column in Table 2.11 displays the sample statistics obtained. The remaining columns exhibit the results obtained for different variables under study.We have used the efficient weighting matrix in both experiments (the Newey-West (1987) matrix for the SNJD sample and the Hansen (1982) weighting matrix for the JD samples). Again, as it is becoming standard through the text, our experiments comprise simulations of 3,000 paths with 1,800 data each one. Finally, the simulation true values are displayed into parenthesis.

Before presenting the results obtained in both experiments, it can be interesting to discuss, as much as possible, what we expect. In principle, it seems to be plausible that the estimates of the different parameters will not be far away from their simulation values. This closeness can be deduced from equations (2.2.3.2)-(2.2.3.5), showing that JD and SNJD models have

Table 2.11: Cross sample estimations

First experiment: JD model - SNJD sample

Statistic	Parameters				
	$\mu \times 10^{-2}$ (0.02)	$\sigma \times 10^{-2}$ (1.26)	$\lambda \times 10^{-2}$ (4.00)	$\beta \times 10^{-2}$ (10.00)	a (0.06)
Mean	0.04	1.24	7.45	8.03	–
Median	0.04	1.26	7.42	8.03	–
Min	-0.11	0.38	5.45	3.63	–
Max	0.16	1.83	10.38	11.35	–
MSE	0.00	0.00	0.12	0.05	–
RMSE	0.04	0.17	3.49	2.17	–

Second experiment: SNJD model - JD sample

Statistic	Parameters				
	$\mu \times 10^{-2}$ (0.02)	$\sigma \times 10^{-2}$ (1.26)	$\lambda \times 10^{-2}$ (4.00)	$\beta \times 10^{-2}$ (10.00)	a (0.06)
Mean	0.03	1.79	3.93	7.24	0.34
Median	0.03	1.78	3.39	8.16	0.04
Min	-0.09	1.18	0.00	0.00	-1.00
Max	0.17	2.37	14.69	16.86	1.00
MSE	0.00	0.00	0.09	0.19	0.30
RMSE	0.04	0.55	3.02	4.39	0.54

GMM estimations for the JD and SNJD models under SNJD and JD samples, respectively. The first and second experiments have been carried out by using the Newey-West (1987) and the Hansen (1982) efficient weighting matrices, respectively. True simulation values are in parenthesis. This Table shows mean, median, minimum, maximum, mean squared error (MSE) and root mean squared error (RMSE). The response function for the SNJD model is $h(t-\tau_k) = e^{-a(t-\tau_k)}$. We have assumed a jump size U lognormally distributed of the form $U = e^{-\beta^2/2+\beta\varepsilon} - 1$, with $\varepsilon \sim N(0,1)$ with mean $E[U] = 0$.

similar central moments for a *certain* range of a values[43]. The further away from these approximations we are, the worse results we will have.

Results from Table 2.11 seem to confirm our guess. Focusing on the first experiment, we observe that the estimates obtained for a JD process under presence of serial correlation[44] are close to the true SNJD simulation values. Looking at the mean and median values obtained, the sample distributions for all the variables seem to be symmetrical[45]. On the whole, the results obtained here apparently confirm our previous expectations: within the range of values for which the numerical approximations work, the results are not extremely different. We will come back to this point in the next subsection.

Concerning the second experiment in Table 2.11, the results go in the same direction that those obtained for cases of low serial dependence (for instance, $a = 0.2$) in Table 2.7. For instance, we observe an asymmetric distribution for the sample estimates, confirmed by the huge differences among median and min-max values. Additionally, high values of σ are also detected. In general, we are inclined to think that in the cases of low (or null) autocorrelation in the sample, the estimate is not capable of distinguishing between extremely low ($a = 0.00$) or high ($a = 10.00$) values of the parameter a. This circumstance may be explained by the order of the a parameter involved in: we have defined some moment equations based on a certain range of values for the parameter a, and $a = 0.00, 10.00$ could be considered outside that interval.

2.6.2 Additional issues on the JD process estimation in samples with autocorrelation

Once examined the results in Table 2.11, it can be interesting to analyze how the serial dependence in the sample affects the estimation of a JD model. The conditions to develop such an experiment are optimal because we have a model (the SNJD one) that permits us to control the degree of persistence in returns. To our knowledge, this question has not yet been addressed in the financial literature.

Table 2.12 shows the results of the estimation for a JD process using a SNJD simulated sample. Basically, this table is an extension to more

[43] Chapter 1 shows that the first moments of JD and SNJD models are equal under a set of numerical approximations. See Section 5 on Chapter 1 for a complete discussion about this point

[44] A value of the parameter a equal to 0.6 implies that first and second lags autocorrelation values are, on average, significant.

[45] This is not the case for the σ value, that presents a slightly asymmetrical sample distribution

cases of the first experiment in Table 2.11. Again, we have used the Newey-West (1987) estimate in all our estimations, due to the presence of serial correlation in data. The simulated sample comprises 3,000 paths of 1,800 data each one.

The first column in Table 2.12 displays the different variables under study and their simulated values in parenthesis. The second column shows the different values for the parameter a. Third and fourth columns exhibit the mean and root mean squared error (RMSE) obtained. Finally, columns fifth to eleventh show different percentiles from the estimates distribution.

Some conclusions arise from inspection of Table 2.12. Firstly, it seems that the estimate distributions are symmetrical, as it is deduced from the mean and median values and the distance between the pairs of percentiles 10-90%, 5-95% and 2.5-97.5% with respect to their median value. Secondly, it is observed a general displacement *to the right* of the estimates distributions as the parameter a increases[46]. This general tendency does not appear to apply to the first percentiles of the parameter σ when $a = 10.00$.

In line with this last feature, a possible explanation to our results could be that as a increases, the total variance in the sample also increases[47]. From our point of view, the JD estimates tend to capture this excess of variance by increasing the parameters associated to the JD variance, as those of the volatility of the Brownian process σ, and the number of jumps λ and their amplitudes β.

2.7 Conclusions

This chapter has analyzed the estimation in the time domain of the SNJD model. We have used an standard methodology, the GMM procedure posited by Hansen (1982), to estimate the parameters of a SNJD process. To our knowledge, this is the first study that tackles the estimation of a SNJD model by using the GMM estimate.

We have followed a sequentially procedure to estimate the SNJD model, from simplest to more complicated cases. In a first stage, we have checked the performance of the GMM methodology by using samples not perturbed by Brownian noise; secondly, we have carried out an estimation of the JD

[46]The parameter μ is an exception, but it will not be analyzed here. This is mainly due to our interest through this chapter on the terms that affect the total variance of the JD or SNJD processes.

[47]Unfortunately, this increment in the total variance in the SNJD sample is not detected by our equations (see expression (2.2.3.3) on this chapter) because of the numerical approximations involved in.

Table 2.12: Percentiles for cross sample estimations

Parameter ($\times 10^{-2}$)	a	Mean ($\times 10^{-2}$)	RMSE ($\times 10^{-2}$)	Percentile ($\times 10^{-2}$)						
				2.5%	5%	10%	50%	90%	95%	97.5%
μ	0.2	0.04	0.05	-0.04	-0.03	-0.01	0.04	0.10	0.11	0.12
(0.02)	0.6	0.04	0.04	-0.03	-0.02	-0.01	0.04	0.09	0.11	0.12
	1.0	0.04	0.04	-0.03	-0.02	0.00	0.04	0.09	0.10	0.12
	10.0	0.04	0.04	-0.02	-0.01	0.00	0.04	0.08	0.09	0.10
σ	0.2	1.05	0.30	0.54	0.67	0.79	1.08	1.26	1.31	1.36
(1.26)	0.6	1.24	0.17	0.84	0.93	1.03	1.26	1.43	1.47	1.51
	1.0	1.32	0.19	0.89	1.00	1.10	1.35	1.50	1.55	1.58
	10.0	1.30	0.24	0.72	0.85	1.01	1.35	1.54	1.59	1.61
λ	0.2	6.88	2.91	6.06	6.20	6.38	6.90	7.36	7.49	7.61
(4.00)	0.6	7.45	3.49	6.50	6.67	6.87	7.42	8.10	8.28	8.52
	1.0	7.89	3.94	6.74	6.96	7.16	7.82	8.74	9.08	9.40
	10.0	8.48	4.55	7.10	7.32	7.53	8.39	9.54	9.95	10.23
β	0.2	8.06	2.18	6.17	6.49	6.83	8.05	9.30	9.63	9.96
(10.00)	0.6	8.03	2.17	6.32	6.55	6.87	8.03	9.16	9.51	9.80
	1.0	8.20	2.03	6.52	6.79	7.02	8.18	9.40	9.74	10.10
	10.0	9.38	1.10	7.68	7.94	8.26	9.33	10.57	10.92	11.25

Percentiles for GMM estimations for a JD model under SNJD sample. All the estimations have been carried out by using the Newey-West (1987) weighting matrix. The first column displays the variables under study and their true simulation values (in parenthesis). Additionally, this table contains four blocks of parameters, each one containing four rows. For the different a cases, each row displays the sample statistics (mean, root mean squared error and percentiles) obtained. Finally, the response function used in the SNJD model is $h(t-\tau_k) = e^{-a(t-\tau_k)}$. We have assumed a jump size U lognormally distributed of the form $U = e^{-\beta^2/2+\beta\varepsilon} - 1$, with $\varepsilon \sim N(0,1)$ with mean $E[U] = 0$.

processes with the purpose of being familiar with the GMM results for this type of processes; finally, some results about the SNJD estimation are provided. In addition to this issue, the last section of this chapter has been devoted to analyze possible problems of misspecification and overparametrization in the JD and SNJD models under SNJD and JD simulated samples, respectively.

As an overall conclusion, the results obtained in this chapter seem to suggest that the GMM methodology is capable of estimating the parameters of a SNJD model. However, this performance of the GMM technique must be taken carefully, because of the following reasons:

- An observed bias is detected in the parameters associated with the total variance of the process.
- The numerical approximations involved in the different moment expressions appear to affect dramatically the results.
- Apparently, the inclusion of more efficient weighting matrices in the SNJD estimation does not improve substantially the quality of the estimates obtained.
- Unfortunately, with the intention of focusing on a controllable set of experiments, we have not extended our study to changes in other parameters of the SNJD model than the a parameter.

As a final remark, it can be worthy to remark that the GMM methodology is a numerically stable procedure in all the computations carried out.

Chapter 3

Estimation of the SNJD process in the Frequency Domain

This chapter[1] is devoted to the estimation of the SNJD process in the frequency domain. We introduce the basic theory of Spectral Analysis, providing some useful tools for carrying out the estimation using this technique. Finally, an intensive Monte Carlo study is provided to explore the capability of the posited procedure to estimate the parameters of the SNJD model.

3.1 Introduction

The basic question that addresses this chapter is the estimation of the SNJD process in the frequency domain. In more detail, we are interested in the kind of solutions that could provide to our estimations a commonly used methodology among econometricians: the estimation in the frequency domain, also named Spectral Analysis.

[1]We want to thank Jose M. Vidal and Alfonso Novales for their helpful comments. I also acknowledges financial support from the Plan Nacional de I+D+I (project BEC2003-02084) and project GIU 06/53 of the University of the Basque Country and Basque Government, and especially to Jose M. Usategui. A previous draft of this chapter have been presented under the name "Estimation of Jump-Diffusion Processes with Shot-Noise Effects" in the XIV Foro de Finanzas AEFIN. I have also benefited from the comments of participants in seminars at Universidad Complutense de Madrid, Universitat de les Illes Balears and Universidad Carlos III de Madrid. Finally, part of this chapter was developed while I was visiting the Department of Business Administration at Universidad Carlos III. I am sincerely grateful to them.

Intuitively, it is clear that a stochastic process can be understood as a sequence of innovations over the time (Hamilton, 1994). Many economic variables are provided as series of events happening in some equally spaced time intervals. For example, the Growth Domestic Product (GDP) or the Consumer Price Index (CPI) are reported with quarterly or monthly frequencies, respectively. In Finance, it is common to show the evolution of a stock index as the time series path of daily or weekly closing prices. We name as *time domain* this kind of approach where the economic variables are displayed as a sequence of data through time.

An alternative way of thinking about the time series is provided by the Spectral Analysis or analysis in the *frequency domain*. Roughly speaking, the idea consists on decomposing a time series extracting their repeated ("common") components and, next, ordering these components in a certain manner. As we will see, these periodic elements will be grouped attending to the information content within a particular time span, it is said, attending to their *frequencies*. Providing some insights that will be detailed later, these cycles can be interpreted as the contribution of each cyclical component to the total variance of the process.

Actually, former idea of understanding the behavior of a time series in this way is not new. In Economics, we are used to hear many expressions related to the "business cycle" or the "procyclicality" of some activities or goods. The Spectral Analysis provides useful tools for adopting this particular scheme.

Regarding the financial area, we can wonder about the interest of this methodology. Basically, two main questions arise:

1. Firstly, we can question about what is the gain of using this approach, especially when compared with the existing (and developed) econometric theory in the time domain.

2. Moreover, to a certain degree, the tools of time domain analysis are usually employed in the financial econometrics of continuous-time processes. So, can we learn something with the Spectral Analysis that we have not yet known?

Concerning to the former question, this is a point that we will try to answer during this chapter. With respect to the second one, we think that some interesting questions on the empirics have not been completely solved in the previous chapter, and the analysis in the frequency domain could result in a interesting answer.

This chapter analyzes the SNJD model using a Spectral Analysis methodology. From a theoretical point of view, we propose an econometric framework based on the frequency analysis to estimate the SNJD processes, by taking profit of their ability to exhibit serial dependence. Additionally, from an empirical point of view, we provide some results about the performance of this technique by means of a extensive Monte Carlo study.

The literature on the Spectral Analysis procedure is quite large. For example, Hamilton (1994) provides an useful and extensive introduction to the spectral theory while more details can be found in some reference textbooks as those of Hannan (1970) or Priestley (1981), which also contains several references about the estimation in the frequency domain. By contrast, the literature about estimation of shot noise processes is not so large and, basically, our work is inspired in the results of Bevan *et al* (1979).

The structure of this chapter is as follows. Section 2 introduces the econometric framework and provides some of the intuitions behind this framework. Section 3 performs an intensive Monte-Carlo study. Some comments and results are included in Section 4. Finally, Section 5 summarizes the main conclusions and suggests some lines for further research.

3.2 Econometric Framework

This section describes the econometric framework proposed. Firstly, we provide some additional intuitions about the spectrum without any special mathematical formulation, with the purpose of fixing some ideas. Then, we define the estimate and some necessary mathematical tools employed[2].

3.2.1 The intuition behind

As it was previously cited, the frequency domain methodology tackles the problem of the estimation of an stochastic process from another point of view. Here, the idea consists on analyzing the contributions of the *cyclical* components of the series: by means of using certain mathematical transformations, we will be able to extract from a certain time series those patterns that are repeated in a periodical way. To put it another words, we will display in an ordered form the contributions of the *repeated* information that appears over time. Mainly, this information is presented using a useful tool in Spectral Analysis named *sample spectrum*.

[2]This section is mainly based on Hamilton (1994) and Hannan (1970). The idea about the estimate employed here is taken from Bevan *et al* (1979).

With the purpose of clarifying former ideas, Figure 3.1 displays the time and frequency domain representation of an economic variable, the adjusted Consumer Price Index (CPI) for Spain[3]. The sample is composed by monthly observations from 01/01/1997 to 01/08/2007. The time series path for the level of the CPI is displayed in the upper graph in Figure 3.1. Additionally, the sample spectrum[4] of log-increments of CPI series is represented in the lower graph.

The lecture of the CPI series through time is usual in Economics: Figure 3.1 shows that CPI increases (upper figure) over time and that, from observation 60 to the end, the CPI path seems to exhibit a waved effect, with successive up-down movements, keeping the upward trend.

On the other hand, the lower graph displays the sample spectrum. The goal here is to determine how important cycles are in accounting for the behavior of the series (Hamilton, 1994). In other words, we are interested in looking at the contributions of the different cyclical components of the series.

To analyze the data, we have decomposed the total sample T of data on small pieces of different lengths; basically, we have ordered the data in periodical parts named frequencies, denoted by w_j, of size $2\pi j/T$, with $j = 1, 2, ..., (T-1)/2$. The lower frequency we have, the higher the time span is: for example, frequencies closer to zero correspond to the information contained in yearly, quarterly observations, and so on.

The axis OX represents this ordered set of frequencies. Finally, the axis OY of the sample spectrum plot in this figure just establishes the degree of importance of each frequency to the total part of the spectrum. More specifically, the interpretation is as follows: the higher the area under the graph around a certain frequency, the higher portion of the global variance of the process is explained by that frequency.

To illustrate former ideas, we will come back to the sample spectrum graph in Figure 3.1. For example, looking at this figure it is observed how the first peak happens around the value $j = 12$ and is also followed by the maximum peak in the figure, that happens around $j = 21$. This last would say that something is happening at $j = 21$ case that partially explains the variability of the series. As $w = 2\pi j/T$, the period corresponds to T/j. The sample includes monthly data with length $T = 128$, so it corresponds to a period of $128/21 \simeq 6.09$ months. One possible explanation to our result

[3]For the sake of the exposition, we have chosen a macro variable. The data has been taken from the Instituto Nacional de Estadística (INE), the national bureau of statistical services for Spain. Web page http://www.ine.es.

[4]Details about the sample spectrum estimate will be provided later.

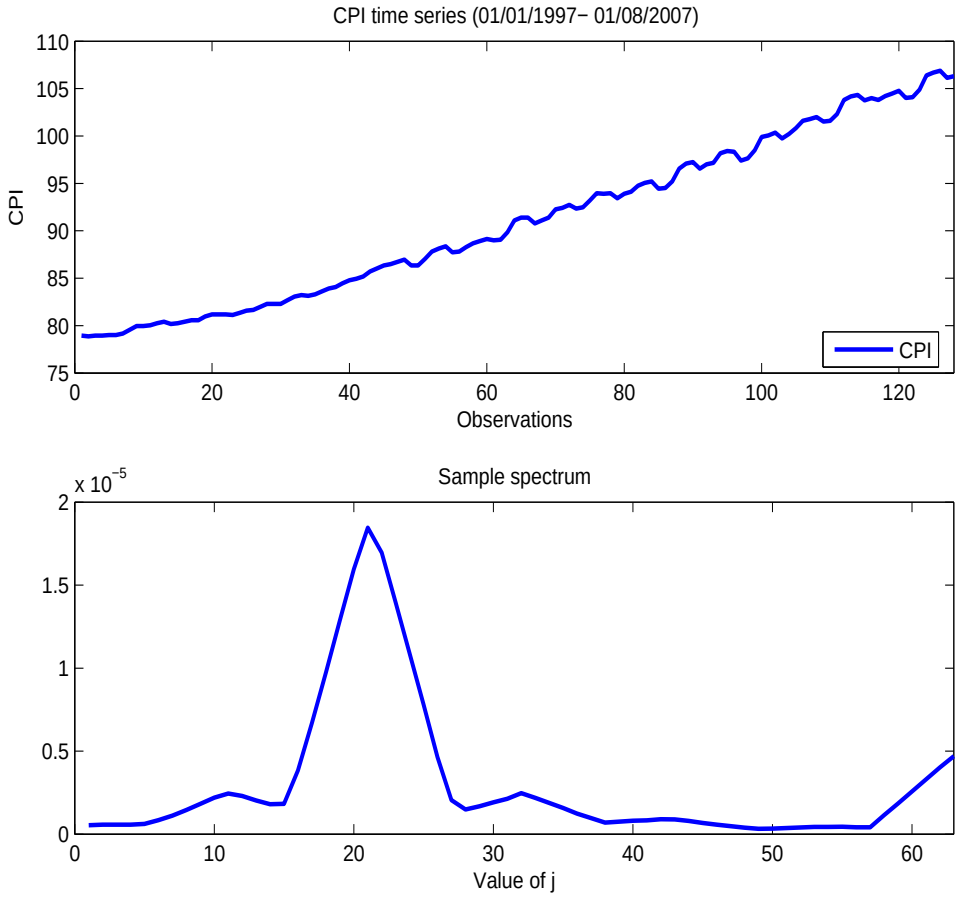


Figure 3.1: Time series path and spectrum for Consumer Price Index (CPI) of Spain. Upper graph displays the evolution over time of the CPI in Spain for the period 01/01/1997 to 01/08/2007. Lower graph exhibits the sample spectrum for CPI log-increments.

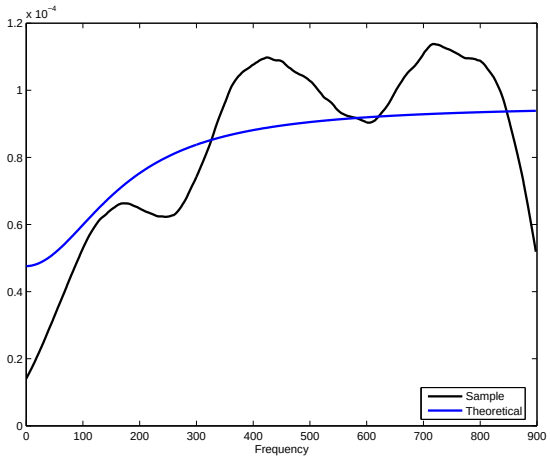


Figure 3.2: Sample and theoretical spectrum for a SNJD process. The sample spectrum has been obtained using a non-parametric estimate posited in Hamilton (1994) with bandwidth parameter fixed to $h = 100$. The SNJD model response function is $h(t-\tau_k) = e^{-a(t-\tau_k)}$ with the random variable U lognormally distributed as $U = e^{-\beta^2/2+\beta\varepsilon} - 1$, $\varepsilon \sim N(0,1)$ and mean $E[U] = 0$. Simulated (annualized) parameters are $\sigma = 0.05$, $\lambda = 10$, $\beta = 0.10$ and $a = 0.60$.

seems to be that a big part of the total variation of the CPI for the Spanish case could be explained in the semi-annual periods[5].

As it has been presented, the spectrum could provide us an additional source of information of what is happening within the dynamics of the process. Concerning to the SNJD model (and, more specifically, to their exponential decay response version), the type of related spectrum is similar to that displayed in Figure 3.2, where the spectrum is clustered on higher frequencies. This may reflect that contribution of short range effects to the global variance of the process must be taken into account when analyzing the behavior of the process.

3.2.2 The procedure

The spectral analysis describes the value of a stochastic process as a weighted sum of periodical functions[6]. These functions are of the sine- or cosine-type (denoted as $\sin(wt)$ and $\cos(wt)$, respectively) where w denotes a particular

[5]For the Spanish case, this issue coincides with the sales periods in January and July and, then, with semi-annual periodicity.

[6]Mathematically, the spectral analysis can be considered as a change of basis from a canonical space to the Fourier one.

frequency, defined as $w_j = 2\pi j/T,\ j = 1, 2, ..., (T-1)/2$ (see Hamilton (1994) for details).

In short, the approach used here involves the following steps:

1. The autocovariance function of the process is computed.
2. The Fourier transform of the autocovariance function must be calculated to obtain the population spectrum.
3. Then, an estimate of the population spectrum must be introduced.
4. Finally, the population spectrum and its sample counterpart must be related by any kind of estimate.

Concerning to the first and second points, let $\Theta \subset \mathbb{R}^n$ be a n-parameter vector of the model. Given the process $Z(t)$, we define the τ-autocovariance of the process $Z(t)$ as

$$c(\tau;\Theta) = Cov(Z(t), Z(t+\tau)) = E[Z(t)\, Z(t+\tau)],$$

Then, the population spectrum is given by the Fourier transform of the autocovariance function $c(\tau;\Theta)$

$$s(w;\Theta) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-iw\tau} c(\tau;\Theta)\, d\tau$$

Focusing on the third and fourth points aforementioned, our estimated parameters will be obtained by two different ways: firstly, our strategy will be to minimize the (squared) distance between the population spectrum and a kind of non-parametric spectrum estimate for all frequencies; this will be our *naive* estimate. Secondly, we will use the Whittle (1953) estimate as an alternative to the former non-parametric estimate.

3.2.3 The estimates

We detail now the two estimates (naive and Whittle) used, explaining their main features and providing also additional references about them.

The naive estimate

As its name denotes, the naive estimate is simple: basically, it consists on minimizing the population spectrum $s(w_i;\Theta)$ with any class of estimated

spectrum[7], $\widehat{s}_Z(w_i)$, over a set of n frequencies $w = \{w_1, w_2, \cdots, w_n\}$ for a certain parameter vector Θ,

$$\min_{\theta} \varphi = \min_{\theta} \sum_{j=1}^{n} [\, \widehat{s}_Z(w_j) - s(w_j; \Theta) \,]^2 \qquad (3.3.2.1)$$

This approach is inspired on Bevan *et al* (1979). However, our estimate differs on that from Bevan *et al* (1979) in

1. The minimization problem, since they minimize the logarithm of the spectra, instead of the spectra directly, and
2. The class of the population spectrum estimate employed.

Concerning to the minimization problem, Bevan *et al* (1979) point out (among other reasons) that applying logarithms to the spectra is a procedure to stabilize their variance. However, we encountered some numerical problems due to the negative sign of the theoretical population spectrum for a given set of parameters. This situation led us to modify the form of the naive estimate.

Regarding the class of the population spectrum estimate used here, we should introduce a previous estimate (the *periodogram*) to distinguish between our procedure and that of Bevan *et al* (1979).

The periodogram, denoted by $\widehat{s}_z(w_j)$, is the simplest version of the population spectrum estimate. It is computed for any frequency w_j as

$$\widehat{s}_z(w_j) = \frac{1}{2\pi T} \left\{ \left[\sum_{t=1}^{T} y_t \cos[w_j (t-1)] \right]^2 + \left[\sum_{t=1}^{T} y_t \sin[w_j (t-1)] \right]^2 \right\}$$

where

- $w_j = \frac{2\pi j}{T},\ j = 1, 2, \cdots, \frac{T-1}{2}$ are the defined frequencies.
- $\{y_1, y_2, ..., y_T\}$ is the set of observations and
- T denotes the sample size.

It is well known (Hamilton, 1994) that the sample periodogram $\widehat{s}_z(w)$ is an asymptotically unbiased estimate of the population spectrum, $s(w; \Theta)$. However, the variance of the sample periodogram does not decrease as the

[7] In a form that will be defined in the next paragraph.

sample size increases. As a result, the periodogram is not a consistent estimate[8]. To obtain consistency, it is common to average the periodograms in a set of neighbourhood frequencies around w (Priestley, 1981). As shown in Bevan *et al* (1979), the periodogram is approximately distributed as a chi-square variable with 2 degrees of freedom; if periodograms of sucessive frequencies are averaged, the asymptotic distribution of estimates φ are approximately independent gamma variables (Bevan *et al*, 1979).

Mainly, our study differs from Bevan *et al* (1979) in the class of the population spectrum estimate employed. To compute the population spectrum, these authors average the periodograms estimated using the Fast Fourier Transform (FFT) algorithm over a set of different time series intervals.

Alternatively, we use a non-parametric estimate specified in Hamilton (1994), which represents a weighted average of frequencies around any fixed frequency w_j,

$$\widehat{s}_Z(w_j) = \sum_{m=-h}^{h} \left[\frac{h+1-|m|}{(h+1)^2} \right] \widehat{s}_z(w_{j+m}) \qquad (3.3.2.2)$$

The value h is a *bandwidth* parameter, which indicates the number of frequencies that are taken for estimating $\widehat{s}_Z$.

To illustrate the behavior of our non-parametric estimate, Figure 3.3 plots the different representations of the population spectrum estimate under different degrees of smoothness, just by changing the bandwidth parameter h. Figure 3.3 exhibits the population spectrum estimate for $h = 0, 50, 100$ and 200. The case $h = 0$ is equivalent to the spectrum obtained by the sample periodogram $\widehat{s}_z(w_i)$.

As Figure 3.3 reveals, the higher the parameter h, the smoother the spectrum. For example, some waves presented on the cases $h = 50, 100$ around frequencies 400 and 600 become regular when $h = 200$, as a direct consequence of the increase in the number of neighbourhood frequencies averaged. Finally, the graph for $h = 0$ displays the irregular shape for any frequency of the spectrum obtained by using the sample periodogram as an estimate of the population spectrum.

Coming back to the discussion of the population spectrum estimate chosen in this chapter, Priestley (1981, p. 582) points out the equivalence between the two alternatives of population spectrum estimates, that is, the averaged FFT of Bevan *et al* (1979) and our non-parametric estimate. Regarding our case, the election of the non-parametric estimate has been determined by two main reasons:

[8] A more detailed discussion about this point can be found in Priestey (1981), p. 432.

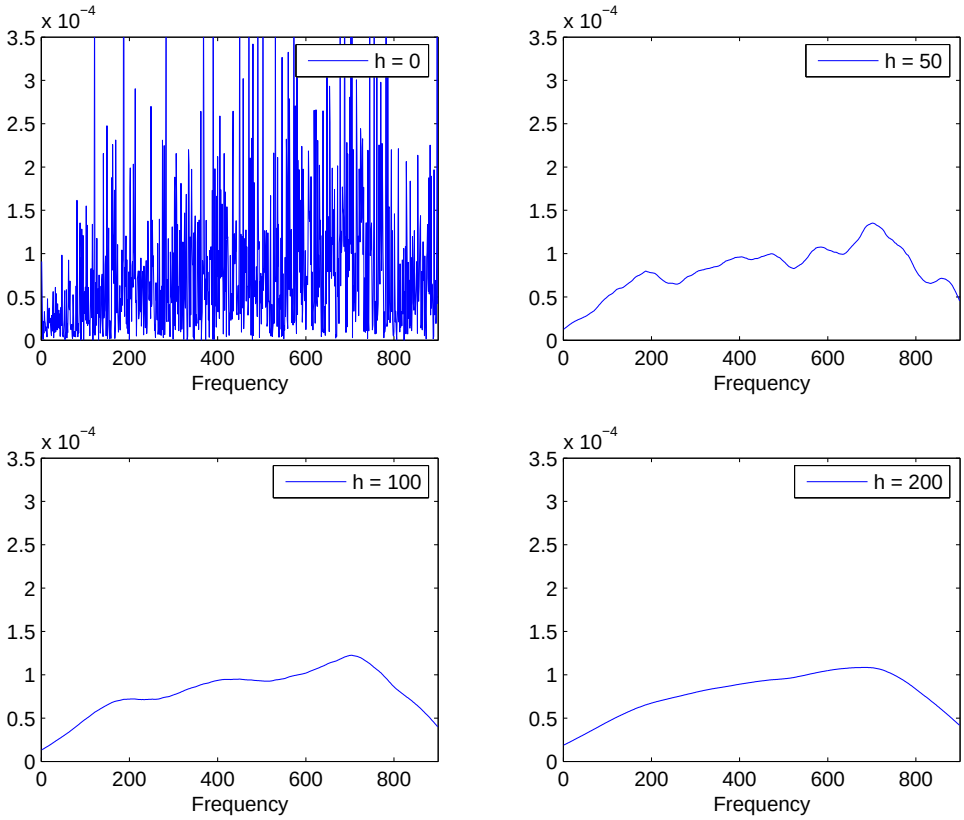


Figure 3.3: Sensitivity of the sample spectrum to the bandwidth parameter h. This figure displays different spectra obtained for $h = 0, 50, 100, 200$. Simulation (annualized) parameters are $\mu = 0.05$, $\sigma = 0.20$, $\lambda = 10.00$, $\beta = 0.10$ and $a = 0.60$

1. It is more intuitive to select the level of smoothness of the population spectrum estimate, just taking different sets of frequencies around w.

2. This new approach provides no substantial computing cost.

Taking into account the nature of our estimate due to the arbitrary selection of the bandwidth parameter, some objections could be made to the distribution of the estimates. In prevision to this, some care has been taken when computing the distribution of the estimates. Thus, to deal with possible problems related to the finite sample distribution, we have simulated numerically this distribution. This is a standard problem that involves the following steps:

1. We simulate N different paths of the posited model with a certain parameter vector Θ_0 -true values-

2. We obtain a vector of estimates $\widehat{\Theta}_i$ of the posited model for each i-path

3. Then, we build the histograms of the variables in the vector $\widehat{\Theta}_i$ by using the N estimated values obtained for each variable.

4. With the empirical distribution of estimates at hand it is possible to compute the desired confidence intervals.

A detailed analysis of the distribution of the estimates we have obtained will be provided in the Monte Carlo study section.

The Whittle estimate (1953)

The Whittle estimate[9] (1953) is based on minimizing the function

$$L_w(\Theta) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \left\{ \ln f(w;\Theta) + \frac{\widehat{s}_z(w)}{f(w;\Theta)} \right\} dw \qquad (3.3.2.3)$$

where w denotes the set of frequencies, $f(w;\Theta)$ is the (theoretical) population spectrum and $\widehat{s}_z(w)$ is the periodogram previously defined.

Generally speaking, the Whittle estimates $\hat{\Theta}$ are consistent, efficient and follow asymptotically a normal distribution. Moreover, it can be shown that minimizing L_w is an alternative to maximum likelihood, as it was pointed out by Robinson and Vidal (2005).

[9]This estimate was suggested by J. Vidal Sanz. We are very grateful to him.

Instead of using the estimate in (3.3.2.3), we will use a discrete estimation procedure inspired on Casas and Gao (2004) or Robinson and Vidal (2005). Basically, it consists on considering an approximation of the integral L_w of the form

$$L_w(\Theta) = \frac{1}{2\pi} \sum_{j=1}^{(T-1)/2} \left\{ \ln f(w_j;\Theta) + \frac{\widehat{s}_z(w_j)}{f(w_j;\Theta)} \right\} \tag{3.3.2.4}$$

with

$$w_j = \frac{2\pi j}{T},\ j = 1, 2, \cdots, \frac{T-1}{2}$$

where T denotes the sample size.

Again, some objections could be done due to the discrete nature of the estimate (3.3.2.4). To avoid this kind of problems, we will proceed similarly to the previous section, where we have simulated numerically the distribution of the estimates[10]. At the same time, this will be a form of checking the distributional behavior of the estimate on a finite sample experiment[11].

Taking into account all the previous considerations, our estimation problem results on the following optimization expression

$$\min_{\Theta} L_w(\Theta) = \min_{\Theta} \sum_{j=1}^{(T-1)/2} \left\{ \ln f(w_j;\Theta) + \frac{\widehat{s}_z(w_i)}{f(w_j;\Theta)} \right\} \tag{3.3.2.5}$$

To the best of our knowledge, the literature about estimation of continuous time processes using the Whittle estimate is scarce. As far as we know, just two papers have dealt with this issue: Gao (2004), which considers an estimation procedure based on the Whittle estimate for long-range dependent Gaussian processes, and Casas and Gao (2004), which could be considered as an application of the former article to real data.

3.2.4 The population spectrum of the SNJD model

We recall here some details of the SNJD model under study. As it was previously developed in Chapter 1, we are interested on estimate the parameters of the SNJD model with exponential decaying function of the form,

$$S_t = S_0 \exp\left[\left(\mu - \frac{\sigma^2}{2}\right)t + \sigma B_t\right] \prod_{j=1}^{N_t} \left[1 + U_j e^{-a(t-\tau_j)}\right] \tag{3.3.2.6}$$

[10] For the sake of brevity, we address the reader to the previous subsection for a detailed exposition of this point.

[11] To the best of our knowledge, this issue has not yet been reported in the literature.

where

- S_j denotes the asset price at time j.
- μ, σ are constants.
- $\{B_t\}_t$ is a standard Brownian motion.
- $\{N_t\}_{t\geq 0}$ is a Poisson process with intensity λ.
- $\{U_j\}_j$ is a sequence of i.i.d. jumps.
- τ_j is the instant in which the j-th jump appears.

The population spectrum of the SNJD model was also computed in Chapter 1 and takes the following form

$$s_Z(w) = \frac{1}{2\pi}\left(\sigma^2 + \lambda E[U^2] \ + \lambda(e^{-a\Delta t} - 1)E\left[U^2\right]\frac{2a}{a^2+w^2}\right)\Delta t \quad (3.3.2.7)$$

As it is deduced from last expression, the SNJD spectrum depends on the set of frequencies chosen. To get a visual idea of what equation (3.3.2.7) represents, Figure 3.4 displays the spectra (blue line) obtained for simulations[12] of the SNJD model with different values of the parameter a. The bandwidth parameter has been fixed to $h = 100$ for all cases.

As it is exhibited in Figure 3.4, the higher the parameter a, the bigger the spectrum on higher frequencies is. This fact seem to be confirmed by the slope of the curve, which increases with a.It is worth to mention the flat spectrum obtained for the case $a = 0.00$, which corresponds to the spectrum of a JD process.

Another interesting consequence of expression (3.3.2.7) is that the spectrum of the SNJD process for "large enough" frequencies ($w \to \infty$) converges to

$$s_Z(w) = \frac{1}{2\pi}\left(\sigma^2 + \lambda E[U^2]\right)\Delta t \quad (3.3.2.8)$$

that is, the spectrum of the JD model. This result suggests that the SNJD model behaves similarly to a JD model in small time spans[13]. However, the main contribution to the SNJD spectrum is done at lower frequencies, that is, by observations related to long-time horizons.

[12] All the simulations have been carried out with the same seed, just modifying the decaying parameter.

[13] This could explain the similar expressions we have obtained for the centered moments of the JD and SNJD process for small Δt.

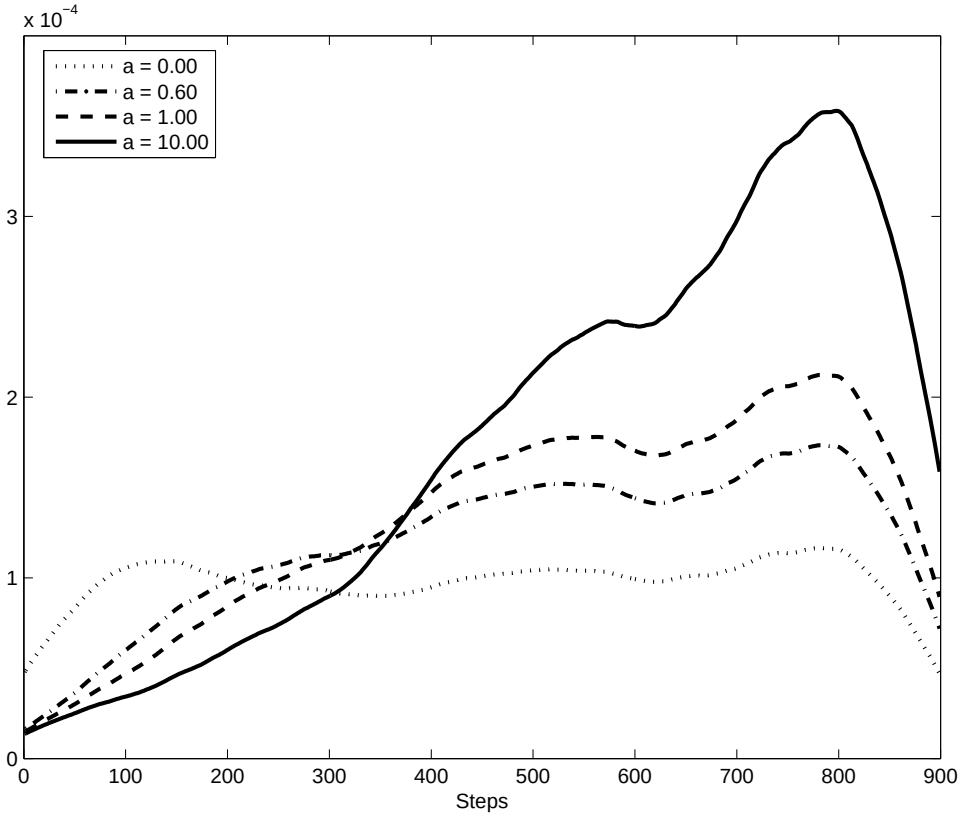


Figure 3.4: Sample spectra, for different values of the parameter a, of a SNJD model with exponential decaying function. We consider the case $h = 100$. The response function of the restricted SNJD model is $h(t - \tau_k) = e^{-a(t-\tau_k)}$ and the random variable U is lognormally distributed with mean $E[U] = 0$. Simulation parameters are $\mu = 0.0005$, $\sigma = 0.0126$, $\lambda = 0.0400$ and $\beta = 0.10$.

This result might be consistent with many empirical findings of the literature, as the presence of long-memory in asset returns cited by Lo (1991). Anyway, this affirmation must be taken carefully, in the sense that the problem of identifying long (or short) memory components in the financial series is not obvious, and it calls for a deep study, a matter that is beyond the scope of this dissertation.

Finally, the parameter vector for log-returns in expression (3.3.2.7) results $\Theta = (\sigma, \lambda, \beta, a)' \in R^4$, where

- σ is the volatility of the Brownian motion.
- λ is the intensity of the Poisson process, measuring the probability of appearance of an extreme event.
- β is related to the mean and variance of the jump size (its amplitude).
- The parameter a refers to the speed of the decaying effect after a jump event.

It is worth to note that the parameter μ does not appear in the parameter vector Θ. This is because the mean of the observations has been subtracted from the observations, a common practice in Spectral Analysis.

3.2.5 A note on the spectral estimation of diffusion processes

Since the spectral analysis has not been extensively employed in financial econometrics, it could be convenient to spend a few lines discussing the approach that will be adopted here. Mainly, we address the question about what could be the gain by using the Spectral Analysis approach in continuous-time finance.

Basically, the analysis in the frequency domain consists on a transformation of the autocovariance function of the stochastic process that provides us an alternative way of capturing the dynamics of the process. To a certain degree, the richer structure of autocovariance, the more useful this technique.

As it was pointed out by Lo (1991) in a context of long-term memory, we are inclined to believe that most of the current models employed in finance are not consistent with the serial dependence evidences suggested in

some financial series[14], which seems to support the idea of getting rid of the frequency domain tools.

To put it another way, a wide range of current continuous-time models does not impose enough structure on cross moments to capture the empirical features. This situation does not incite to adopt the spectral analysis approach, which reveals potentially usefulness when dealing with the dynamics of the process.

On the other hand, the spectral analysis does not overcome the difficulties on solving some standard problems in continuous-time finance as the estimation of JD processes[15]. Traditionally, the maximum likelihood estimation of JD processes has revealed inconsistent (Honoré, 1998) due to problems with the boundaries of the likelihood function[16]. This problem is not solved by the spectral analysis, as equation (3.3.2.8) reveals: assuming that the terms σ, λ and β are constants, the spectrum of a JD process is also constant *independently* of the frequency chosen. Indeed, we are facing an indetermination problem since an infinite set of parameter values lead to a solution of expression (3.3.2.8).

To the best of our knowledge, just some recent papers have pointed out the estimation of continuous-time processes with spectral tools with the ability of generating serial dependence. For example, this is the case of the fractional Brownian motion (fBm)[17], a generalization of the standard Brownian motion that is able to introduce autocovariance in the increments of the process. Gao (2004) and Casas and Gao (2004) consider two versions

[14]See, for instance, Benito *et al* (2007) for evidences of autocorrelation in daily changes in EONIA rates; Campbell *et al* (1997) find statistically significant correlation coefficients for the few first lags of Value and Equal Weighted stock Indexes of the Center for Research of Security Prices (CRSP), respectively, at different sampling frequencies. Finally, Singleton (2006) provides an interesting discussion about the testing of serial correlation in asset returns

[15]This point was generously noticed by E. Sentana. We are grateful to him for this comment.

[16]Currently, the estimation of JD processes could be tackled by different techniques, as GMM (Das and Sundaram, 1999), spectral GMM (Chacko and Viceira, 2003) or Efficient Moment Methods (Andersen *et al*, 2002). A more extensive discussion about this point can be found in Chapter 2.

[17]The fractional Brownian motion $B_H(t)$, with $t \in [0,T]$ is a Gaussian, continuous-time process with zero mean and covariance function equal to

$$E[B_H(t), B_H(s)] = \frac{1}{2}(|t|^{2H} + |s|^{2H} - |t-s|^{2H})$$

where the real number $H \in [0,1]$ is called the Hurst index. For $H = 1/2$, the process is an standard Brownian motion. The cases $H > 1/2$ and $H < 1/2$ generate positive or negative serial dependence, respectively.

(discrete and continuous, respectively) of a spectral estimate, the Whittle estimate, for capturing the parameters of a fBm[18].

On the other hand, the continuous-time autorregressive moving average (CARMA) model seems to be also an alternative to generate serial persistence, but its treatment is beyond the scope of this dissertation.

3.3 Some details on the estimation procedure

Similarly to Chapter 2, we detail here some aspects about the estimation procedure employed. Basically, we refer some characteristics of the model employed and the optimization routine.

All the experiments are developed for the SNJD model (see (3.3.2.6)) with lognormally distributed jump sizes $U = e^{-\beta^2/2+\beta\epsilon} - 1, \varepsilon \sim N(0,1)$. Some remarks also mentioned in previous Chapters 1 and 2: the assumption of a lognormal distribution for the jump sizes is common in the jump-diffusion literature (see, for instance, Merton (1976)). To decrease the number of parameters under estimation, we impose $E[U] = 0$, another common assumption. As usual through this dissertation, the response function considered is $h(t) = e^{-at}$, which nests the Merton (1976) model[19]. To the best of our knowledge, the functional form of $h(t)$ has not been reported yet in the financial literature. Then, for analytical and intuition purposes, we assume an exponential decreasing form for this function.

For simplicity and without loss of generality, the estimation of the parameter μ in the model (3.3.2.6) has been omitted. As a standard procedure in spectral analysis, the parameter vector results $\Lambda = (\sigma, \lambda, \beta, a)' \in R^4$, where σ is the volatility of the Brownian motion, λ is the intensity of the Poisson process, β is related to the mean and variance of the jump size, and the parameter a refers to the speed of the decaying effect after a jump event.

3.4 Monte Carlo study for the naive estimate

This section analyzes the capacity of the naive estimate proposed previously to estimate the parameters of the SNJD process. To check the estimation procedure, we implement some Monte Carlo experiments: firstly, we study the estimation methodology considering different values for the smoothness parameter of the nonparametric spectrum estimate, with the aim of selecting

[18] For a more detailed treatment of fBm we address the reader to Gao and Casas (2004) and references therein.

[19] The parameter a is zero in Merton (1976).

a suitable value of the bandwidth parameter h. Once fixed the h value for the subsequent experiments, we reproduce the experiment for the classical Shot Noise and restricted SNJD process, similarly to that carried out on Chapter 2. Finally, we extend the Monte Carlo study for the naive estimate to different values of the SNJD process itself.

This section is partially based on Bevan *et al* (1979). We refer the reader to section 2.3.1. for a detailed discussion about the estimate employed here.

3.4.1 Sensitivity to the bandwidth parameter h

Due to the non-parametric nature of our estimate, it is necessary to check the response of the non-parametric spectrum estimate (3.3.2.2) to changes in the bandwidth parameter h. This subsection carries out this analysis.

Previous work[20] indicate that the SNJD model with exponential response function presents statistically significant autocorrelation coefficients on lags 1 and 2 for the values $a = 0.60, 1.00$. To make things easier, we have simulated a sample of SNJD processes with $a = 0.60$ and estimated their parameters for a set of different values of the coefficient h. The results obtained are shown in Table 3.1.

Table 3.1 exhibits the mean, median, minimum, maximum, mean squared error (MSE) and root mean squared error (RMSE) obtained for a SNJD sample of 3,000 paths with 1,800 data each one. To control as much as possible the effects of the bandwidth parameter h in the estimation, we have used the same sample in all experiments carried out. Finally, the set of values of h we have chosen are $h = 0, 25, 50, 100$ and 200. The case for $h = 0$ corresponds to the sample periodogram estimate.

Some interesting results arise from Table 3.1. For example, the best estimates in terms of RMSE errors are obtained for the highest values of h. Moreover, the estimates for σ and λ parameters seem to present a considerably bias with respect to their true simulation values, being worth to mention that the magnitude of this bias is slightly reduced when we increase the bandwidth parameter h. Finally, just mention that the periodogram estimate for $h = 0$ could provide negative estimates of the Brownian motion volatility σ, as it is reflected in the minimum value for the parameter σ in this case.

From the inspection of Table 3.1, it seems that the smoother the spectrum, the better results are obtained on the whole. However, an important bias is detected on the parameters σ and λ. Regarding the former, the better

[20] These results are available upon request.

Table 3.1: Changes in the estimates with respect to bandwidth parameter h.

	h	$\widehat{\sigma} \times 10^{-2}$ (1.26)	$\widehat{\lambda} \times 10^{-2}$ (4.00)	$\widehat{\beta} \times 10^{-2}$ (10.00)	$\widehat{a}$ (0.60)
Mean	0	1.36	11.24	11.80	1.29
	25	1.98	10.67	9.84	0.62
	50	1.97	10.76	9.73	0.59
	100	1.95	10.56	9.65	0.56
	200	1.95	10.11	9.43	0.54
Median	0	1.92	9.86	11.12	0.75
	25	1.97	10.36	9.53	0.59
	50	1.96	10.35	9.41	0.57
	100	1.94	10.13	9.33	0.54
	200	1.94	9.74	9.17	0.52
Min	0	−2.23	1.58	5.01	0.01
	25	1.37	2.42	4.97	0.30
	50	1.47	2.24	4.62	0.31
	100	1.39	2.01	2.12	0.29
	200	1.46	2.15	1.71	0.30
Max	0	2.46	26.31	23.09	3.43
	25	2.41	20.00	18.90	2.00
	50	2.41	20.00	19.84	1.96
	100	2.80	20.00	18.82	2.00
	200	2.48	20.00	16.98	2.00
MSE	0	0.01	0.76	0.18	1.32
	25	0.01	0.59	0.04	0.02
	50	0.01	0.61	0.04	0.02
	100	0.00	0.58	0.04	0.01
	200	0.00	0.50	0.04	0.02
RMSE	0	1.06	8.73	4.19	1.15
	25	0.72	7.71	2.08	0.15
	50	0.71	7.82	2.11	0.14
	100	0.70	7.63	2.11	0.12
	200	0.70	7.07	2.03	0.14

Estimation results for the non-parametric estimates (3.3.2.2) with respect to changes in the bandwidth parameter h. The first row contains the parameters of the SNJD model as given by expression (3.3.2.6). True simulation values are in parenthesis. This table contains six blocks corresponding to the following statistics for these estimates: mean, median, minimum, maximum, mean squared error (MSE) and root mean squared error (RMSE).Each block comprises five rows, attending to the different values of the parameter h under study.

results could be easily understood taking into account that by smoothing the spectrum, we are reducing the variability of the sample spectrum, that is, we are using a sample estimate with "less noise", as Figure 3.3 reflects.

Regarding the bias, Hamilton (1994) mentions that, when the periodograms are averaged at different frequencies, the estimates obtained reduce the variance of the periodogram but they tend to introduce some bias, depending on the size of the bandwidth. This author suggests a complementary experiment to that realized here, mainly based on a subjective election of the bandwidth parameter[21].

Finally, we do not find substantial differences between using $h = 100$ or 200 instead. Then, we will select (arbitrarily) a bandwidth parameter $h = 100$ to carry out our estimations.

3.4.2 Previous study

Some details

Directly inspired on Chapter 2, we start by analyzing the performance of the naive estimate under two different experiments:

1. First, we study the estimation results for a simulated sample of the classical shot noise process referred on section 4.1 of Chapter 2.

2. Then, we repeat the analysis for a restricted version of the SNJD model, denoted as restricted SNJD model, where the μ and σ parameters have been fixed to zero.

Before showing the results of the estimations for the classical shot and restricted shot noise processes, we provide here the expressions of their population spectra. Again, we address the reader to section 4.1 of Chapter 2 for a more detailed discussion about the reasons why both processes have been selected for our study.

On one hand, the classical shot noise is defined as

$$H_t^c = \sum_{i=1}^{N_t} U_i e^{-a(t-\tau)} \tag{3.3.4.1}$$

[21] Implicitly, this idea has been already developed in Figure 3.3, where we have represented the spectra under different bandwidth (smoothing) parameters.

where $\{N_t\}_t$ and $\{U_j\}_j$ are mutually independent. The population spectrum for this process[22] takes the form

$$s_Z^c(w) = \frac{\lambda E[U^2]}{a^2+w^2} \tag{3.3.4.2}$$

where w denotes the set of frequencies under study.Note from equation (3.3.4.2) that the spectrum of a classical shot noise exhibits a concentration of power in lower frequencies and tends to zero as w parameter increases[23]. A simulation of the typical shape of the spectrum for a classical shot noise model (blue line) is provided in Figure 3.5.

Regarding the restricted version of the SNJD model, we recall that its expression is

$$H_t^r = \sum_{j=N_t+1}^{N_{t+\Delta t}} \ln[1+U_j e^{-a(t+\Delta t-\tau_j)}] + \sum_{j=1}^{N_t} \ln\left[\frac{1+U_j e^{-a(t+\Delta t-\tau_j)}}{1+U_j e^{-a(t-\tau_j)}}\right] \tag{3.3.4.3}$$

The spectrum for this process is easily obtained by substituting $\mu = \sigma = 0$ in equation (3.3.2.7), which leads to the following expression

$$s_Z^r(w) = \frac{1}{2\pi}\left(1+(e^{-a\Delta t}-1)\frac{2a}{a^2+w^2}\right)\lambda E[U^2]\Delta t \tag{3.3.4.4}$$

with w is the usual frequencies set.

Some remarks should be done about former equation (3.3.4.4):

1. Note this expression leads to negative values of the spectrum for a certain set of parameters a and w, a problem that could be easily avoided by a careful selection of boundary values and frequencies.

2. Contrary to the spectrum of the classical shot noise, this expression allocates the main part of power in *higher* frequencies, instead of lower ones.

Figure 3.5 also includes the graphical representation of the spectrum (red line) obtained with a restricted version of the SNJD model.

[22]This population spectrum has been obtained by computing the Fourier transform over the autocovariance function of the classical shot noise. For a detailed derivation of the autocovariance of the classical shot noise, see Ross (1996).

[23]Such a concentration of power in lower frequencies is usually put down to the presence of long-memory components in the series (Lo, 1991).

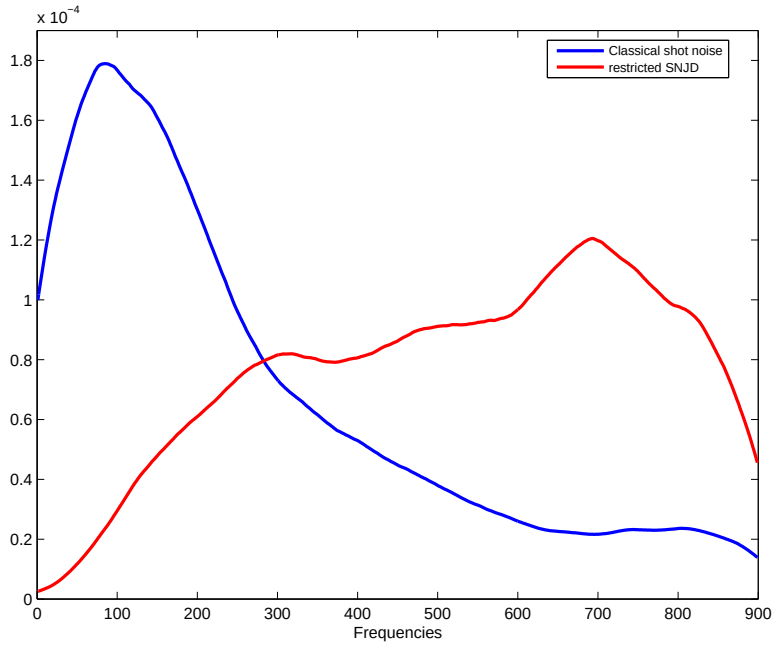


Figure 3.5: Nonparametric estimate of the population spectrum for classical Shot noise (blue line) and restricted SNJD processes (red line). The response function is $h(t-\tau_k) = e^{-a(t-\tau_k)}$ in both cases. The random variable U_k is lognormally distributed as $U = e^{-\beta^2/2+\beta\varepsilon} - 1$, with $\varepsilon \sim N(0,1)$ and mean $E[U_k] = 0$. Simulation parameters are $\lambda = 0.04$, $\beta = 0.10$ and $a = 0.50$. The bandwidth parameter h of the estimate has been fixed to 100.

Results

Table 3.2 exhibits the results obtained for the two main experiments carried out: the estimation of classical and restricted shot noise, respectively. As it is usual in this dissertation, the results correspond to sets of 3,000 paths composed by 1,800 simulations each one. This table is structured in two different blocks, corresponding to the first and second experiments, respectively.

The first column in Table 3.2 displays the mean, median, minimum, maximum, mean squared error (MSE) and root mean squared error (RMSE). The three next column blocks comprise the results for the three parameters of each model (λ, β and a). At the same time, each column blocks is divided into four columns, corresponding to each of the different degrees of serial dependence ($a = 0.20, 0.60, 1.00$ and 10.00). Finally, the true simulation

values are given in parenthesis.

Concerning to the first experiment, parameters λ and β exhibit a higher bias as parameter a increases, which results in a loss of efficiency of the estimates. With respect to parameter a, their estimates are slightly biased to the right; anyway, in terms of efficiency and bias, these results seem much better than those obtained for the remaining parameters.

With regard to the restricted SNJD model, Table 3.2 seems to present slightly different conclusions compared to those for the classical Shot noise model: it seems to be clear the shifted (to the right) behavior of the parameters λ and β, which generally present smaller bias than the classical Shot noise ones. By contrast, the observed bias for the estimates of the parameter a is much bigger (cases of $a = 0.20, 0.60$ and 10.00) than that for the classical Shot noise ones.

3.4.3 Main results

We present now the central result of this section: the estimation of the SNJD model with exponential response function by using the naive estimate of expression (3.3.2.2). The estimates we have obtained are displayed in Table 3.3.

Table 3.3 is structured in four vertical blocks. The first and second columns display the parameters under study and their true values. The third vertical block is divided into seven columns, containing the mean, median, standard deviation (Std), minimum (Min), maximum (Max), mean squared error (MSE) and root mean squared error (RMSE) for the spectral estimates of the SNJD model. Finally, the last block includes some percentiles of the estimates distribution.

To keep tractable our study, we detail separately the main features of the obtained results:

- **Parameter σ:** A bias for this parameter is observed in all the cases. Letting the $a = 10.00$ case aside, the observed bias in this parameter decreases slightly with the parameter a. Additionally, the most efficient estimates are obtained for $a = 0.60$ and $a = 1.00$, as it is deduced from their RMSE and the percentiles values. Finally, negative estimates for σ have been encountered in no more than 10% of the obtained results.

- **Parameter λ:** Results in Table 3.3 seem to exhibit a detectable pattern for λ parameter: the higher the parameter a, the less efficient and more biased the λ value. This behavior is confirmed by a general

Table 3.2: Naive estimates for classic Shot noise and SNJD process

Parameters (True value)	$\lambda \times 10^{-2}$ (4.00)				$\beta \times 10^{-2}$ (10.00)				a (below)			
a	0.20	0.60	1.00	10.00	0.20	0.60	1.00	10.00	0.20	0.60	1.00	10.00
I.- Classic Shot noise model												
Mean	11.36	14.61	15.92	24.20	13.33	12.80	15.36	49.25	0.38	0.86	1.42	8.66
Median	11.40	14.25	15.24	21.96	12.55	12.37	15.04	48.52	0.38	0.86	1.43	8.35
Min	1.98	1.48	1.57	2.06	7.28	5.21	4.57	1.07	0.31	0.58	0.68	2.24
Max	20.00	30.00	32.00	62.00	27.90	28.04	30.00	99.76	0.49	1.19	1.92	15.00
MSE	0.73	1.40	1.78	5.07	0.23	0.15	0.41	18.17	0.03	0.07	0.20	11.68
RMSE	8.56	11.83	13.33	22.51	4.82	3.91	6.37	42.63	0.18	0.27	0.44	3.42
II.- Restricted SNJD model												
Mean	8.81	10.19	10.37	13.80	10.61	11.28	14.23	17.27	0.10	0.29	1.21	2.14
Median	8.99	10.32	10.48	11.96	10.16	10.96	14.24	16.76	0.09	0.29	1.15	1.52
Min	2.36	2.94	3.92	2.42	6.20	6.54	8.53	8.84	0.06	0.18	0.33	1.37
Max	15.02	16.77	16.19	30.00	15.00	15.00	20.64	30.00	3.50	2.12	2.20	14.49
MSE	0.31	0.46	0.44	1.39	0.04	0.05	0.22	0.72	0.02	0.10	0.43	68.39
RMSE	5.59	6.81	6.66	11.81	2.06	2.15	4.64	8.50	0.15	0.32	0.66	8.27

Naive estimates for classic Shot noise and SNJD processes. This table shows the mean, median, minimum, maximum, mean square error (MSE) and root mean square error (RMSE) of the estimated parameters, respectively. Simulation parameters are in parenthesis. The response function is $h(t-\tau_k) = e^{-a(t-\tau_k)}$ in both cases. The random variable U_j is lognormally distributed $U = e^{-\beta^2/2+\beta\varepsilon} - 1$, with $\varepsilon \sim N(0,1)$ and mean $E[U] = 0$.

Table 3.3: Results for the naive spectral estimate with SNJD samples

Parameter	True value	Statistics							Percentiles						
		Mean	Median	Std	Min	Max	MSE	RMSE	2.5%	5%	10%	50%	90%	95%	97.5%
$\sigma \times 10^{-2}$	1.26	1.98	1.98	0.23	0.29	2.50	0.01	0.75	1.57	1.68	1.77	1.98	2.22	2.33	2.46
$\lambda \times 10^{-2}$	4.00	9.17	8.60	3.96	1.40	25.53	0.42	6.51	3.08	3.63	4.50	8.60	14.73	16.58	17.98
$\beta \times 10^{-2}$	10.00	8.09	7.79	2.10	2.43	17.53	0.08	2.84	4.81	5.24	5.66	7.79	10.92	12.01	13.02
a	0.20	0.80	0.28	0.92	1.20	3.00	1.21	1.10	0.13	0.15	0.17	0.28	2.49	2.72	2.94
$\sigma \times 10^{-2}$	1.26	1.95	1.95	0.13	1.26	2.40	0.00	0.69	1.69	1.73	1.78	1.95	2.11	2.18	2.21
$\lambda \times 10^{-2}$	4.00	10.54	10.22	3.86	2.44	22.28	0.58	7.59	4.22	5.00	5.84	10.22	15.99	17.57	18.76
$\beta \times 10^{-2}$	10.00	9.68	9.38	2.00	5.17	17.76	0.04	2.03	6.52	6.84	7.32	9.38	12.44	13.33	14.18
a	0.60	0.57	0.54	0.18	0.31	2.00	0.03	0.18	0.38	0.41	0.43	0.54	0.71	0.81	0.93
$\sigma \times 10^{-2}$	1.26	1.91	1.91	0.15	0.34	2.43	0.00	0.66	1.66	1.71	1.75	1.91	2.07	2.15	2.17
$\lambda \times 10^{-2}$	4.00	11.64	11.16	3.86	3.38	28.15	0.73	8.53	5.45	6.15	6.95	11.16	17.24	18.53	19.39
$\beta \times 10^{-2}$	10.00	11.22	11.03	2.17	5.29	19.02	0.06	2.49	7.61	8.05	8.60	11.03	14.21	15.21	15.80
a	1.00	0.86	0.83	0.18	0.41	2.46	0.05	0.23	0.62	0.65	0.69	0.83	1.04	1.13	1.18
$\sigma \times 10^{-2}$	1.26	0.13	0.01	0.47	-1.77	2.41	0.02	1.23	-0.61	-0.44	-0.29	0.01	0.89	1.15	1.29
$\lambda \times 10^{-2}$	4.00	15.99	15.55	4.34	5.31	29.98	1.63	12.75	9.04	9.91	10.81	15.55	22.21	23.51	24.70
$\beta \times 10^{-2}$	10.00	16.27	16.14	2.50	5.19	24.55	0.46	6.75	11.84	12.50	13.15	16.14	19.67	20.43	21.12
a	10.00	1.75	1.70	0.72	1.20	11.61	68.55	8.28	1.46	1.54	1.60	1.70	1.78	1.81	1.84

Results for the naive spectral estimates with SNJD samples for some values of the parameter a. The first two columns display the parameters under study and their true values. Third to ninth columns exhibit the mean, median, standard deviation, minimum, maximum, mean squared error and root mean squared errors, respectively. The tenth to the last columns provide different percentile values of the estimate distribution. We have assumed a jump size U lognormally distributed of the form $U = e^{-\beta^2/2+\beta\varepsilon} - 1$, with $\varepsilon \sim N(0,1)$, and response function of the form $h(t) = e^{-at}$.

displacement of the estimate distributions (percentiles increase with the parameter a) to the right. Moreover, the estimate distributions appear not to be symmetrical and they exhibit a fatter right tail, as deduced from the percentile values for each value of a.

- **Parameter β:** At light of results displayed, the β estimates show the most stable behavior among the different set of parameter under study. For the sample analyzed, the higher the parameter a, the more biased β. By contrast, the RMSE values do not exhibit the same pattern as the bias, since the better results (in terms of RMSE) are obtained for $a = 0.6$ and $a = 1.00$.

- **Parameter a:** Maybe, the central parameter in this study. As Table 3.3 exhibits, the more efficient and less biased results are obtained for $a = 0.60$ and $a = 1.00$. The methodology seems to capture the case of $a = 0.2$, but a high percentage of the results (at least the 10%) are clustered on values bigger than 2.00, as the percentile values indicate.

  Moreover, the clear differences between the mean and median values obtained for $a = 0.20$ and the high value of the standard deviation (the biggest one across the a estimates) suggest us that, when $a = 0.20$, the estimate distribution has not a regular shape, probably asymmetric or even with two modes. Finally, concerning to the $a = 10.00$ case, the estimation values seem to be bounded around the point 2.00, as the minimum, median and percentile values confirm.

With the intention of presenting additional insights on former results, Figures 3.6 and 3.7 represent graphically, for each parameter, the true simulation value and the estimates obtained under the different a cases.

Figure 3.6 represents the estimates for the parameters σ and λ (upper and lower graphs, respectively) obtained for the different degrees of serial correlation we are considering ($a = 0.20, 0.60, 1.00$ and 10.00). The true simulation values are represented by a black horizontal line. On the whole, the closer the point clouds to the black line, the better the performance of the estimations.

Concerning to the parameter σ (upper graph), the existing bias is clearly observed in this figure. Moreover, we can also appreciate how the best σ estimates are obtained when $a = 0.20$ and $a = 0.60$. For $a = 10$, we observe negative estimates of the σ value and a higher dispersion in these estimates.

With respect to the λ value (lower graph in Figure 3.6), all the points are concentrated on the upper area of the graph, indicating a general bias

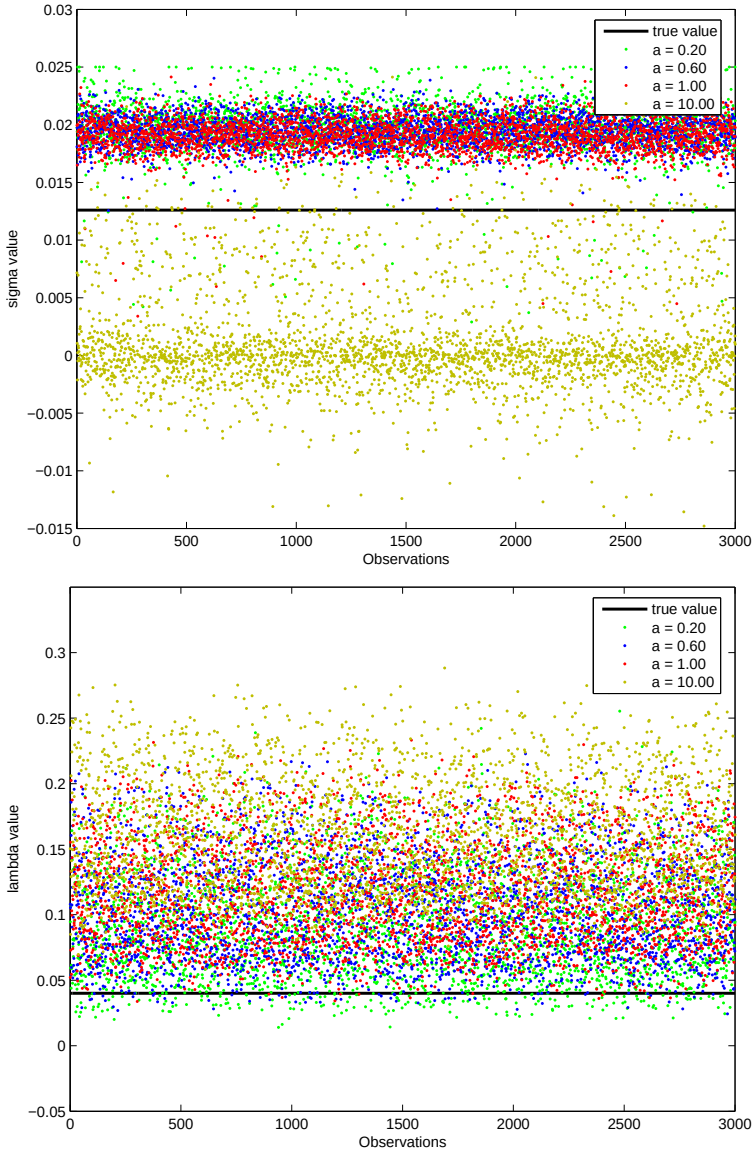


Figure 3.6: Estimated σ and λ parameters using spectral analysis for the different values of a. Upper graph results correspond to $\sigma = 1.26 \times 10^{-2}$ and lower graph to $\lambda = 4.00 \times 10^{-2}$, respectively. The true simulation value is represented by a black horizontal line. Different point colours correspond to the estimations obtained for the different cases of the parameter a under study.

to the right. Moreover, it seems to be observed that higher the parameter a, the more biased results. Contrary to the graph for the parameter σ, it is hard to distinguish the dispersion of the estimates for the different a cases.

In a similar way, Figure 3.7 represents the point clouds of the estimates for the parameters β (upper graph) and a (lower) parameters. As it was previously noted, the estimates of β seem to exhibit the most stable behavior across the different parameters under study: lower dispersion and clustered disposition around the true value (black line).

Regarding the lower graph in Figure 3.6, the different black horizontal lines correspond to the simulation values of the parameter a done. Moreover, the horizontal line corresponding to the $a = 10.00$ case has been dropped from this figure as it is far away from their estimates, confusing considerably the results obtained.

As Figure 3.7 reveals, the same conclusions of former figures apply here: the better estimates (less biased, less dispersed) correspond to the $a = 0.60, 1.00$ values. Although the case of $a = 0.20$ is almost placed around its true value (green points), we detect a higher dispersion on its estimated values up to value 3.00, corresponding with the optimization upper bound for this case. Finally, the methodology does not capture correctly the $a = 10.00$ case as we observe that points are located around the 1.60 value[24].

Finally, we also provide a representation of the histograms for the different estimates obtained for the SNJD process. For the sake of brevity, Figure 3.8 just displays the histogram for $a = 1.00$ parameter[25]. This figure also shows the normal distribution fit (red line) to the obtained results.

As Figure 3.8 exhibits, the estimates λ, β and a do not seem to be normally distributed: we observe that the different distributions show a certain asymmetry to the left, with right tails fatter than normal distribution ones. By contrast, these observations are not fully appreciated in the σ parameter case.

3.5 Monte Carlo study for the Whittle estimate

We present now the results concerning to the optimization of the Whittle estimate in expression (3.3.2.4). We start by detailing some of the problems encountered during the estimation procedure; then, we comment the main results obtained for the SNJD model.

[24]The upper bound of the optimization routine for this parameter has been placed in 12.00.

[25]The histograms for the remaining a cases are available upon request.

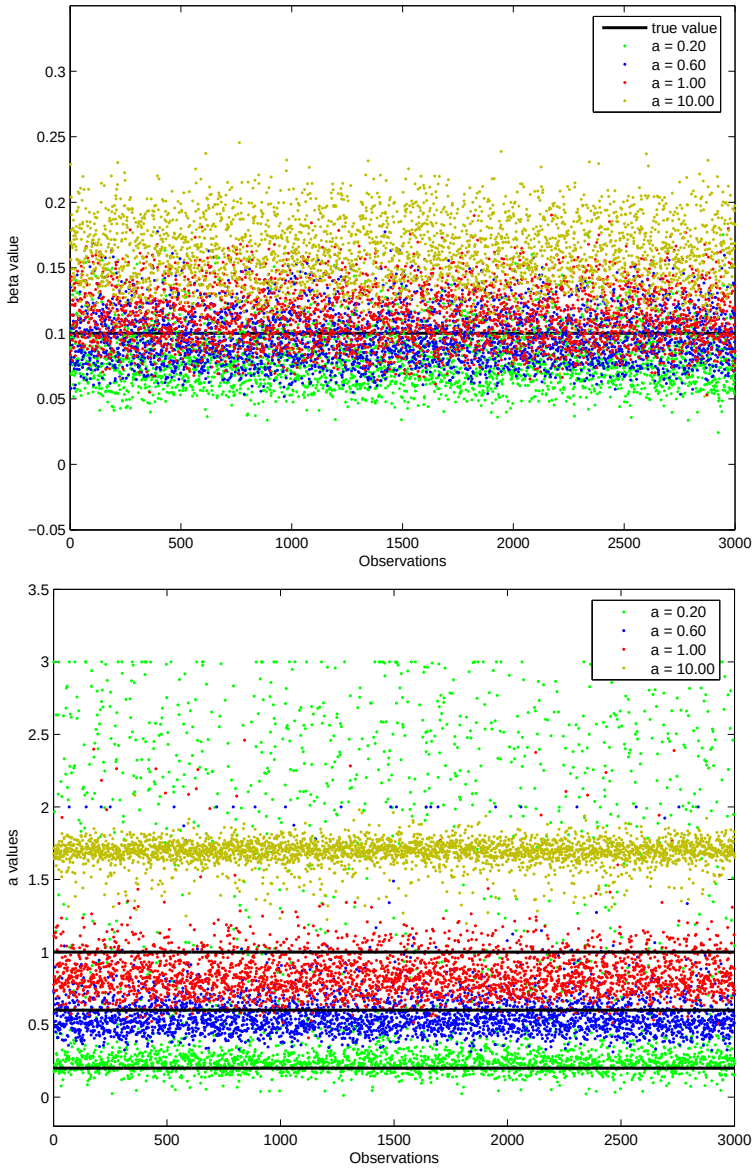


Figure 3.7: Estimated β and a parameters using spectral analysis for the different values of a. Upper graph results correspond to $\beta = 10.00 \times 10^{-2}$ and lower graph to different values of the parameter a, respectively. The true simulation values are represented by a black horizontal line. Different point colours correspond to the estimations obtained for the different cases of the parameter a under study. Finally, note that the black line corresponding to $a = 10.00$ case does not appear because it is outside the figure limits.

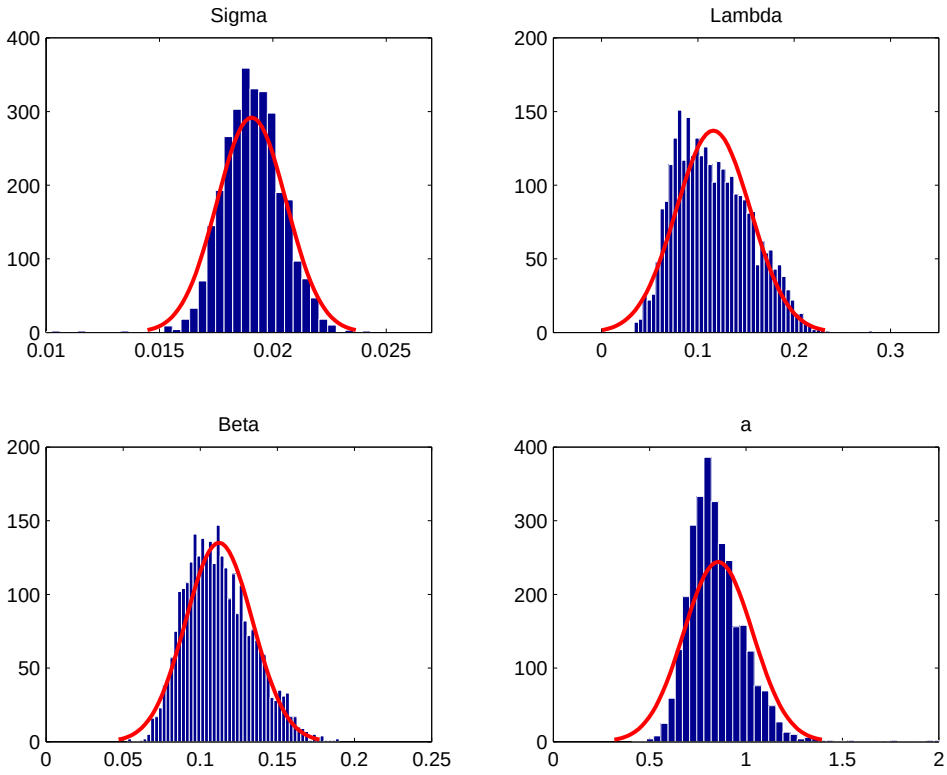


Figure 3.8: Sample distribution for spectral naive estimates. The red line corresponds to the normal distribution fit of the data. The sample is composed by estimations over 3,000 paths with 1,800 data each one, corresponding to daily frequency. Simulations parameters are $\mu = 0.0002$, $\sigma = 0.0126$, $\lambda = 0.0400$ and $\beta = 0.1000$. The parameter a has been fixed to 1.00. The jump size U is log-normally distributed of the form $U = e^{-\beta^2/2+\beta\varepsilon} - 1$, with $\varepsilon \sim N(0,1)$ and mean $E[U] = 0$.

3.5.1 Previous considerations

Contrary to the general procedure in this dissertation, this section just provides some results about the SNJD process itself, without passing through the previous studies offered for the different estimates proposed. The main reason for this is because of some numerical problems encountered during the computation of the estimate (3.3.2.4).

Until now, every estimate used here comprises the minimization of a quadratic form. For instance, the method of moments with identity weighting matrix of Chapter 2 does not differ much from the naive estimate posited in this chapter: both methods minimize the sum of quadratic errors between certain sample and population estimates[26]. All these computations were carried out by means of an *ad hoc* algorithm suitable for this kind of procedures.

By contrast, the optimization problem of the Whittle estimate (1953) in expression (3.3.2.4) differs substantially from that employed in the quadratic form environments: instead of using a non-least squared procedure, we are forced to use more suitable routines as the Simplex method or gradient techniques as the Newton-Raphson algorithm one, among others[27].

Focusing on the estimate employed here, we have chosen the Newton-Raphson routine. Basically, this election has been done attending to the computational time criteria[28].

Mainly forced by the expensive computing time, we have restricted our experiments to 1,000 paths with 1,800 data each one. Although this amount of sample is not enough to test the asymptotic properties of our estimate, we think it can us some idea of its behavior. Again, although our results might not be conclusive, it could give us some intuitions about the performance of the Whittle (1953) estimate.

---

[26] In the same line, the generalized moment method (GMM) of Hansen (1982) could be intended as the optimization of a quadratic form whose terms have been conveniently weighted

[27] Many standard textbooks of econometrics include some chapters devoted to these topics. See, for instance, Hamilton (1994) or Green (1998) and references therein.

[28] Taking into account that each optimization spent about 30 seconds for a tolerance level on the parameters variation of 10^{-5}, 3000 optimizations lead to about 25 hours (!). We are not considering here simulation procedure, which is also time-consuming. With respect the Simplex method, the optimizations work even worse than the Newton-Raphson procedure, unfortunately.

3.5.2 Main results

Table 3.4 summarizes the results obtained for the SNJD model using the Whittle (1953) estimate. The structure of this table is as follows: the first two columns display the parameters and their true (simulation) values. Then, we have a block of columns named 'Statistics', divided in seven columns, containing the mean, median, standard deviation (std), minimum, maximum, mean squared error (MSE) and root mean squared error (RMSE) of the estimates distribution. Finally, the block 'Percentiles' provides different percentiles of the estimates distribution. The sample size is composed by 1,000 paths composed by 1,800 data each one.

For the sake of exposition, we focus separately on the different parameters of the SNJD model:

- **Parameter σ:** The estimate distributions are clearly asymmetrical for $a = 0.20$ and 10.00, at light of their respective lower percentile values. This circumstance leads us to doubt about the performance of the estimate for this pair of values. By contrast, the distribution of the parameter σ seems to be symmetrical ($a = 1.00$) or almost symmetrical ($a = 0.60$). Again, σ parameters for the former a cases are biased and present equal RMSE values.

- **Parameter λ:** Once again, it seems to be clear the behavior for this parameter: the higher the parameter a, the higher the λ parameter . With respect to the bias behavior of the estimates, it seems to be clear that for a 95%-confidence level, the estimate is completely biased in almost all the a cases.

- **Parameter β:** For the lower a cases ($a = 0.20, 0.60$ and 1.00) we observe that the true value falls within the 95%-confidence interval. The closeness about the mean and median values, jointly with the almost symmetrical behavior of the percentiles, seem to indicate that the estimate distribution are symmetrical or almost symmetrical. Again, an observed increasing bias is reported when the parameter a increases.

- **Parameter a:** On the whole, the standard deviation of the estimates are much lower for the a parameters than in the previous cases. It is detected a different behavior for lower values of the parameter a $(0.20, 0.60)$ with respect to the higher ones $(1.00, 10.00)$. Thus, we can see that the former cases have estimate distributions that seem to be clustered on the same point, at light of their percentile values. Just

Table 3.4: Results for the Whittle estimate with SNJD samples.

Parameter	True value	Statistics							Percentiles						
		Mean	Median	Std	Min	Max	MSE	RMSE	2.5%	5%	10%	50%	90%	95%	97.5%
$\sigma \times 10^{-2}$	1.26	1.24	1.15	0.30	0.95	2.60	0.00	0.30	0.95	0.95	0.95	1.15	1.71	1.80	1.94
$\lambda \times 10^{-2}$	4.00	7.52	7.62	1.11	1.00	11.00	0.14	3.69	3.55	5.44	7.05	7.62	8.37	8.77	8.94
$\beta \times 10^{-2}$	10.00	11.72	11.72	1.09	5.30	14.21	0.04	2.04	9.43	10.00	10.49	11.72	12.97	13.03	13.16
a	0.20	0.40	0.41	0.04	0.09	0.45	0.04	0.20	0.25	0.34	0.38	0.41	0.42	0.42	0.43
$\sigma \times 10^{-2}$	1.26	2.03	2.05	0.31	0.96	2.60	0.01	0.83	1.37	1.54	1.62	2.05	2.46	2.55	2.60
$\lambda \times 10^{-2}$	4.00	7.97	8.01	1.60	1.00	11.00	0.18	4.28	3.77	4.86	6.27	8.01	10.09	10.82	11.00
$\beta \times 10^{-2}$	10.00	10.98	11.14	1.87	1.08	15.00	0.04	2.11	6.71	7.54	8.56	11.14	13.05	13.37	13.91
a	0.60	0.85	0.89	0.09	0.33	0.95	0.07	0.26	0.57	0.64	0.72	0.89	0.91	0.91	0.92
$\sigma \times 10^{-2}$	1.26	2.08	2.08	0.28	0.95	3.17	0.01	0.87	1.51	1.62	1.75	2.08	2.48	2.55	2.63
$\lambda \times 10^{-2}$	4.00	8.75	8.79	0.95	1.20	13.51	0.23	4.84	6.70	7.20	7.70	8.79	9.86	10.15	10.62
$\beta \times 10^{-2}$	10.00	13.03	13.17	1.49	1.21	17.27	0.11	3.38	9.81	10.22	11.20	13.17	14.73	15.06	15.56
a	1.00	1.40	1.40	0.27	1.39	1.41	0.16	0.40	1.40	1.40	1.40	1.40	1.40	1.40	1.40
$\sigma \times 10^{-2}$	1.26	1.41	0.95	0.54	0.95	2.69	0.00	0.56	0.95	0.95	0.95	0.95	2.18	2.35	2.44
$\lambda \times 10^{-2}$	4.00	11.20	11.67	1.53	7.40	14.28	0.54	7.36	7.58	7.91	8.85	11.67	13.01	13.28	13.47
$\beta \times 10^{-2}$	10.00	15.20	15.29	1.04	12.29	20.57	0.29	5.38	13.31	13.51	13.88	15.29	16.39	16.68	17.13
a	10.00	8.00	8.00	0.09	8.00	8.00	4.00	2.00	8.00	8.00	8.00	8.00	8.00	8.00	8.00

Results for the Whittle spectral estimate with SNJD samples for some values of the parameter a. The first two columns display the parameters under study and their true values. Third to ninth columns exhibit the mean, median, standard deviation, minimum, maximum, mean squared error and root mean squared errors, respectively. The tenth to the last columns provide different percentile values of the estimate distribution. We have assumed a jump size U lognormally distributed of the form $U = e^{-\beta^2/2+\beta\varepsilon} - 1$, with $\varepsilon \sim N(0,1)$, and response function of the form $h(t) = e^{-at}$.

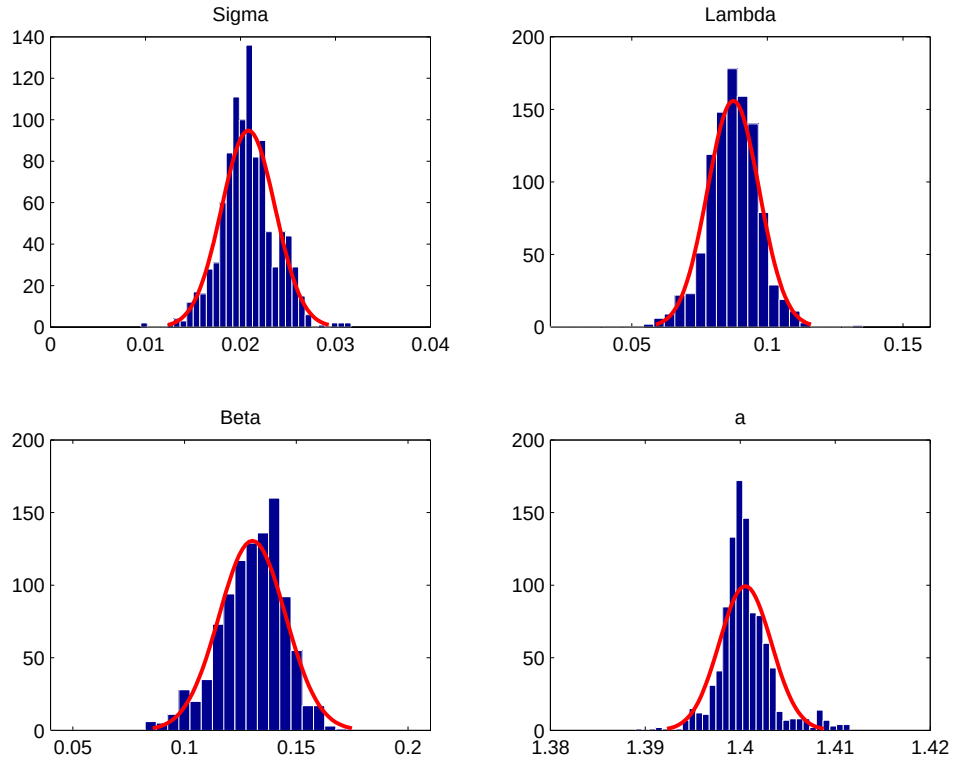


Figure 3.9: Sample distribution for spectral Whittle estimates. The red line corresponds to the normal distribution fit of the data. The sample is composed by estimations over 3,000 paths with 1,800 data each one, corresponding to daily frequency. Simulation parameters are $\mu = 0.0002$, $\sigma = 0.0126$, $\lambda = 0.0400$ and $\beta = 0.1000$. The parameter a has been fixed to 1.00. The jump size U is lognormally distributed of the form $U = e^{-\beta^2/2+\beta\varepsilon} - 1$, with $\varepsilon \sim N(0,1)$ and mean $E[U] = 0$.

for the case of $a = 0.60$, the true value of the estimate falls into a 95%-confidence interval.

Although the sample under study is not large enough to infer the asymptotic distribution of the estimates, it could be interesting to look at their sample distribution. To provide additional tools for analyzing the results, Figure 3.9 shows the sample distribution of the Whittle estimates for the case of $a = 1.00$. We have also included the fitted normal distribution plot of the data (red line).

As Figure 3.9 reflects, the sample estimate distributions do not seem to be far away from the normal one: almost every histogram in this Figure

appears to be symmetrical. Moreover, the normal fit appears to offer a good fit to the estimate sample distributions, just observing a clear deviation from the normal distribution in the case of the a parameter. Again, these results must be taken carefully, due to the short size of the sample used.

3.6 Comments on spectral estimate results

Once checked the individual performance of the spectral estimates under study, it could be interesting to compare them by looking at their sample distributions. In addition to this, since the results for the time domain estimate on Chapter 2 are at our disposal, a more global perspective can be provided by comparing the results obtained in the frequency domain estimates with those results obtained in Chapter 2.

Then, this section compares the main results obtained for the frequency and time domain chapters. Firstly, we focus on the case of the frequency domain estimates (naive, Whittle) and, then, we extend our study by including the GMM estimates of Chapter 2.

3.6.1 Previous considerations

As it was previously cited, we start by analyzing the obtained results across spectral estimates. Our procedure simply consists on computing the histograms for each sample distribution of the different estimates -two frequency domain and one time domain estimates-, and fitting the obtained data by means of a cubic spline data interpolation. The obtained results are represented in Figures 3.10 and 3.11, respectively.

Figures 3.10 and 3.11 display the sample distribution for the spectral and time domain estimates obtained for a SNJD process with exponential response function with parameters $a = 0.60$ and 1.00, respectively. Each graph in these figures exhibit the fitted histograms for the Whittle (red line), naive (blue line) and GMM (black line) estimates[29].

Moreover, each graph also corresponds to the different parameter estimates under study: upper graphs, from left to right, correspond to σ and λ values; lower graphs, also from left to right, correspond to β and a parameters, respectively.

Before discussing these figures, some considerations should be taken into account:

[29]The GMM estimate results correspond to those obtained by using the Newey-West (1987) estimate as weighting matrix. We have chosen this estimate as it is the suitable election in presence of serial correlation in the sample.

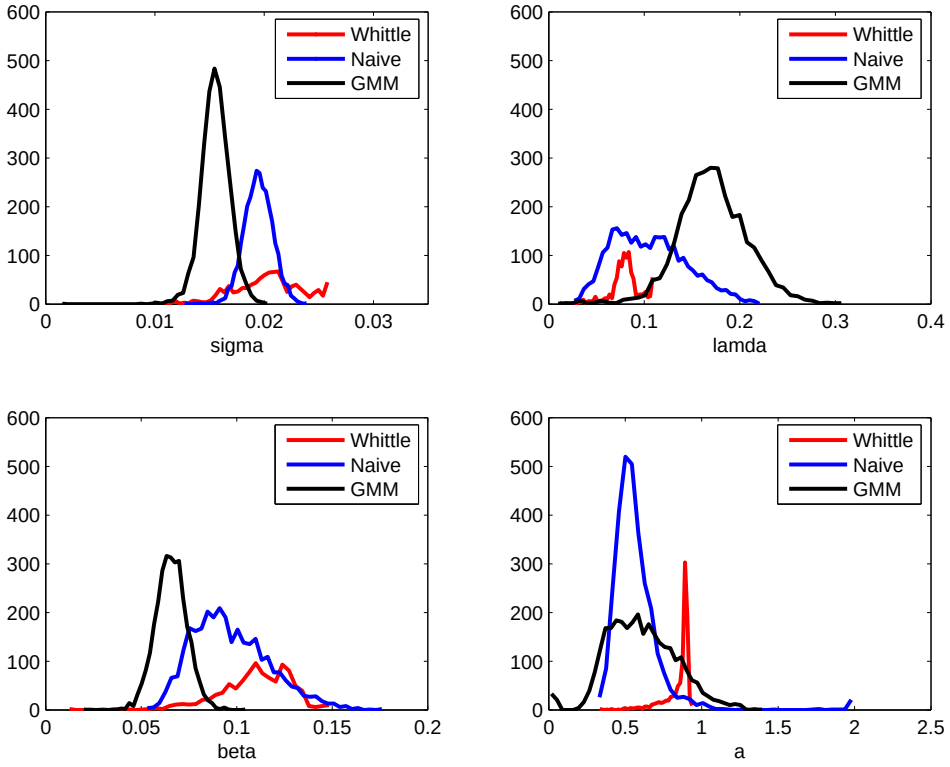


Figure 3.10: Sample distribution for spectral and time domain estimates. Red, blue and black lines correspond to Whittle, Naive and GMM estimate distributions, respectively. The sample is composed by estimations over 3,000 paths with 1,800 data each one, corresponding to daily frequency, unless the Whittle estimate, which are estimations over 1,000 paths. Simulations parameters are $\mu = 0.0002$, $\sigma = 0.0126$, $\lambda = 0.0400$ and $\beta = 0.1000$. Parameter a has been fixed to 0.60. The jump size U is lognormally distributed of the form $U = e^{-\beta^2/2+\beta\varepsilon} - 1$, with $\varepsilon \sim N(0,1)$ and mean $E[U] = 0$.

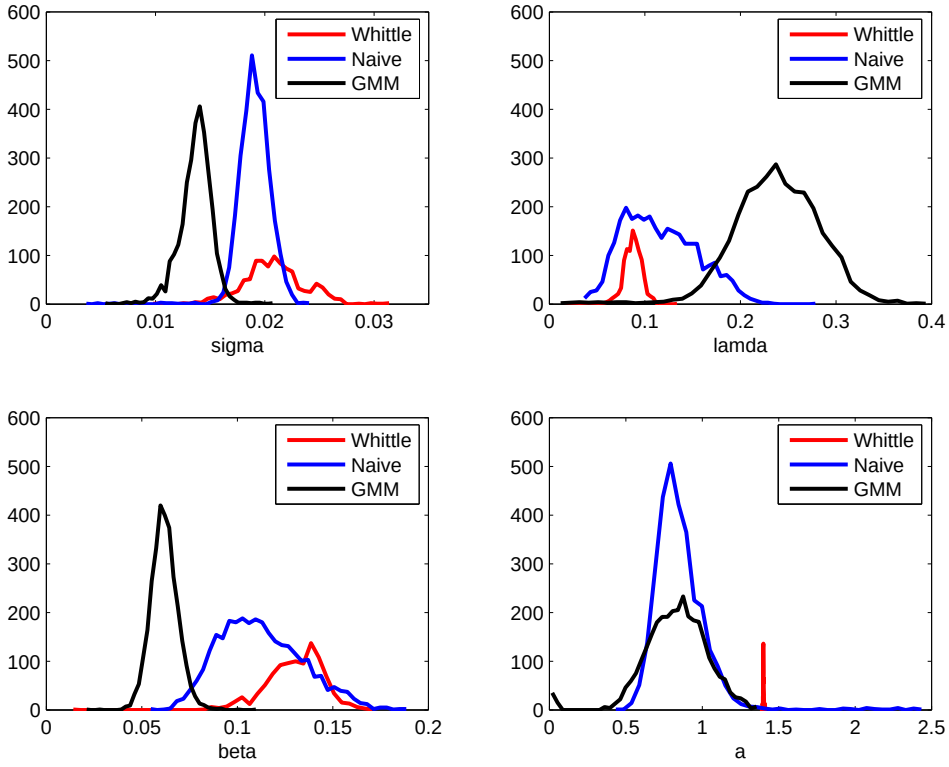


Figure 3.11: Sample distribution for spectral and time domain estimates. Red, blue and black lines correspond to Whittle, Naive and GMM estimate distributions, respectively. The sample is composed by estimations over 3,000 paths with 1,800 data each one, corresponding to daily frequency, unless the Whittle estimate, which are estimations over 1,000 paths. Simulations parameters are $\mu = 0.0002$, $\sigma = 0.0126$, $\lambda = 0.0400$ and $\beta = 0.1000$. Parameter a has been fixed to 1.00. The jump size U is lognormally distributed of the form $U = e^{-\beta^2/2+\beta\varepsilon} - 1$, with $\varepsilon \sim N(0,1)$ and mean $E[U] = 0$.

1. First, as mentioned previously in Section 5, the problems encountered during the computation of the Whittle estimate forced us to restrict the number of optimizations to 1,000 values, instead of the 3,000 ones obtained for the naive and GMM estimates.

   Nevertheless, we are inclined to believe that the inclusion of the Whittle estimate results in Figures 3.10 and 3.11 could be useful, *even though* these results are not conclusive[30].

2. Second, the cases that we will compare directly are those corresponding to the SNJD processes with exponential response function for the cases $a = 0.60$ and 1.00. Two main reasons motivate this approach:

   (a) Due to the numerical approximations involved when computing the moments and spectra of the SNJD process (see Chapter 1), we think that the SNJD process with exponential decaying function is reasonably well approximated in a certain range of parameter a values.

   (b) As a consequence of the former, the results for the restricted SNJD processes in chapters 2 and 3 may indicate that the best performance of the estimates is achieved for $a = 0.60$ and 1.00 values.

Of course, some objections could be done to this procedure, as we are just taking the cases with best performance. Although this affirmation could be true, we believe that by analyzing these cases we can learn more about the estimates under study, since it is *the best* we can do for estimating the SNJD process, at light of the results obtained in the previous chapter. In our opinion, exploring the regions where we have evidences that the model partially (or even not) goes well instead of those situations with better behavior just could lead us to erroneous conclusions.

3.6.2 Comparison between the spectral estimates

We start by analyzing the two spectral estimates under study in this chapter: the naive and Whittle estimates. Some interesting points about the behavior of both estimates are revealed in Figures 3.10 and 3.11.

On the whole, we observe that the naive and Whittle estimates seem to present (almost) the same mode for the parameters σ and λ although it

[30] Again, although the Whittle sample is not very large, these results could indicate how things might go in larger samples.

considerable differs on the β and a cases. This fact could be reflecting an equal pattern behavior in the case of the former parameters, contrary to the last ones.

Secondly, the parameters σ and λ appear to be biased in both $a = 0.60, 1.00$ cases across estimates. Contrary to this point, the observed bias in the case of the β and a parameters is considerably smaller (almost non-existing) for the naive estimate than for the Whittle one.

This last issue refers to the dispersion of the different sample distributions. For example, the sample distribution for the Whittle estimate tends to be smaller than the naive one in the cases of λ and a parameters. By contrast, the Whittle estimate presents a higher dispersion than the naive estimate in the σ parameter. Regarding the parameter β, we are not able to obtain additional conclusions.

Taking everything into account, the naive estimate reveals as a simple and adequate tool for the estimation in the frequency domain, under the sample studied. Noting the simplicity of its implementation, we are inclined to think that it should be considered as a previous step on any kind of spectral estimation, in a similar way as the Identity matrix does in GMM estimations, a suggestion done by Cochrane (2005).

With a similar importance, although results are not conclusive, our estimations seem to show that the Whittle estimate could be a candidate for the spectral estimation, as many of its properties -consistence, asymptotical efficiency- seem to be reflected even in a short sample study as that presented here.

3.6.3 Comparison between frequency and time domain estimates

We analyze now the results concerning both frequency and time domain estimates. As it was previously cited, we will focus on the naive and Whittle estimates (frequency domain) and the GMM estimate (time domain). Due to the presence of serial dependence, we have chosen the GMM results based on the Newey-West (1987) weighting matrix.

The main interesting features that arise from Figures 3.10 and 3.11 can be summarized as follows:

1. Firstly, the GMM estimate seems to be a considerably less biased estimate that those of the frequency domain in the case of σ. Moreover, we do not appreciate a bias reduction when the parameter a increases.

Apparently, it is not noticeable a reduction on the distribution dispersion of GMM estimate with reference to the naive estimate.

2. Secondly, regarding the parameter λ, the GMM estimate presents a higher bias than their frequency domain competitors. Moreover, the bias of the GMM estimate tends to increase with the parameter a. Finally, the GMM estimate dispersion is comparable with that of the naive estimate in both a cases under study, and much higher than the Whittle estimate.

3. Concerning to the parameter β, the dispersion of the GMM estimate is lower than the other spectral estimates. By contrast, its bias is higher, again.

4. Finally, the GMM estimate and naive estimates seem to present a similar performance for the parameter a, in terms of bias and dispersion. By contrast, the Whittle estimate presents a higher bias and a lower dispersion than its competitors.

All things considered, the most interesting conclusions concern the usefulness of the spectral analysis tools when dealing with serial dependent process. As Figures 3.10 and 3.11 reveals, the behavior of the Whittle and (specially) the naive estimates seem to be more stable (the modes remain almost constant under changes in the parameter a than their time domain competitor. On the whole, the frequency domain estimates are less biased in all cases except when considering the parameter σ. With regard to their dispersion, this aspect varies across estimates and parameters.

As it was noticed in the previous subsection, the naive estimate reveals again as an interesting tool for estimating the parameters of the SNJD process. However, due to the simple nature of this estimate, we think that its efficiency could be improved. Finally, the Whittle estimate may provide interesting results (although not conclusive) on larger sample studies.

Once again, we would like to point out that the comparison among the time and frequency domain estimates of this subsection should not be taken as definite, as the samples under study must be enlarged to capture other possible scenarios in the simulation data. However, as far as we are concerned, we tend to interpret the following results as a starting point for further studies.

3.7 Conclusions

This chapter explores the estimation in the frequency domain of the SNJD process posited by Altmann *et al* (2007). In this way, this chapter presents two main tools for carrying out this task: firstly, the naive estimate, directly inspired on Bevan *et al* (1979), and an alternative technique, based on Whittle (1953).

The first part of this chapter is devoted to introduce the spectral approach and its intuitions behind. We also discuss the estimates under study, presenting some of their features. To study the ability of the posited estimates to capture the SNJD parameters, we have also performed an exhaustive Monte Carlo study.

To sum up, the main results of this paper are threefold: firstly, by using the analytical expression for the spectrum of the SNJD process of Chapter 1, we provide an econometric methodology based on the spectral analysis to estimate their parameters. Secondly, the spectral approach reveals as a useful tool for the estimation of the posited process and its extension to more general continuous-time models should be explored. Thirdly, by a direct comparison with the GMM results obtained in Chapter 2, our results seem to confirm the suitability of this procedure against other time domain alternatives as GMM ones.

It should be emphasized that this chapter must be intended as a first attempt to deal with this kind of processes, since the financial literature about shot noise effects and its estimation has not been previously reported. Although the comparison between time and frequency domain estimates of this chapter should not be taken as definite, we think that the obtained results support the idea that the spectral estimation should be considered as an additional tool for estimating continuous-time dependent processes.

Part II

Credit Risk Valuation Techniques

Chapter 4

Pricing Tranched Credit Products with Generalized Multifactor Models

This chapter[1] is devoted to pricing a very recent type of credit derivative named Collateral Debt Obligation, or CDO. We present some of their main features, focusing on its valuation. We will also discuss the standard model used for the pricing of CDO, generalyzing some of their assumptions. Finally, an empirical application with Moody's data is also included.

4.1 Introduction

The market of credit tranched products is one of the fastest growing segments in the credit derivative industry. As an example, Tavakoli (2003) reports an increase in market size from almost $19 billions in 1996, to $200 billions in 2001. Recent reports estimate market size to be $20 trillions in

[1]I would like to thank to A. Novales and J.I. Peña for their helpful comments and suggestions. I also acknowledges financial support from the Plan Nacional de I+D+I (project BEC2003-02084) and the Gobierno Vasco-UPV project (GIU 06/53), especially to Jose M. Usategui. Previous drafts of this chapter have been accepted to be presented on XV Foro Finanzas AEFIN, 15th-16th November 2007, Palma de Mallorca, Spain, and XXXII Simposio de Análisis Económico, 13rd-15th December 2007, Granada, Spain. I have also benefit from the comments of participants in seminars at University of Murcia, Murcia, Spain and Universidad Complutense, Madrid, Spain. Finally, part of this work was developed during a visiting scholar position at Business Administration department at University Carlos III, Madrid, Spain. I am sincerely grateful to them.

2006.[2] As a result, an increasing attention of the financial sector audience has focused on the pricing of these new products.

The development of pricing models for multiname derivatives is relatively recent. As pointed out by Hull and White (2004), the standard approach on the credit risk literature tends to subdivide the pricing models for multiname derivatives in two groups: structural models, which are those inspired in Merton's (1974) model or Black and Cox (1976), or the intensity based approach, like Duffie and Garlenau (2001). Roughly speaking, their differences remain on how the probability of default of a firm is obtained: using their fundamental variables - assets and liabilities - as in the case of the structural models, or using directly market spreads, as in the intensity models approach. Up to a point, the structural based approach has been extensively implemented by the financial sector, maybe due to the extended use of industrial models like Vasicek (1991) or Creditmetrics model of Gupton *et al* (1997).[3] However, recent academic literature analyze the prices of CDO tranches using intensity models, as Longstaff and Rajan (2006). We refer to Bielecki and Rutkowski (2002) for a general presentation of structural and intensity based models.

This paper presents an extension of the standard Gaussian model of Vasicek (1991), in line with the structural models literature. Basically, Vasicek's (1991) model assumes that the value of a firm is explained by the weighted average of one common factor for every asset plus an independent idiosyncratic factor. By means of linking the realization of one systematic factor to every firm's values, Vasicek (1991) provides a simple way to reduce the complexity of dealing with dependence relationships between firms. Gibson (2004) or Gregory and Laurent (2004, 2005) provide additional insights about risk features of this model.

Our approach relies on the connection between the changes of value of a firm and the sum of two factors: systematic and idiosyncratic. Our approach overcomes the limitations of the standard Gaussian model: the different areas or *regions* of correlation that could compose a credit portfolio (see Gregory and Laurent, 2004). This article presents a model that captures some of the facts found in real data. Motivated by this fact, this article proposes an extension to the two Gaussian asset classes as in Schönbucher (2003). Our paper extends the existing literature in three ways: firstly, the

[2] See BBA Credit Derivatives Report (2006).

[3] It is worth mentioning that the appearance of techniques within the Structural framework that diminishes the traditional high computing cost of multiname credit derivatives (see Andersen *et al* (2003) or Glasserman and Suchitabandid (2006), among others) has contributed to the widely usage of structural models.

assumption of asset homogeneity is relaxed by introducing two asset classes. Secondly, we consider an additional source of systematic risk by including another common factor related with industry factors. Finally the normality assumption on common factors is relaxed.

This paper is divided as follows: Section 2 presents the model. Section 3 studies the sensitivity to correlation and to changes in credit spreads. Section 4 addresses the econometric modelling. Finally, some conclusions are presented on section 5.

4.2 The model

To motivate our model we discuss some empirical features found in CDO data that are not captured by the standard Gaussian models and propose an extension to the asset class models posited by Schönbucher (2003). Notation is taken from Mardia *et al* (1979).

4.2.1 The standard Gaussian model

The Gaussian model introduced by Vasicek (1991) has become a standard in the industry. Basically, it addresses in a simple and elegant way the key input in CDOs price: the correlation in default probabilities between firms affects the price of the CDO.

Usually a CDO is based on a large portfolio of firms bonds or CDS.[4] Let $\mathbf{V}_{n\times 1}$ (subscript denotes matrix dimension) be a random vector with mean zero and covariance matrix $\mathbf{\Sigma}$. As standard notation in multivariate analysis, we will define the p-factor model as

$$\mathbf{V}_{n\times 1} = \mathbf{\Lambda}_{n\times p}\mathbf{F}_{p\times 1} + \mathbf{u}_{n\times 1} \tag{4.4.2.1}$$

where $\mathbf{F}_{p\times 1}$ and $\mathbf{u}_{n\times 1}$ are random variables with different distributions.

The interpretation of the model (4.4.2.1) is the following:

- The vector $\mathbf{V}_{n\times 1}$ represents the value of the assets for each of the individual n-obligors.
- The vector $\mathbf{F}_{p\times 1}$ captures the effect of systematic factors - business cycle, industry, etc. - that affect to the whole economy.
- By contrast, $\mathbf{u}_{n\times 1}$ represents the idiosyncratic factors that affect each of the n-companies.

[4] CDO tranches of NYME are composed by 100 firms.

- Finally, $\mathbf{\Lambda}_{n\times p}$ is called the *loading* matrix, and determines the correlation between each of the n-firms.

By assumption, we have that

$$E(\mathbf{F}) = \mathbf{0}, \qquad Var(\mathbf{F}) = \mathbf{I}, \tag{4.4.2.2}$$

$$E(\mathbf{u}) = \mathbf{0}, \qquad cov(\mathbf{u_i}, \mathbf{u_j}) = 0, \qquad i \neq j, \tag{4.4.2.3}$$

and

$$cov(\mathbf{F}, \mathbf{u}) = \mathbf{0} \tag{4.4.2.4}$$

where $\mathbf{I}$ is the identity matrix.

We will also assume that vector $\mathbf{u}_{n\times 1}$ is standardized to have zero mean and unit variance.

Using a simplified form of equation (4.4.2.1), Vasicek (1991) assumes that firm's values in the asset pool backing the CDO are affected by the sum of two elements: on one hand, a common factor to every firm which represents the systematic component represented by the factor F; on the other hand, an idiosyncratic component modelled by a noise ε_i. Both are assumed to be standard N(0,1) random variables

$$V_i = \rho_{i1} F_1 + \sqrt{1 - {\rho_{i1}}^2}\varepsilon_i, \quad \text{with } i \leq n \tag{4.4.2.5}$$

By means of equation (4.4.2.5) it is possible to capture the correlation between different firms in a portfolio. As equation (4.4.2.5) reveals, Vasicek (1991) assumes that correlation coefficient is homogeneous for each pair of firms.[5] Additionally, the simplicity of their assumptions permits a fast computation of CDO prices under this framework.

As an immediate consequence of equation (4.4.2.5), a further step is given by generalizing the number of factors that affects firms values. Thus, Lucas *et al* (2001) consider the following factor model

$$V_i = \sum_{j=1}^{p} \rho_{ij} F_j + \sqrt{1 - \sum_{j=1}^{p} {\rho_{ij}}^2} u_i, \text{ with } i \leq n \tag{4.4.2.6}$$

where V_i represents the value of the i-company, F_j, $j = 1, ..., p$ capture the effect of p-systematic factors and u_i is the idiosyncratic factor associated

[5]For a detailed exposition of assumptions in structural models see Bielecki and Rutkowski (2002).

to the i-firm. Needless to say, assumptions on the distribution of common and idiosyncratic factors (F and u, respectively) could be imposed. Hull and White (2004) also includes an extension to many factors, including the t-Student distributed case. Finally, Glasserman and Suchitabandid (2006) implements in a recent paper numerical approximations to deal with these multifactor structures within a Gaussian framework.

With the purpose of getting intuition about the behaviour of these different models, Figure 1 exhibits the loss distribution generated by three alternative models:

- The standard Gaussian model (Vasicek, 1991),

$$V_i^G = \rho_{i1} F_1 + \sqrt{1 - \rho_{i1}{}^2}\varepsilon_i, \quad \text{and } F_1, \varepsilon_i \sim N(0,1)$$

- The one factor double-t model (Hull and White, 2004),

$$V_i^S = \rho_{i1} F_1 + \sqrt{1 - \rho_{i1}{}^2}\varepsilon_i,$$

  where F_1 and ε_i follows t-Student distributions, with 6 degrees of freedom.

- The two-factor Gaussian model (Hull and White (2004), or Glasserman and Suchitabandid, 2006),

$$V_i^{DG} = \rho_{i1} F_1 + \rho_{i2} F_2 + \sqrt{1 - \rho_{i1}{}^2 - \rho_{i2}{}^2}\varepsilon_i$$

  with $F_1, F_2, \varepsilon_i \sim N(0,1)$ and $\rho_{i1}^2 + \rho_{i2}^2 < 1$.

As we will see later, the loss distribution plays a role crucial in the CDO pricing. Different distributions are computed for a portfolio of 100 firms, with constant default intensities of 1% per year. Correlation parameters are 0.3 for all models, except for two-Gaussian factor, with 0.1 and 0.3. The double-t model considers a common and idiosyncratic factors each distributed as t-Student with 6 degrees of freedom. The picture shows that, under the same correlation parameters, the standard Gaussian model assigns lower probability to high losses (see for instance the case of 20 firms) than the double-t model. By considering an additional factor, as in the 2-Gaussian case, the probability of high losses is substantially higher than in previous cases.

Table 1 presents the spreads (in basis points) of a 5-years CDO portfolio composed by 100 firms. The individual default probabilities are constant and

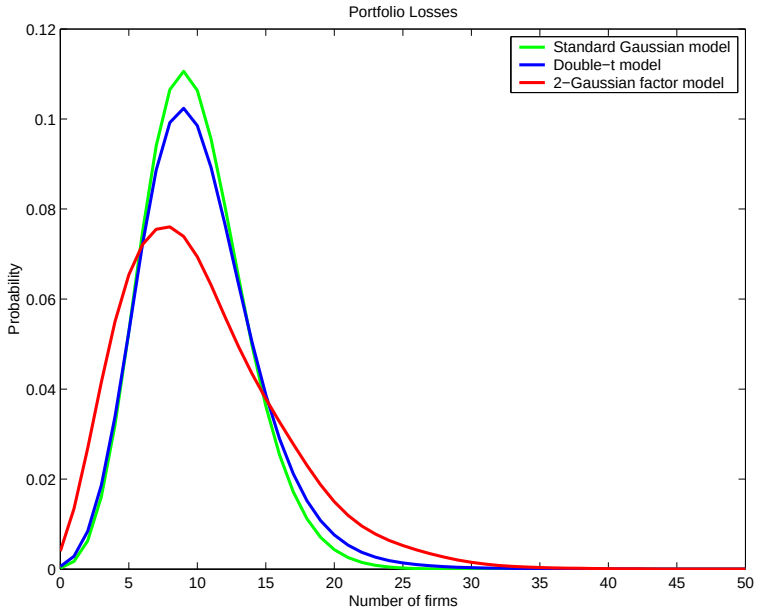


Figure 4.1: Total loss distribution of a 100-firms portfolio for different models. Individual default probabilities are fixed at 1%. Correlation for the Standard Gaussian and t-Student distribution (with 6 degrees of freedom) is 0.3. Correlations for the Double Gaussian model are 0.1, 0.3.

Table 4.1: Spreads of different tranches for a 100-names CDO

	Tranche Spread (basis points)					
Model	Standard Gaussian		2 Gaussian factors		t-Student	
Correlation	0.1	0.3	0.1/0.1	0.1/0.3	0.1	0.3
0–3	2117	1320	2186	1656	2485	2135
3–10	86	110	508	653	131	233
10–100	0.02	1	7	29	1	9

Spreads of different tranches for a 100-names CDO under different correlation parameters. Attachment points (in percentage) are in bold.

fixed at 1%. The recovery rate is 40%, a standard in the market. Finally, the different correlation parameters are displayed in the table. We remember that all spreads are obtained considering only one asset class. As Table 1 shows, results are consistent with the loss distribution lines presented in Figure 1: the double t-Student distribution gives prices systematically bigger than those obtained for the standard Gaussian model, keeping constant the correlation. In the 2 Gaussian factor model prices, we observe a mixture of effects due to different combinations of correlations in the portfolio.

4.2.2 A 2-by-2 model

Generally, as considered by Schonbucher (2003) or Lando (2004), a credit portfolio is composed by different asset classes or buckets, attending to criteria of investment grade, non-investment grade assets or industry, among others. The exposure of a credit portfolio to a set of common risk factors could be significant between groups, but should be homogeneous within them. In line with this, the idea of two groups of assets treated in different ways is a more realistic assumption.

As pointed out by Gregory and Laurent (2004), the one-factor model imposes a limited correlation structure on the credit portfolio, which is not realistic. Initially, one can argue that increasing the number of factors could be enough to capture a richer structure of correlation within the portfolio. However, this is not yet consistent with the idea of heterogeneity correlation among groups, due to the fact that every asset is exposed to the same degree of correlation. By contrast, a richer correlation structure could be imposed

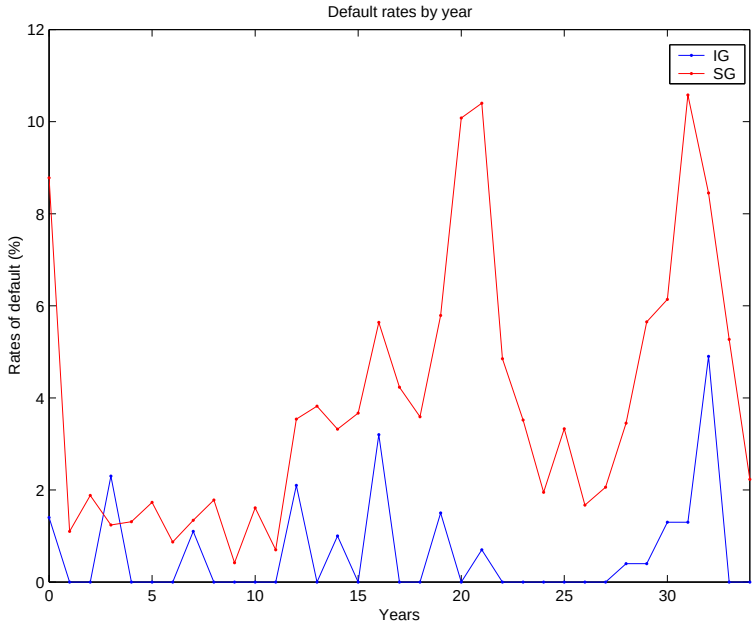


Figure 4.2: Default (yearly) rates for Investment and Non-investment grades. Rates for investment data have been multiplied by ten with the purpose of comparing the data. Source: Moody's.

in the portfolio if these two groups are treated in different ways.

To illustrate these ideas, Figure 2 shows the yearly percentage default rates for investment and non-investment grades. Rates for investment data have been multiplied by ten with the intention of clarifying the exposition. Figure 2 reveals that correlation between these two assets groups varies through time: periods with high degree of default in non-investment grade assets do not match with high default rates for investment grade. With reference to this idea, Figure 3 displays the correlation coefficients computed using a moving window of five years. This also provides some additional insights on the degree of correlation between asset classes: the picture shows how the correlation among different groups changes during the sample period, from negative correlation (1975, 1977 or 1990), to zero (1993, 1996) or highly positive (1982, 1995 or 2000). These differences in default rates through time could reveal the existence of an idiosyncratic component between asset classes. This fact could support the idea of modelling in a different way the behaviour of different assets.

This article considers a family of models that take account the existence

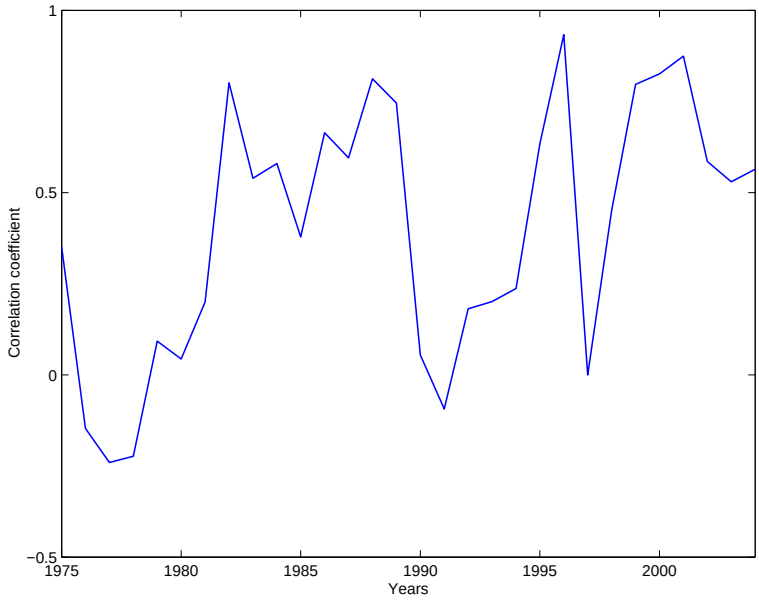


Figure 4.3: Correlation coefficients using a moving window of 5 years. Sample period comprises from 01/01/1975 to 31/12/2004. Source: Moody's.

of these different asset classes or *regions*, in line with the suggestions of Gregory and Laurent (2004). We analyze a model in the line of the two assets-two Gaussian factor model of Schonbucher (2003), where no distinction is made between the obligors which belongs to the same class. Our model generalizes that posited by Schonbucher (2003) by considering a t-Student distribution, which assigns a higher probability to high default events. Empirical evidence seems to go in this direction.[6] Our work contributes to the existing literature in the analysis of these asset class models. To the best of our knowledge, no similar studies has been reported yet in this direction.

The model we propose is a two-by-two factor model as follows[7]

$$\begin{aligned} V_{i1} &= \alpha_{11}F_1 + \alpha_{12}F_2 + \beta_1 u_{i1} \\ V_{i2} &= \alpha_{21}F_1 + \alpha_{22}F_2 + \beta_2 u_{i2} \end{aligned} \tag{4.4.2.7}$$

where

[6] See, for example, Mashal and Naldi (2002).

[7] For the sake of brevity, we omit the graph of th loss distribution implied by this model.

- V_{i1}, V_{i2} $(i < n)$ represents the value of the i-company which belongs to the different asset class

- F_j, $j = 1, 2$ captures the effect of systematic factors - business cycle and industry - with independent t- distributions with n_j degrees of freedom, and

- u_{i1}, u_{i2} are idiosyncratic factors distributed also with n_{i1}, n_{i2} degrees of freedom, respectively.

Under the assumptions of factor models, F_j are scaled to have variance 1, then $\alpha_{11} = \rho_{11}\sqrt{\frac{n_1-2}{n_1}}$ and so on. Idiosyncratic errors are also scaled, for instance $\beta_1 = \sqrt{\frac{n_{i1}-2}{n_{i1}}}\sqrt{1-\alpha_{11}^2-\alpha_{12}^2}$. Finally, we assume the same default barriers K_{i1}, K_{i2} for the obligors of the same class.

Figure 4 displays the simulated distribution for a portfolio of 100 obligors under the standard Gaussian model of Vasicek (1991), the double-t model of Hull and White (2004) and the 2-by-2 double-t model proposed in equation (4.4.2.7). Risk neutral default intensities have been fixed at 10% for asset class A and 15% for asset class B. We consider homogeneous firms of class A for Gaussian and double-t models. The portfolio in the 2-by-2 Double-t model is equally composed of asset classes A and B. Finally, correlation parameters are 0.2 for the standard Gaussian and double-t models, and 0.1, 0.2 for the 2-by-2 double-t model, respectively. As Figure 4 exhibits, the 2-by-2 model allocates more probability in the tail of the loss distribution than one factor models, which results in an increasing (decreasing) value for risky (safe) tranches.

Needless to say that the model could be easily generalized to the case of m-asset classes, as follows:

$$V_{i,m} = \sum_{h=1}^{p} \rho_{mh} F_h + u_{i,m}\sqrt{1 - \sum_{h=1}^{p} \rho_{mh}{}^2}$$

where $V_{i,m}$ represents the value of the i-company which belong to the m-asset class, F_j, $j = 1, ..., p$ capture the effect of systematic factors and $u_{i,m}$ is the idiosyncratic factor corresponded to i-firm of the $m-$asset class. Generally speaking, assumptions relying on distribution factors or more asset classes could also be proposed. However, a trade-off between accuracy, parsimony and computing efficiency must be considered.

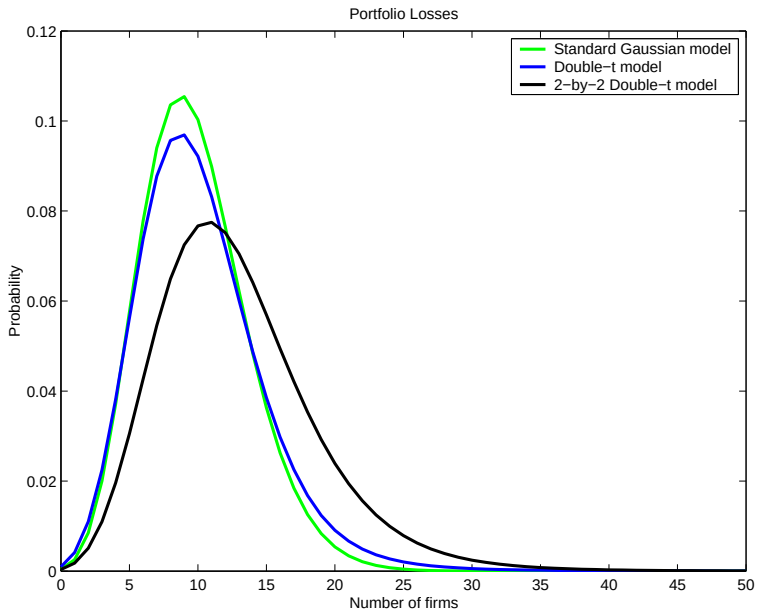


Figure 4.4: Total loss distribution of a 100-firms portfolio for standard Gaussian, Double-t and 2-by-2 Double-t model. Degrees of freedom in the double-t models have been fixed at six for every variable. Individual default probabilities are fixed at 10% for asset class A and 15% for asset class B. All firms in Gaussian and Double-t cases belong to asset class A. The portfolio in the 2-by-2 Double-t model is equally composed of asset classes A and B. Finally, correlation for the standard Gaussian model and Double-t is 0.2, and correlations for the Double-t model are 0.1, 0.2, respectively.

4.2.3 Conditional Default Probabilities

Without loss of generality we omit the subscript that refers to the i-firm for the ease of exposition. We want to study the probability of default for the i-firm which belongs to an asset class m, with $m = 1, 2$,

$$P\left[V_m \leq K|\mathbf{F} = \mathbf{f}\right] = P\left[u_m \leq \frac{K - \alpha\mathbf{F}}{\beta_m}|\mathbf{F} = \mathbf{f}\right] = T_m\left(\frac{K - \alpha\mathbf{f}}{\beta_m}\right)$$

where T_m denotes the distribution function of a t-student with n_m degrees of freedom for the i-firm, and $\mathbf{F}$, α are the common vector factors and their coefficients, respectively.

It is usual to calculate the probability of having k default events, conditional to realization of vector factor $\mathbf{F}$ as

$$P[X = k|\mathbf{F}] = \sum_{l=0}^{k} b\left(l, N_1, T_1\right) b\left(k - l, N_2, T_2\right)$$

where $b\left(l, N, T\right)$ denotes the binomial frequency function, which gives the probability of observing l successes with probability T, where N represents the number of firms which belong to each asset class. In the same manner, the unconditional probability of k default events is obtained considering all possible realizations of factors $\mathbf{F}$

$$P[X = k] = \sum_{l=0}^{k} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} b\left(l, N_1, T_1\right) b\left(k - l, N_2, T_2\right) \psi\left(\mathbf{f}\right) d\mathbf{f}$$

where $\psi\left(\mathbf{f}\right)$ denotes the probability density function of $\mathbf{F}$.

Finally, the total failure distribution is just obtained as the sum of all the defaults up to level k,

$$\begin{aligned} P[X \leq k] &= \sum_{r=0}^{k} P[X = r] \\ &= \sum_{r=0}^{k} \sum_{l=0}^{r} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} b\left(l, N_1, T_1\right) b\left(r - l, N_2, T_2\right) \psi\left(\mathbf{f}\right) d\mathbf{f} \end{aligned} \quad (4.2.8)$$

Lando (2004) refers to this problem as a *different buckets* problem in sense that by means of multinomial distributions we compute the total loss distribution of the portfolio. Our approach to this point would be the computation of this set of bucket probabilities but we will adopt a different approach.

4.2.4 Loss distribution

As is pointed out in Lando (2004), calculation of multinomial expressions like (4.4.2.8) is burdensome. Instead of computing (4.4.2.8) by brute force, we will use a simple idea due to Andersen *et al* (2003) that could improve the efficiency in terms of computing cost.[8]

Andersen *et al* (2003) provides an efficient algorithm which has become widely used by the industry. Roughly speaking, the main idea behind is to observe what happens to the total loss distribution of a portfolio when we increase its size by one firm.

Consider a portfolio that includes n credit references and let $p_n\left(i|\mathbf{f}\right)$ be the probability of default of i firms in this portfolio conditional on factors $\mathbf{f}$. Let $q_{n+1}\left(\mathbf{f}\right)$ be the default probability of an individual firm that is added to this portfolio. These two probabilities are conditional on the realization of the common factor vector $\mathbf{f}$.

Consider the total Loss Distribution (LD) in this portfolio. Intuition says that the probability of i-defaults in this portfolio - conditional on $\mathbf{f}$ - can be written as

$$LD\left(i|\mathbf{f}\right) = p_n(i|\mathbf{f}) \times \left(1 - q_{n+1}\left(\mathbf{f}\right)\right) + p_n(i-1|\mathbf{f}) \times q_{n+1}(\mathbf{f}), \quad 0 < i < n+1$$

where the first term reflects that default is due to i-firms included in the initial portfolio while the new reference (just included in the portfolio) survives. In a similar way, the second term reflects the new firm (just included in the portfolio) defaults while the other $i-1$ defaulted firms were previously included in the original portfolio.

Moreover, for the extreme cases of default firms, we have

$$\begin{aligned} p_{n+1}(0|\mathbf{f}) &= p_n(0|\mathbf{f}) \times \left(1 - q_{n+1}\left(\mathbf{f}\right)\right) \\ p_{n+1}(n+1|\mathbf{f}) &= p_n(n|\mathbf{f}) \times q_{n+1}(\mathbf{f}) \end{aligned}$$

Then, taking into account the last firm just included in the portfolio, these equations reflect that no firm defaults or all of them default, respectively.

As an example, consider an initial portfolio including two firms. Adding a third firm, the total default probabilities are given by

$$\begin{aligned} p_3\left(0|\mathbf{f}\right) &= p_2\left(0|\mathbf{f}\right)\left(1 - q_3\left(\mathbf{f}\right)\right) \\ p_3\left(1|\mathbf{f}\right) &= p_2\left(1|\mathbf{f}\right)\left(1 - q_3\left(\mathbf{f}\right)\right) + p_2\left(0|\mathbf{f}\right) q_3\left(\mathbf{f}\right) \\ p_3\left(2|\mathbf{f}\right) &= p_2\left(2|\mathbf{f}\right)\left(1 - q_3\left(\mathbf{f}\right)\right) + p_2\left(1|\mathbf{f}\right) q_3\left(\mathbf{f}\right) \\ p_3\left(3|\mathbf{f}\right) &= p_2\left(2|\mathbf{f}\right) q_3\left(\mathbf{f}\right) \end{aligned}$$

[8] Their contribution has been also explored and extended in Hull and White (2004).

Using this iterative procedure, we can compute the unconditional total Loss Distribution by considering all possible realizations of $\mathbf{f}$.[9]

4.3 Results for the 2-by-2 model

In this section we present some results of the model (4.4.2.7). Firstly, we discuss the spreads obtained using the two-by-two approach. Secondly, we give an approach useful for cases of high degrees of freedom based on Cornish-Fisher expansions, which is useful in terms of computing cost.

4.3.1 Numerical results

To give some results of model (4.4.2.7), we analyze different cases for a two asset classes, 5-year CDO with 100 firms with quarterly payments. We also assume that the recovery rate is fixed and equal at 40%. Risk-neutral default individual default probabilities are fixed at 1% and 5% for assets which belong to class 1 and 2, respectively. For ease of explanation, the size of each asset class in the portfolio is the same (50% for each one). More results concerning the size of the portfolio will be provided in Section 4.

Table 2 displays the main results obtained. First row corresponds to the three simulated base cases: the standard Gaussian model[10] with one factor, two asset classes; the two assets-two Gaussian factor; finally, the two assets-two t-Student factors.[11] t-Student distributions have been fixed at 5 degrees of freedom for idiosyncratic and systematic factors. Second row displays the correlation parameters of *each* asset class with *both* factors. For example, 0.1/0.3 in the two Gaussian case refers to a correlation of 0.1 (0.3) for elements of class 1 (2) with *the two* systematic factors. In line with Gregory and Laurent (2004), our idea is to check the behaviour of the CDO to a portfolio exposed to two different degrees of correlation.

All the simulations have been carried on for different tranches values. Looking at the riskier tranche (equity tranche), we observe that, for the same degree of correlation, its spread is systematically bigger for the two t-Student case than for the 2 Gaussian factors. One conclusion that arises from Table 2 is that the two assets-one Gaussian factor spreads for equity tranche are close to those values obtained for two assets-two t-Student factors.

[9] See Andersen *et al* (2003) or Gibson (2004) for more details.

[10] The standard Gaussian model has been computed using a Gauss-Hermite quadrature with 8 nodes.

[11] The t-Student simulations have been computed using a Simpson's quadrature with 25 points.

Table 4.2: Spreads of different tranches for a 100-names CDO

	Tranche Spread (basis points)					
Model	Standard Gaussian		2 Gaussian factors		t-Student	
Correlation	0.1/0.1	0.1/0.3	0.1/0.1	0.1/0.3	0.1/0.1	0.1/0.3
0–3	7166	5016	5421	3154	6899	4977
3–10	1040	929	1581	1401	1254	1302
10–100	5	11	43	122	27	65

Spreads of different tranches for a 100-names CDO under different correlation parameters. Attachment points (in percentage) are in bold.

As expected, the same does not apply for mezzanine tranches: the addition of more factors provides more weight to extreme default events, which results in an increase in spread of mezzanine tranches. The same conclusions apply to senior tranche.

4.3.2 Approximation for n infinite

The asymptotic relationship between a t-distributed random variable and a normal random variable by means of the Cornish-Fisher expansion could be interesting for cases of big degrees of freedom. Shaw (2006) provides the Cornish-Fisher expansion for a t-distributed random variable, with zero mean and unit variance, in terms of a standard normal random variable distribution. This reduces considerably the computational time as, in this case, it is possible to use a double Gauss-Hermite quadrature, instead of a Simpson quadrature. Basically, Shaw (2006) provides the relationship

$$s = z + \frac{1}{4n} z\left(z^2 - 3\right) + \frac{1}{96n^2} z\left(5z^4 - 8z^2 - 69\right) + ... \qquad (4.4.3.1)$$

where s is the t-distributed random variable with n degrees of freedom, and z is a standard normal random variable. To check the accuracy of the approximation (4.4.3.1), Table 3 displays the spreads (in basis points) for different tranches in a two asset classes, 50-named CDO with quarterly payments under different correlation parameters for various degrees of freedom. As in the previous section, correlation parameters correspond to factors of

Table 4.3: Spreads of different tranches for a 50-names CDO

	Tranche Spread (basis points)				
Model	2 Gaussian factors	t-Student		CF t-Student	
Correlation	0.1/0.3	0.1/0.3	0.1/0.3	0.1/0.3	0.1/0.3
n_1/n_2	∞/∞	10/10	15/15	10/10	15/15
0–3	3015	3924	3711	3270	3296
3–10	1380	1236	1198	1025	1083
10–100	289	160	160	166	160

Spreads of different tranches for a 50-names CDO under different correlation parameters. Attachment points (in percentage) are in bold. Coefficients n_1/n_2 correspond to the number of degrees of freedom.

each asset classes. As in previous examples, risk neutral default probabilities have been also fixed at 0.1 and 0.3 for each asset class. The recovery rate is fixed at 40%.

Basically, two general models have been computed: two Gaussian factors and a two t-Student factor. The last column refers to the two t-Student factor approximation using the Cornish-Fisher expansion. Correlation coefficients and degrees of freedom for each one are displayed in the table.[12] The Gaussian case is presented to get intuition of how far we are from the asymptotic result. As expected, the larger the degree of freedom, the higher the accuracy of the results. The differences between the equity tranche range from 20% for 10 degrees of freedom to 12% for a 15 degrees of freedom case. In a similar way, considering the mezzanine cases, differences go from 21% to 11%. There are no substantial changes in the case of senior tranche. It is worth to remember that computations under the *exact* t-Student distribution have been done using a numerical quadrature and, so, they are exposed to numerical errors.

[12]The two Gaussian factor model is equivalent to a two t-Student factor model with infinite degrees of freedom.

4.4 Sensitivity analysis

Now, we are interested in the prices of the CDO under two different scenarios: changes in portfolio size and correlation. Firstly, we present some standard measures in risk management as the Value at Risk and the Conditional Value at Risk ones. Secondly, we analyze the sensitivity of different tranches to changes in correlation.

4.4.1 VaR and CVaR

Value at Risk (VaR) and Conditional Value at Risk (CVaR) are usually taken as representative risk measures for a portfolio. VaR is defined as the percentile of the distribution of portfolio losses given a certain level of confidence.[13] Artzner, Delbaen, Eber and Heath (1999) enumerates some limitations of the VaR measure and discuss some interesting properties of a proper measure of risk. According to this, we also include the CVaR measure,[14] defined as the expected loss in a portfolio conditional to a certain loss threshold u, that is,

$$CVaR_u = E\left[x|x > u\right]$$

where the sign of the inequality has been changed because we are working directly with the total loss distribution.

Table 4 includes the VaR and CVaR measures for a 100-named CDO composed by two different risky asset classes under different correlation parameters and different proportions in the portfolio. The (yearly) risk-neutral default probabilities for each obligor of asset class 1 have been fixed at 0.01, and 0.05 for those which belong to asset class 2. For the sake of simplicity, all the factors - systematic and idiosyncratic - in the model (4.4.2.7) have been fixed at 5 degrees of freedom. The first column includes the correlation coefficients corresponding to both factors of each class (i.e. 0.1/0.3 means a correlation coefficient of 0.1 (0.3) for *both* factors in asset class 1 (2)). Correlation coefficients equal to zero refers to independence case between obligors. With the purpose of analyze the response of the loss distribution generated by the model (4.4.2.7) with respect to different sizes of the portfolio, the remaining columns show different percentage sizes of the portfolio. The first term refers to the portfolio percentage of class 1, and so on.

Needles to say that two main results arise form Table 4. Firstly, the higher the correlation the higher expected extreme loss as measured by VaR

[13] See Duffie and Pan (1997) for details.

[14] This measure was posited by Acerbi, Nordio and Sirtori (2001).

Table 4.4: VaR and CVaR measures for a 100-named CDO

	VaR 99%			CVaR 99%		
Correlation	25/75	50/50	75/25	25/75	50/50	75/25
0.0/0.0	10	9	7	18	9	9
0.1/0.1	25	19	13	35	27	19
0.1/0.3	57	40	23	67	48	30
0.3/0.3	61	47	33	72	61	50

VaR and CVaR measures for a 100-named CDO composed by two different risky asset classes under different correlation parameters and different proportions in the portfolio. Degrees of freedom are fixed at 5 for all factors. The first (second) coefficient refers to the percentage proportion of Asset Class 1 (2) in the portfolio. Yearly risk-neutral probabilities have been fixed to 0.01 for each obligor of Asset Class 1, and 0.05 for those of Asset Class 2.

and CVaR, as expected. Secondly, an increase in the percentage of the risky asset (asset class 2) produces an increase in the losses of the portfolio.

4.4.2 Sensitivity to correlation

To analyze the sensitivity to correlation of the model (4.4.2.7) we have created a CDO based on a portfolio of 50 names. Individual default probabilities have been fixed at 1% for asset class 1 and 5% for asset class 2, respectively. To reduce the computational cost of the implementation, we have set the distribution of the two systematic factors as t-Student ones with 15 degrees of freedom. We have used the results of Shaw (2006) developed in Section 2.5, without loss of generality. All simulations have been performed using a double Hermite quadrature with 64 nodes. Idiosyncratic factors have been fixed at 5 for each asset class. To search differences in the portfolio composition, we have applied our study to two different sized portfolio: equally weighted portfolio (50% asset class 1, 50% asset class 2) and risky portfolio (25% asset class 1, 75% asset class 2).

Figures 5 and 6 display the spreads obtained under different sets of correlations for the equally weighted and risky portfolios. In general, the convexity pattern of the equity-senior curves remains constant in both cases, which is consistent with the preference (aversion) for risk on equity (senior)

tranche investors, as expected.

Regarding changes in correlation, Figure 5 reveals that correlation with risky asset classes are, by large, responsible of changes in the value of equity spreads. When it comes to the risky portfolio (Figure 6), it is interesting to observe how changes produced by the correlation in asset class 1 or 2 (see Equity and Mezzanine tranches) produce almost the same effects.

As also expected, an increase in global correlation raises the spreads of senior tranche. A higher correlation increases the probability of big losses, which is reflected in the senior spreads. This feature could be also mentioned (in a different scale) to the case of the risky portfolio in Figure 6.

4.5 Econometric Framework

This section focuses on the parameters estimation of the model (4.4.2.7). As pointed out in Embretchs, Frey and McNeil (2005), the statistical estimation of parameters in many industrial models are *simply assigned by means of economic arguments or proxies variables.* We will develop an exercise of formal estimation using some well known econometric tools as logit-probit regressions.[15] Due to the features of our data, some cautions must be taken to understand our results. This is due to the shortage of relevant data (for instance, rates of default of high-rated companies) or the sample size, as was also noticed by Embretchs, Frey and McNeil (2005). These authors also provide a more general discussion on the statistical estimation of portfolio credit risk models.

4.5.1 Estimation Techniques

As suggested by Schonbucher (2003) or Embretchs, Frey and McNeil (2005), the estimation of parameters in the expression (4.4.2.7) will be carried on using the models for discrete choice of proportions data. Basically, the idea consists in explaining the sample rates of default p_i (where i refers to the asset class or group) as an approximation to the population rates of default P_i plus an error term, ε_i. The idea behind is to link the population probability with some function $F(\cdot)$ over a set of explanatory factors $\mathbf{x}_i$ and their coefficients β, as follows:

$$p_i = P_i + \varepsilon_i = F\left(\mathbf{x}_i'\beta\right) + \varepsilon_i \tag{4.4.5.1}$$

[15] Standard references on this type of regressions using grouped data can be found in Novales (1993) or Greene (2003).

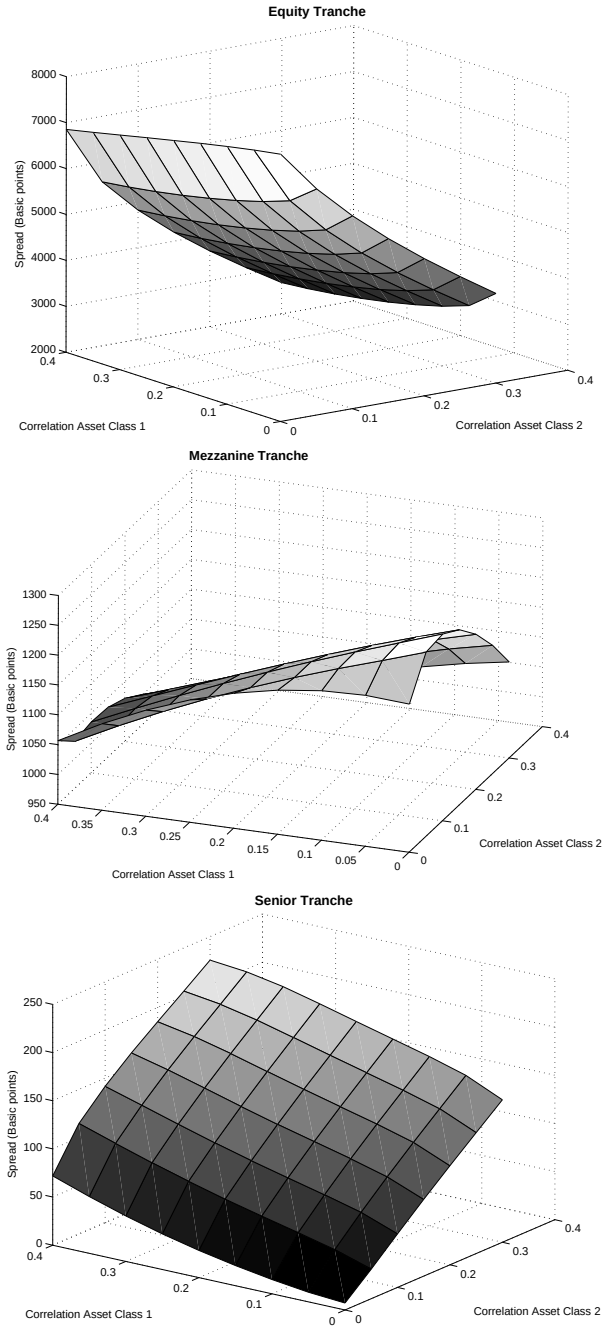


Figure 4.5: Spreads under different correlations for a equally weighted (50%-50%) portfolio.

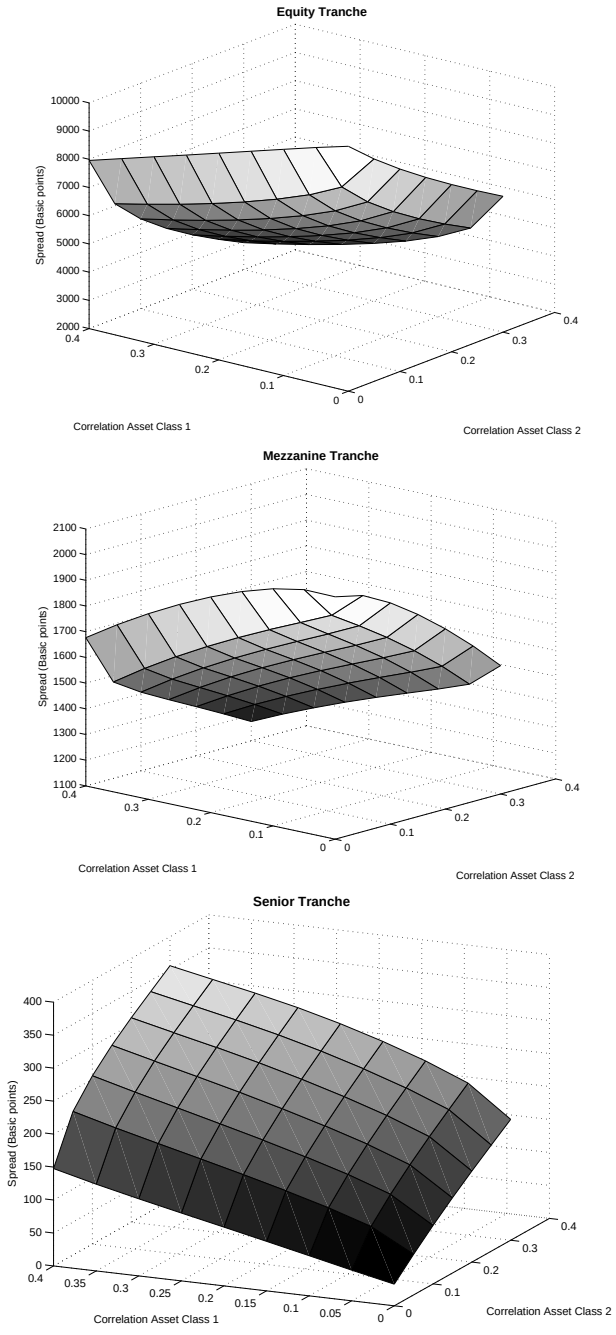


Figure 4.6: Spreads under different correlations for a 25%-75% portfolio.

To be interpreted as a probability, the function $F(\cdot)$ must be bounded and monotonically increasing in the interval $[0,1]$. Some widely used functions for F are the standard Normal distribution, which corresponds to the probit model, or the uniform distribution, which results in the linear probability model.

As suggested by Greene (2003), we could use regression methods as well as maximum likelihood procedures to estimate the set of coefficients β of the expression (4.4.5.1). For example, in the case of the probit regression, the relationship between the sample rates of default p_i and their population counterparts are

$$p_i = P_i + \varepsilon_i \rightarrow \Phi^{-1}(p_i) = \Phi^{-1}(P_i + \varepsilon_i)$$

which could be aproximated by (Novales, 1993)

$$\Phi^{-1}(p_i) \simeq \mathbf{x}_i'\beta + \frac{\varepsilon_i}{f(\mathbf{x}_i'\beta)}$$

where $\Phi(\cdot)$ denotes the distribution function of a standard Normal variable. As mentioned in Novales (1993), the last expression suggests that we can estimate the parameter vector β by regressing the sample probits $\Phi^{-1}(p_i)$ on the variables x. Considerations about the heteroskedasticity of the residual can be found in the cited reference.

To check the model's goodness of fit, Novales (1993) also provides a comparison of different regressions (probit, logit or lineal) in terms of the mean square error (MSE) . The statistic s is defined as

$$s = \sum_{1}^{T} \frac{n_i \left(p_i - \widehat{P}_i\right)^2}{\widehat{P}_i \left(1 - \widehat{P}_i\right)} \sim \chi^2_{T-k} \qquad (4.4.5.2)$$

where n_i represents the sample size of the data (subscript i refers to asset class or group) and $p_i, \widehat{P}_i$ are the observed and estimated frequencies, respectively. The statistic s follows a chi-square distribution with $T-k$ degrees of freedom, sample length T and k restrictions.

4.5.2 Variables and estimation

With the intention of illustrating the estimation of the model (4.4.2.7), we have chosen a set of six explanatory variables for the rates of default: the real Growth Domestic Product (GDP), the Consumers Price Index (CPI), the annual return on the S&P500 index (SP_ret), its annualized standard

Table 4.5: Descriptive Statistics

Statistics	Mean	Median	Std	Skew	Kurtosis	Max	Min
IG	0.00	0.00	0.00	2.25	8.21	0.01	0.00
SG	0.04	0.04	0.03	1.02	3.10	0.11	0.00
SP_ret	0.03	0.04	0.07	-0.61	2.51	0.13	-0.12
SP_std	89.47	39.40	106.34	1.42	3.72	390.19	7.55
GDP	0.01	0.02	0.01	-0.61	2.88	0.03	-0.01
CPI	0.02	0.02	0.01	1.23	3.78	0.06	0.01
10_rate	7.76	7.42	2.42	0.84	3.23	13.92	4.01
IPI	0.01	0.01	0.02	-0.46	3.29	13.92	-0.04

Descriptive statistics for the set of explanatory variables. Table exhibits the mean, median, standard deviation, skewness, kurtosis, maximum and minimum of the variables under study.

deviation (SP_std), the 10-year Treasury Constant Maturity Rate (10_rate) and the Industrial Production Index (IPI).[16] As dependent variables we have the annual rates of default for two investment grades: non investment grade (SG) and investment grade (IG), both collected from Hamilton *et al* (2005) . Due to the availability of default rate data, the sample period has been taken from 1970 to 2004, which results in 35 observations. A summary of the main statistics and the correlation coefficients is presented in Tables 5 and 6, respectively.

To visualize the influence of the proposed explanatory variables in the default rates, Figure 7 represents the scatter plots of non-investment grade rates of default versus different explanatory variables. This figure seems to corroborate what we could guess departing from the correlation parameters included in Table 6: the standard deviation of the S&P 500 returns, the GDP and the CPI can be good candidates for explaining the default rate in the case of the non-investment firms. Additionally, at a certain degree, the S&P 500 return can be added to this list as a possible explanatory variable in the case of the investment firms.[17]

[16] All data are available from the Federal Reserve Bank of St. Louis webpage (www.stlouisfed.org) except the S&P 500 index level, which has been taken from Bloomberg.

[17] A similar figure, available upon request, with investment grade rates of default was

Table 4.6: Correlation matrix

	IG	SG	SP_ret	SP_std	GDP	CPI	10_rate	IPI
IG	1.00							
SG	0.41	1.00						
SP_ret	-0.24	-0.22	1.00					
SP_std	0.18	0.35	-0.04	1.00				
GDP	-0.1431	-0.35	0.50	0.00	1.00			
CPI	-0.1847	-0.32	-0.28	-0.50	-0.49	1.00		
10_rate	-0.10	-0.23	0.06	-0.61	-0.11	0.61	1.00	
IPI	-0.15	-0.29	0.27	-0.04	0.67	-0.33	-0.13	1.00

One step ahead is to compute how much of the sample can be explained by the set of variables under study. To answer this question, we regress the non-investment and investment rates on these explanatory variables. Table 7 shows the results. The first row corresponds to the different independent variables under study. The first column contains the model under study - linear, probit, logit - and the different regressed variables (SG and IG default rates). Second to eighth columns display different betas obtained under different models. Finally, the last column shows the s statistic defined in (4.4.5.2), which will be used as a naive benchmark: if the whole set of independent variables explains some quantity of the sample, two variables would explain *"less"*: the pair of variables whose s value are closest to the benchmark could be good candidates as common factors in the model (4.4.2.7).

We start with regressions on SG rates. One main reason recommends this procedure: their data are more relevant to determine which factors may cause default. Up to a point, conclusions on the factors will be more robust. Previous regressions suggest choosing the variables GDP, CPI, IPI and SP_stdas common factors in the model (4.4.2.7). These variables minimize the statistic (4.4.5.2) with respect to other pairs of alternatives. Finally, we select GDP and CPI as common factors according to two main reasons:

built but is omitted for the sake of brevity. However, this figure reveals that, due to the high number of null observations in the IG sample, conclusions about the factors affecting IG rates should be taken carefully.

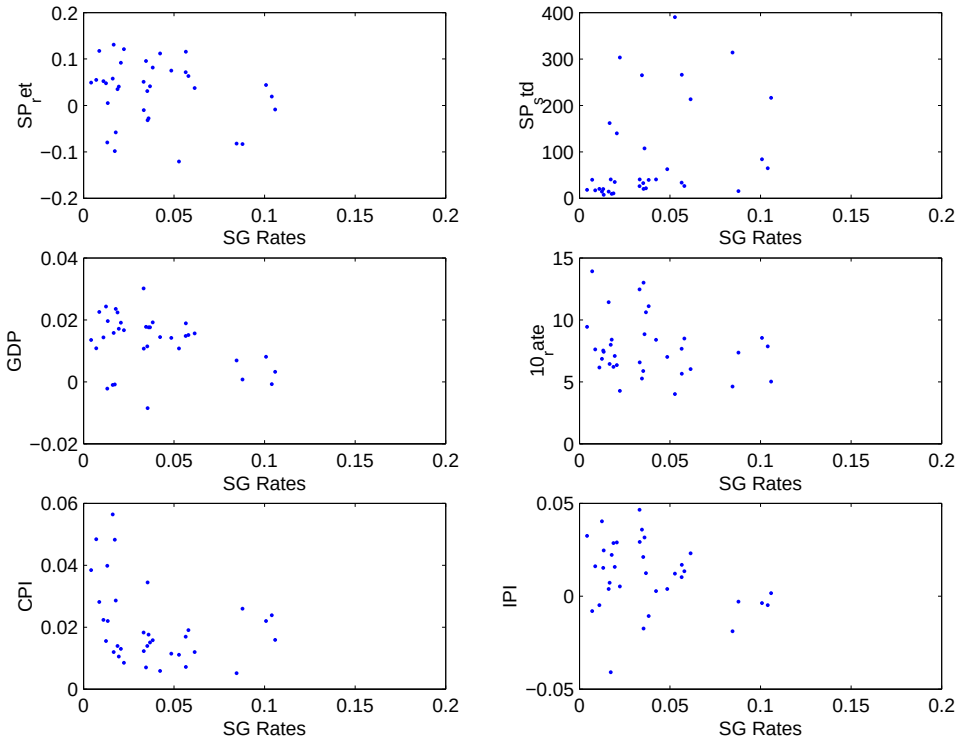


Figure 4.7: Non-investment grade rates of default (SG) versus different explanatory variables

Table 4.7: Regression of SG and IG default rates with respect to different explanatory variables.

Model	β_0	SP_ret	SP_std	GDP	10_rate	CPI	IPI	s
linear								
SG	0.08 [0.03 0.12]	−0.05 [−0.19 0.09]	0.00 [−0.00 0.00]	−1.96 [−3.44 −0.47]	0.00 [−0.00 0.01]	−1.69 [−2.67 −0.71]	−0.12 [−0.71 0.46]	11.55
IG	0.00 [−0.00 0.00]	−0.01 [−0.01 0.00]	0.00 [−0.00 0.00]	−0.018 [−0.09 0.05]	0.00 [−0.00 0.00]	−0.04 [−0.09 0.01]	−0.01 [−0.03 0.02]	1.78
probit								
SG	−1.38 [−1.86 −0.90]	−0.88 [−2.38 0.62]	0.00 [−0.00 0.00]	−23.74 [−39.74 −7.73]	0.05 [−0.00 0.10]	−26.80 [−37.34 −16.25]	−1.37 [−7.65 4.92]	13.94
IG	−3.40 [−4.32 −2.47]	−1.21 [−4.08 1.66]	0.00 [−0.00 0.00]	−17.15 [−47.79 13.49]	0.05 [−0.05 0.15]	−16.66 [−36.85 13.53]	3.63 [−8.40 15.66]	6.05
logit								
SG	−2.34 [−3.46 −1.23]	−2.16 [−5.62 1.30]	0.00 [−0.00 0.00]	−54.85 [−91.78 −17.91]	0.12 [−0.00 0.24]	−64.76 [−89.09 −40.42]	−3.25 [−17.75 11.25]	15.04
IG	−8.12 [−11.49 −4.74]	−4.20 [−14.67 6.27]	0.00 [−0.01 0.01]	−64.24 [−176.03 47.55]	0.18 [−0.18 0.54]	−60.29 [−133.94 13.36]	14.57 [−29.33 58.46]	6.79

Regression of Non-investment (SG) and Investment (IG) default rates with respect to different explanatory variables. SP_ret and SP_std are the yearly returns and standard deviation of the S&P500 Index. GDP, 10_rate, CPI and IPI are the Growth Domestic Product, ten year constant maturity rate, Consumer Price Index and Industrial Production Index, respectively. Finally, s is the fit statistic defined in (4.4.5.2).

1. Firstly, the Industrial Production Index could be seen as a proxy of the GDP and its information could result redundant. Moreover, regressions of probit-logit models using these two variables support the choice of GDP against the IPI.

2. Secondly, regressions on the parameter SP_std give a beta close to the precision imposed to our estimated parameters (10^{-4}).

Table 8 presents in columns the OLS[18] estimates for betas of independent term, GDP variable and CPI variable, respectively, using the SG rates. Confidence intervals at the 95% level are displayed into brackets. The rows in this table also display regression results for linear, probit and logit models. The last row contains the value of the statistic (4.4.5.2) obtained for each case. Attending to the goodness-of-fit criteria using s, the OLS probit model regression provides the best fit to the sample[19]. Overall, all the beta estimates corresponding to OLS regressions are negative, except for the independent term of the linear model, which leads to higher s statistic. Results concerning to OLS regressions can be interpreted as follows: a negative beta implies an increasing on default probabilities. In line with this, as expected, a decrease in GDP rates may produce an increase on SG default rates. Surprisingly, an increase in the CPI rate could diminish the rates of default, which might result counter-intuitive.

Table 9 shows regression results for linear, probit and logit models using IG rates as dependent variable. It is worth to notice that results are not conclusive as 63% of the sample under study are zeros. GLS estimations do not make sense in this context. In order to avoid numerical problems in the estimation, we have added the quantity 0.00005 to the sample, as suggested by Greene (2003). The first row displays the independent term and explanatory variables. Each pair of the following rows contains firstly the different values of betas obtained using two variables (GDP and CPI); secondly, their values using *only* the GDP variable. Maximum likelihood estimates (available upon request) for the probit and logit models are close to those parameters obtained for respective OLS models. We have estimated GDP variable alone with the intention of analyzing the explanatory power of the GDP on IG rates. First to second rows show the model and procedure used. The last column displays the value for the statistic s. At a certain degree, results on Table 8 could support the inverse relationship between

[18] GLS estimates have not been computed due to the sample size.

[19] Maximum Likelihood estimates (available upon request) for models probit and logit are close to those parameters obtained for respective OLS models.

Table 4.8: OLS regressions for SG default rates

SG	Model	β_0	GDP	CPI	s
linear	OLS	0.10 [0.07 0.12]	−2.24 [−3.25 −1.23]	−1.47 [−2.16 −0.78]	21.80
probit	OLS	−1.08 [−1.36 −0.80]	−26.77 [−38.18 −15.36]	−21.63 [−29.37 −13.89]	14.77
logit	OLS	−1.63 [−2.28 −0.98]	−61.90 [−88.46 −35.34]	−51.78 [−69.80 −33.76]	15.40

Results for regressions of Non-investment (SG) default rates on real Growth Domestic Product (GDP) and Consumer Price Index (CPI) as explanatory variables. s correspond to the fit statistic defined on expression (4.4.5.2).

the explanatory variables and the IG rates of default, as previously noted for the SG case.

4.5.3 Interpretation of coefficients

As pointed out by Elizalde (2005), due to the difficulty of interpreting what the correlation term represents, estimating the correlation term in factor models is not an evident task. Looking at equation (4.4.5.1), the estimate β describes the effect from the explanatory factor x through a non-linear transformation of the firm's asset value, which itself is unobserved, as it is also noticed by Elizalde (2005). As this fact complicates understanding the proper correlation term, the author enumerates some measures used ad hoc by practitioners, as equity return correlations, to conclude about the insufficiency (and scarcity) of papers that deals with this problem.

Our interpretation of coefficients in the model (4.4.2.7) goes in the direction of the econometric explanation for the coefficients of the linear, logit and probit models, that is, the influence of the exogenous variables on the endogenous one. In other words, the (relative) impact of the explanatory variables on the probability of default. Following Novales (1993), this interpretation of estimates for the linear model must differ to that for the logit and probit models.[20] This is the main reason why we split our results in

[20]For example, the relationship between the explanatory and explained variables in the probit model is non-linear while the linear probability model implies linearity between dependent and independent variables.

Table 4.9: OLS regressions of IG default rates.

IG	Model	β_0	GDP	CPI	s
linear	OLS	0.00 [0.00 0.00]	−0.04 [−0.09 0.01]	−0.03 [−0.06 0.01]	5.09
	OLS	0.00 [0.00 0.00]	−0.02 [−0.06 0.03]	–	2.26
probit	OLS	−3.14 [−3.65 −2.64]	−14.16 [−34.97 6.65]	−11.24 [−25.36 2.89]	6.44
	OLS	−3.48 [−3.77 −3.19]	−5.99 [−24.52 12.54]	–	8.63
logit	OLS	−7.18 [−9.03 −5.33]	−50.80 [−126.73 25.13]	−40.74 [−92.26 10.79]	7.17
	OLS	−8.40 [−9.45 −7.35]	−21.16 [−88.75 46.43]	–	9.64

Results for regressions of Investment (IG) default rates on real Growth Domestic Product (GDP) and Consumer Price Index (CPI) as explanatory variables. s correspond to the fit statistic defined on expression (4.4.5.2).

two tables, Tables 10 and 11, that include - respectively - the estimation of the linear and logit-probit models.

Regarding the estimates of the linear probability model, Table 10 reflects the contribution of the two explanatory random variables to the probability of default. The main conclusions are obtained from the default probabilities of non-investment grade assets (SG), but can also be extended to the investment grade (IG) ones. Looking at Table 10, it is interesting to observe the sign of the coefficients, which is negative: the more we decrease the GDP or the CPI, the more we increase the rates of default. Given the value of the coefficients, the variables have the same contribution to the default probability. With reference to the fit of the model to the data, under the null hypothesis that the goodness-of-fit to the sample is good, we cannot reject that the linear probability model could explain the results obtained.

Table 11 includes the ratio between estimates for SG and IG series for probit and logit models, respectively. By and large, the conclusions are the same for all the series under study. According to Novales (1993), the ratio between the estimated betas measures the relative contribution of the explanatory variables on the default probability. Results are consistent to those obtained for the linear probability model: the negative sign of the ex-

Table 4.10: Results for the linear probability model.

Linear model	β_0	GDP	CPI	s	$\chi^2_{95\%}(32)$	$\chi^2_{99\%}(32)$
SG	0.10	-2.24	-1.47	21.80	No reject	No reject
IG	0.00	-0.04	-0.03	5.09	No reject	No reject

Estimated coefficients for the linear probability model. s correspond to the fit statistic defined on expression (4.4.5.2).

Table 4.11: Estimated coefficients for probit and logit models.

	β_0	$\frac{GDP}{CPI}$	s	$\chi^2_{95\%}(32)$	$\chi^2_{99\%}(32)$
Probit model					
SG	-1.0785	1.2377	14.7653	No reject	No reject
IG	-3.1421	1.2606	6.4347	No reject	No reject
Logit model					
SG	-1.6293	1.1954	15.4018	No reject	No reject
IG	-7.1801	1.2470	7.1658	No reject	No reject

Estimated coefficients for probit and logit models. s correspond to the fit statistic defined on expression (4.4.5.2).

planatory variables, which reflects an opposite effect between default ratios and the macroeconomic variables. Moreover, the relative contribution of the explanatory variables remains equal, as was also derived from Table 10. Finally, we do not reject the goodness-of-fit of the model using confidence levels of 95% and 99%.

4.6 Conclusions

The current success of the credit derivatives market for tranched products is one of the biggest ones seen within the financial industry. The standard pricing model, widely used by the practitioners, is the Gaussian one-factor model (Vasicek, 1991). However, some assumptions underlying this model are probably too restrictive. These features concern, among others, to those of homogeneity of asset classes involved, or the exposure to one sources of

systematic risk.

In a more realistic setting, Schonbucher (2003) or Lando (2004) pointed out that a credit portfolio is composed by different asset classes or buckets, attending to criteria as, for example, investment grade, non-investment grade assets or industry. The exposure of a credit portfolio to a set of common risk factors could be significant between groups, but should be homogeneous within them. In line with this, the idea of two groups of assets treated in different ways could become a more realistic assumption than that used previously in the literature.

With the aim of contributing to the current literature, this article has considered a family of models that takes into account the existence of different asset classes or regions of correlation. In more detail, this paper has generalized the two assets-two Gaussian factor model of Schonbucher (2003) by proposing a two by two model (two factors and two asset classes). We assume two driving factors (business cycle and industry) with independent t-Student distributions, respectively, and allow the model to distinguish between portfolio assets classes. It may be worth noting that one of the main implications from considering the t-Student distribution is that we assign a higher probability to high default events.

Our work contributes to the existing literature in the analysis of these asset class models. To the best of our knowledge, no similar study has been reported yet in this direction. Regarding the distributional assumptions, we have extended the standard Gaussian model by considering the t-Student distribution. In this way, we have dealt with a more general model with the additional advantage that includes the Gaussian model as a particular case. We have also provided the econometric framework for assessing the parameters of the posited model. Finally, an empirical application with Moody's data has been presented as an illustration of the methodology proposed.

After proposing several explanatory variables, it seems that the standard deviation of the S&P 500 returns, the GDP and the CPI can be good candidates for explaining the default rates in non-investment firms. Additionally, at a certain degree, the S&P 500 return can be added to the previous ones to explain the default rates for investment firms. Additionally, a more detailed study leads to select GDP and CPI as common factors.

Focusing on OLS estimates, it is seen that - as expected - a decrease in GDP rates may produce an increase on SG default rates. Surprisingly, it is also obtained that an increase in the CPI rate can diminish the default rates Regarding the estimates of the linear probability model, we see that the more we decrease the GDP or the CPI, the more we increase the rates of default.

Finally, the ratio between estimates for SG and IG series for probit and logit models has been analyzed. The conclusions are the same for all the series under study and the results are consistent to those for the linear probability model: there is a negative relationship between default ratios and the macroeconomic variables. Deeper explanation for these empirical findings will be the subject of further research.

Chapter 5

Numerical Pricing of Collateral Debt Obligations: A Monte Carlo Approach

This chapter[1] addresses in a simple and illustrative way one of the different alternatives that can be used to price and manage this type of derivatives. In more detail, we explore the CDO pricing model proposed recently in Longstaff and Rajan (2007). Then, this chapter analyzes this model considering different multi-factor versions and, additionally, some extensions are proposed considering jumps in the default process or a random loss distribution.

Our simulation analysis shows evidence that a three-factor version of this model with constant losses impact is flexible enough to reproduce the spreads given by the market. In addition to this, the inclusion of jumps to the default process results in an high arrival of credit events. Finally, it is shown that the alternative of random losses can be helpful when dealing with one- and two-factor models.

[1]I would like to thank to J.I. Peña for his helpful comments and suggestions. I also acknowledges financial support from the Plan Nacional de I+D+I (project BEC2003-02084) and the Gobierno Vasco-UPV project (GIU 06/53), especially to Jose M. Usategui. Previous drafts of this chapter have been accepted to be presented on XV Foro Finanzas AEFIN, 15th-16th November 2007, Palma de Mallorca, Spain, and XXXII Simposio de Análisis Económico, 13rd-15th December 2007, Granada, Spain. I have also benefit from the comments of participants in a seminar at Universidad Complutense, Madrid, Spain.

5.1 Introduction

The pricing of derivatives based on portfolios composed by assets with default risk has received an increasing attention during the last years. Up to a point, this may be due to the recent appearance of the highly successful Credit Default Swap (CDS) market, where the irruption of indexes like the DJ iTraxx or DJ CDX, and their tradable tranches series have contributed to their fast growing over the last years. For example, the International Swaps and Derivatives Association (ISDA) reports an increase in the notional principal outstanding volume[2] of CDS from $26.0 trillion at June 30, 2006 to $34.5 trillion at December 31, 2006, which represents a 33% increase during the second half of 2006.[3] In view of this fast development of standardized credit markets, the recent financial literature has focused their interests on the pricing of these new credit linked products, with attention to the modelling of some interesting features presented in these assets.

Probably, the main characteristic in basked credit derivatives products is their exposure to default correlation, intended as the different relationships of dependence between defaults that could be given among firms: the higher the correlation between the assets in a credit portfolio, the higher the probability of suffering a big loss in the portfolio. Then, the role of the correlation is crucial to price assets as, for instance, First-to-Default Swaps (FtD) or Collateral Debt Obligations (CDO).

The recent literature on credit risk refers, at least, four main approaches to achieve correlation among different obligors. As pointed out by textbooks of Schönbucher (2003), Lando (2004) or Duffie and Singleton (2003), the academic literature provides different frameworks to introduce correlation. Those more cited in the literature are the following:

- Structural models
- Intensity models
- Copulas
- Mixture models

[2]ISDA news release, April 18, 2007. Available at http://www.isda.org. This report refers to credit default swaps on single-names, baskets and portfolios of credits, baskets and index trades.

[3]CDS notional volume growth rates were 52 percent during the first half of 2006. Those reported for the whole years 2006 and 2005 were 102 percent, and 103 percent, respectively. Source: www.isda.org.

Roughly speaking, the main difference between these different pricing models for CDOs consists on the form that they use to achieve correlation among different obligors. We will focus here in the two main approaches to introduce default dependence between assets, in line with the two main paradigms for analyzing credit risk:[4]

1. On one hand, the structural based models, which take into account the capital structure of a firm by relating the value of the issuing company to its liabilities (Schmidt and Stute, 2004). This family of models includes those of Merton (1974), Black and Cox (1976) or Leland and Toft (1996), among others.

2. On the other hand, the reduced form approach is based on the relative valuation of credit derivatives by means of market data, with no care about any firm features. Jarrow and Turnbull (1995) or Duffie and Singleton (1999), among others, can be included within this second group.

To put it briefly, the main basic difference between structural and reduced form models consists on *how* they consider the causes of default:

1. The structural models explain the event of default as the impossibility of the firm to respond to their liabilities, that is, in an endogeneous form.

2. By contrast, the intensity based model analyze the dynamics of default from an "exogenous" point of view, extracting directly the probabilities of default from the observation of market prices (Giesecke, 2004).[5]

Due to its direct connection with the literature of bond pricing, the intensity approach has became an standard in the valuation of corporate bonds and their credit derivatives counterparts (CDS, for example). Although, the intensity based literature has not been capable to extend its influence to the pricing of credit basket derivatives, and specially to tranched products like the Collateral Debt Obligations (CDOs). Some remarkable contributions are those of Duffie and Garlenau (2001), which basically considers correlations between the default times of individual firms. Nevertheless, as

[4] See Schönbucher (2003) or Duffie and Singleton (2003) for an overview of additional alternatives for modelling default dependencies.

[5] Recently, Duffie and Lando (2001) has demonstrated that both approach are connected by information asymmetries.

pointed out by Hull and White (2004) or Mortensen (2006), this procedure is computationally burdensome.[6]

In addition to the lack of tractability of intensity models, another plausible reason that may explain the wide-range usage of factors models by the practitioners may be given by the popularity of some models like the Credit + of Moody's or the KMV (see Schönbucher (2003) or Schmidt and Stute (2004) for a digression of these models). Directly inspired in Merton (1974), both models formalize the CDO pricing under an *individual* level firm framework by modelling the default of individual firms that trigger some threshold previously determined, and combining their different possibilities. Furthermore, the development of techniques that improve the computation times of former models (usually highly time consuming) has contributed to extend the scope of these models.[7]

In contrast to this *individual* approach, where the total number of defaulted firms is conditioned by the level of correlation among the obligors, Longstaff and Rajan (2007) posited a joint default approach for the pricing of CDOs: basically, they assume that defaults could be clustered, in the sense that joint defaults could happen *simultaneously*, an approach closely linked to the contagious scheme of Davis and Lo (1999). Contrary to Duffie and Garlenau (2001) approach, where defaults are considered at an individual level, the assumption of joint default simplifies dramatically the computation of CDOs prices, as will be shown later.

The aim of this chapter is to study the pricing of CDO tranches using the intensity based scheme of clustered defaults developed by Longstaff and Rajan (2006). Our intention is to explore the capability of the model to generate reasonable values for CDOs tranches under different sets of circumtances, particularly those referred to the number of factors used and their impact in the loss distribution. We also explore some natural extensions of the original model, extending its range mainly in two directions: firstly, by analyzing the consequences of adding jumps to the default process; then, by randomizing the impact of defaults in the loss distribution. To the best of our knowledge, no similar studies has yet been reported.

This chapter is organized as follows. Section 2 reviews the mechanics of the CDS and their related indexed products like Itraxx or CDX. Section 3 describes the model posited in Longstaff and Rajan (2007) and introduces

[6]Recently, Mortensen (2006) has implemented a tractable extension of the Duffie and Garlenau (2001) framework by considering the algorithm of Andersen, Sidenius and Basu (2003) for factor models.

[7]We can mention the algorithm of Andersen, Sidenius and Basu (2003) as an example of this type of methods.

some modifications in that paper. Section 4 performs a Monte Carlo study to illustrate numerically some extensions we propose in the original model by Longstaff and Rajan (2007). Finally, the main conclusions are summarized in Section 5.

5.2 CDS indexes

This section is devoted to the CDS indexes. We outline here their main features, beginning with their most basic component, the Credit Default Swap; then, we review their pricing mechanics, and provide some standard pricing formulas. This part is mainly based on Felsenheimer, Gisdakis and Zaiser (2004), Jakola (2006) and references therein.

5.2.1 CDS

Prior to show the mechanism of a CDX index, it is necessary to describe briefly the main component of this index, the Credit Default Swap (CDS). A CDS is a financial product used to hedge mainly fixed income assets (for instance, bonds) against certain types of credit events.[8] Then, a CDS is based on an agreement between two parties where one of them - the protection buyer - pays a defined, periodical amount to the other - the protection seller - contingent to the occurrence of a given credit event (usually a default). If there is no default, the protection buyer must pay a defined premium until maturity. By contrast, if default happens, the protection seller must pay to the protection buyer an amount equal to the difference between the face value of the asset (for instance, a bond) and its market value after default. As pointed out by Duffie and Singleton (2003), a CDS can be seen as a default insurance on loans or bonds.[9]

Up to a point, the success of the CDS market could be explained as these assets provide to investors with a simple, liquid tool to remove the credit risk exposure of their portfolios in an easy way. Blanco, Brennan and Marsh (2006) underlines that the cost of shorting credit risk in the corporate bond market is high, mainly due to illiquidity costs. By using CDS, investors can short credit risk over a longer time period at a known cost by buying protection.

[8] For instance, Moody's refers bankruptcy, delayed disbursement of principal or some legal distresses, among others.

[9] As noted by Blanco, Brennan and Marsh (2006), the difference between a CDS and an insurance contract is of legal type: under a CDS, holding an insured asset to claim compensation is not necessary.

5.2.2 CDS Indexes

The product

A direct consequence of the fast growing of the credit market has been the appearance of standardized products that diminish the hedging cost of portfolios composed by a large number of firms, by means of transferring their credit risk without necessity of buying CDS contracts on each individual firm. An example of these kind of products is the Credit Default Swap Index.

A CDS index is a basket of equally weighted CDS single names that serves as underlying for some tranched products, swaps, etc. Felsenheimer, Gisdakis and Zaiser (2004) refers that a CDS index can not be characterized as a classical index, in the sense that no index level is computed. They also noticed that quoted prices for these products are a result of supply and demand within credit markets.

Jakola (2006) refers two main tradable index families: the Dow Jones CDX and the International Index Company Itraxx. Basically, both indices are portfolios that present similar features as, for instance, the following:[10]

- Both indices are composed by the 125 most liquid firms in the CDS market.
- These firms are investment grade rated.
- The index composition is revised every six months.

With respect to differences between these indices, we can mention that the Dow Jones CDX only include US-names while the Itraxx only takes into account European names.

The Mechanism

To explain the payment mechanism in a CDS index, we illustrate briefly here its two main situations, depending whether there is default or not. We address the reader to Felsenheimer *et al* (2004) for a more detailed exposition.

Investors on CDS index can find any of the following alternative situations:

[10] See Jakola (2006) for references to the Dow Jones CDX and Felsenheimer, Gisdakis and Zaiser (2004) for the Itraxx Index.

- **No credit event happens**. Here, the protection buyer pays to the protection seller a constant premium. This premium is computed as a percentage (named the spread) on a definite notional amount and is paid on a quarterly basis until maturity.

- **A credit event happens**. Now, the magnitude of the event and the recovery value determines the cash payment between the parties involved. As an example, we will assume that 1.6% of the firms included in the total portfolio default (2 over 125) and a recovery rate of 50%. Here, the protection seller pays the amount of 1.6% $\times$ notional $\times$ 50% to the protection buyer. As could be intended, the protection seller is exposed to recovery risk. With respect to the protection buyer, he continues paying the constant premium on the remaining notional amount, which is now 98.4% of the original. This quantity should be paid until maturity or until other credit event.

The existence of a tranched products market besides each CDS index has provided a liquid, simple way of removing the credit exposure associated to fixed, standardized fractions - tranches - of the underlying portfolio losses. In other words, without necessity of creating the full capital structure, an investor can hold individual index tranches according to their needs. Longstaff and Rajan (2007) argues that this huge flexibility has contributed to the fast increase in the trading volume of the index tranches.

5.2.3 Pricing formulas

To complete this overview of tranched products, this section outlines the basic formulas for pricing a CDO. For a more detailed exposition, we refer the reader to Gibson (2004) or Elizalde (2005) and references therein.

Roughly speaking, pricing a CDO means to compute the fair value of the percentage on a definite notional that must be paid for the CDO issuer - protection buyer - to their counterpart, the CDO investor or protection seller. As previously mentioned, this percentage is called the *spread*. With small differences with respect to a swap, a CDO is a contract between two parties where the CDO issuer pays periodically a fixed coupon - based on the spread - to her counterparty. This payoff scheme is close to that of a Credit Default Swap in the sense that the payments are contingent to the occurrence of a credit event on the underlying. However, its difference comes from the following facts:

1. We operate over a pool of firms in the case of a CDO, instead of a single name as in a CDS

2. We can define the loss amount of the underlying portfolio in percentage terms (tranches) that the protection seller agrees to offer protection, in contrast to a CDS, where the protection seller must face to the total amount of the (single) loss.

Usually, a CDO is issued over a portfolio of bonds, loans, CDS or other defaultable instruments (Elizalde, 2005). The exposure to the losses in the portfolio is sold in tranches, that is, percentages on the total portfolio losses. Their bounds are usually standardized and called attachment points. For instance, usual attachment points are 0%-3% for the equity tranche - riskier - and 15%-30% for the senior tranche - smaller risk.[11] Intermediate tranches are called mezzanine and their attachment points vary from 3%-7%, 7%-10%, and 10%-15%.

An additional feature of a CDOs is the hierarchy in the payoffs: cash-flows generated by the pool of firms are assigned by seniority to the different tranches of the CDO; first, the owners of senior tranches, and mezzanine and equity tranches subsequently. It is worth to say that equity investors receive the last payments of the credit portfolio. As a result, these investors are highly exposed to the first losses that may occur in the underlying portfolio.

As pointed out by Elizalde (2005), the pricing of a CDO is similar to that of a swap contract in the sense that the spread is fixed as there is no up-front payment between the parties. In line with this argument, we must compute the cash flows received for each part and equate - in present value terms - both amounts.

Let s be the spread of the CDO, and let U and L be the upper and lower attachment points that define the tranche, respectively. Let R also be the recovery rate, and let X_t be a random variable which defines the total distribution of losses in the portfolio at time t. For the ease of the exposition we will define later the function $f(\cdot)$ that defines the losses of the tranche holders over the total losses in the portfolio.[12]

On one hand, the cash flows received by the protection seller will be the (unknown) spread s over their associate tranche $(U - L)$ computed on a principal amount. If any firm of the portfolio defaults, the protection seller will receive the same spread *but* over the new (reduced) notional amount,

[11] Bespoke tranches are also available in Over-The-Counter markets.

[12] To be consistent with our notation, recovery rates, attachment points and the losses in the portfolio are given in percentage terms.

$U - L - E[f(X_t)]$.[13] The accrual factor for the payment t is given by Δ_t. Computing the corresponding present value for each date of payment, this results

$$P_{seller} = s \sum_{t=1}^{T} e^{-rt} \Delta_t \left(U - L - E[f(X_t)]\right) \tag{5.5.2.1}$$

On the other hand, the protection buyer expects to receive the discounted amount of the total losses produced on the tranche during the different periods, which are

$$P_{buyer} = \sum_{t=1}^{T} e^{-rj} \left(E[f(X_t)] - E[f(X_{t-1})]\right) \tag{5.5.2.2}$$

By equating expressions (5.5.2.1) and (5.5.2.2), we get the spread

$$s = \frac{\sum_{t=1}^{T} e^{-rj} \left(E[f(X_t)] - E[f(X_{t-1})]\right)}{\sum_{t=1}^{T} e^{-rt} \Delta_t \left(U - L - E[f(X_t)]\right)} \tag{5.5.2.3}$$

Finally, the exposure to the total portfolio losses faced by the tranche investors is given by

$$\begin{aligned} f(X_t) &= \min\left[(1-R)X_t, U\right] - \min\left[(1-R)X_t, L\right] \\ &= \max\left[\min\left[(1-R)X_t, U\right] - L, 0\right] \end{aligned} \tag{5.5.2.4}$$

where, of course, the recovery rate affects directly the total portfolio losses.

5.3 The model

This section is devoted to the model. First of all, we justify the idea of correlation among firms, and cite the main approaches presented in the literature to model it. Secondly, we analyze some interesting aspects of the Longstaff and Rajan (2007) model, incorporating also some extensions.

5.3.1 Previous models

The distribution of losses in basket products like CDS indexes or tranches plays a crucial role, because of the amount of cash flows delivered between the protection seller and the protection buyer depends dramatically on the number of default events occurred during the contract: the higher thee

[13] Of course, until the total losses of the portfolio overcome the tranche.

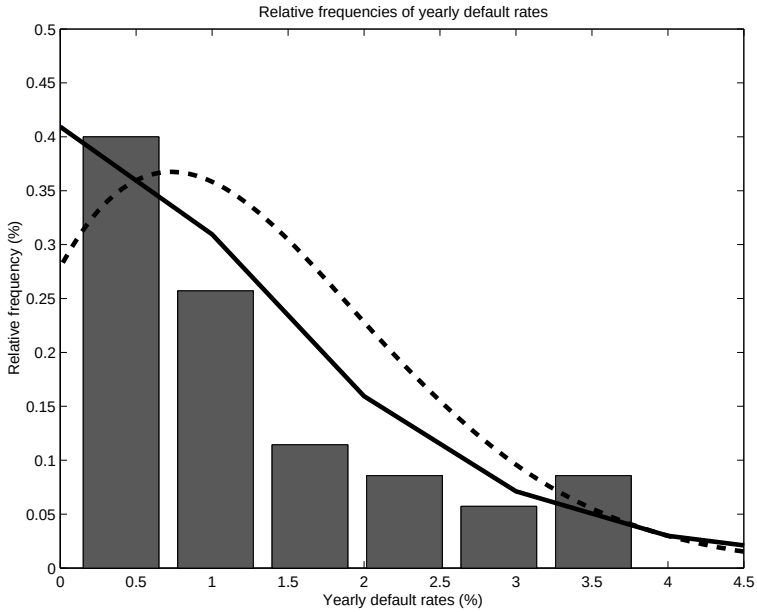


Figure 5.1: Relative frequencies of yearly default rates. The empirical yearly default rates are represented in bars. The dashed-line corresponds to simulated default rates for a portfolio of one hundred independent obligors. The solid line represents the default rates obtained with the Vasicek (1991) model for the same portfolio but considering a correlation parameter of 10.00%. Individual default probabilities are fixed at 1.20%. Simulation parameters have been previously estimated from real data.

number of firm defaults in the portfolio, the higher the losses. In some cases where the prioritization schemes are embedded in some products like the standard tranches, the losses on a definite percentage of the portfolio can be viewed as a call spread on the total losses of the portfolio.[14]

One of the stylized facts generally assumed in the credit risk literature is that firms usually default *together*, in the sense that a default in an individual firm seems to increase the risk of default in some others. A simple empirical exercise can illustrate this idea.

With the intention of checking this effect in real data, Figure 1 shows (with an histogram) the distribution of default rates for US firms from 1970 to 2004.[15]

Figure 1 also displays the default rate distribution obtained in a simu-

[14]See Duffie and Singleton (2003) or Longstaff and Rajan (2007).

[15]The data has been taking from Hamilton *et al* (2005).

lated portfolio (dashed line) composed by one hundred *independent* obligors. For the sake of simplicity, all firms within the portfolio have the same characteristics -even the same probability of default. This is usually called the *homogeneous portfolio* assumption. It is clear that, in absence of correlation among firms, the total distribution losses will be described by a binomial distribution with parameters $N = 100$ and p equal to the individual default probability of each obligors. To put it another way, it reduces to an experiment repeated (independently) N times with probability of success equal to p. Finally, this figure also exhibits the simulated rates of default for a portfolio taking into account correlation among firms (solid line).

Simulations of default rates have been carried on using the Vasicek (1991) model, an standard model used by the industry for pricing CDOs. Simulation parameters for yhe independent ($\rho = 0$) and dependent cases correspond to those previously obtained from the yearly default rates sample of Hamilton, Varma, Ou and Cantor (2005). Estimated parameters of correlation and individual default probability for the Vasicek (1991) model are respectively $\rho = 10.73\%$, $p = 1.27\%$. Finally, parameters have been estimated using the Large Homogeneous Portfolio hypothesis (LHP).[16]

As figure 1 reveals, a model that incorporates correlation among firms seems to offer a better fit to the sample data than a model without correlation. Additionally, Schönbucher (2003) also presents evidence of default clustering through time. The intuition behind this fact is that default correlation among entities results in a higher risk of joint default. That is to say, the correlation between firms affects crucially to the total number of defaults. As a result, models used for pricing portfolios of firms should take into account this empirical feature of data.

Our work is focused on the recent proposal of Longstaff and Rajan (2007), which is inserted within the intensity models family. The next section analyzes some of its main features.

5.3.2 A simultaneous default model

Longstaff and Rajan (2007) assumes the occurrence of simultaneous defaults within the obligors of the credit portfolio. The main idea behind this approach consists on modelling the loss distribution directly, by using a monotonically decreasing process. The intuition underlying this approach is that the arrival of a credit event (usually triggered by an underlying intensity process) produces a fractional loss in the portfolio. Longstaff and Rajan

[16] See Schönbucher (2003) for a detailed discussion of the former points.

(2006) posit that the impact of credit events on portfolio losses are constant through time, with a range that implies defaults from one to various tens of firms.

In view of the former, three main components could be noticed in the Longstaff and Rajan (2007) model: first of all, the loss process; secondly, the process that activates the losses; and finally, the magnitude of the losses themselves. What follows is a more detailed exposition of this point, where we also include some extensions to the original model.

The loss process

Using the notation given in Longstaff and Rajan (2007), let L_t denote the total portfolio losses per \$1 notional amount, with $L_0 = 0$. The dynamics of the loss distribution is given by the expression

$$\frac{dL_t}{1 - L_t} = \sum_{i=1}^{n} \overline{\gamma}_i dN_{it}, \tag{5.5.3.1}$$

where

- N_{it} is a Poisson process with time-varying intensity λ_t
- γ_i is the magnitude of the jump size
- n represents the number of factors that affects the losses.

As shown in Longstaff and Rajan (2007), the solution of equation (5.5.3.1) is given as

$$L_t = 1 - \exp\left(- \sum_{i=1}^{n} \gamma_i dN_{it} \right) \tag{5.5.3.2}$$

The mechanism of the model is simple: an underlying intensity process, which captures the instantaneous default probabilities given by the market, causes a jump in the Poisson process, which takes value 1. This activates a (percentage) loss in the portfolio of magnitude $e^{-\gamma}$.

Figure 2 illustrates the evolution through the time of the losses based on a extended version of Longstaff and Rajan (2007) model for different parameters.[17] We have simulated one thousand different scenarios of portfolio losses and have collected in different histograms their evolution. As Figure

[17] Particularly, we have considered jumps in the intensity process. Moreover, we have taken the impact of losses as a random variable.

2 exhibits, the number of bigger losses in the portfolio increases when times goes by, with a high number of zero counts in the early stages that decreases with time.

Based on Principal Component Analysis over the CDX index and its associated tranches, Longstaff and Rajan (2007) consider that a three factor model is enough to capture the whole variability of prices in the CDX index tranches. This point is also tested by Longstaff and Rajan (2007) analyzing simplest versions of the model. Their results suggest that the three factor model version presents a close fit to the data.

Intensity process

To capture the instantaneous default probability observed in the market, we assume a stochastic process for the intensity of the Poisson process, λ_t. By definition, the process must be guarantee positive values. Then, it is common to assume an squared-root type process as that proposed by Cox, Ingersoll and Roll (1985) for interest rates. We also consider in our study the potential appearance of jumps in the intensity process, as previously used by Duffie and Garlenau (2001) in the context of individual firms:

$$d\lambda_{it} = (\alpha_i - \beta_i \lambda_{it})\, dt + \sigma_i \sqrt{\lambda_{it}} dZ_{it} + J_i d\pi \qquad (5.5.3.3)$$

where

- β is the mean-reversion speed of the process
- α/β determines the long-run mean value at which the intensity converges
- σ is the volatility of the intensity process

Finally, we assume that jumps are driven by a Poisson process $d\pi$ with constant intensity δ_i. The size of the jump is given by a random variable J exponentially distributed with mean μ_i.

The objective that we pursue by adding jumps is to increase the volatility of the intensity process. The intuition, is that, in line with that posited by Merton (1976) for stock markets, the arrival of abnormal information - modelled by the Poisson process $d\pi$ - could produce an increasing, non-marginal change of magnitude J in the instantaneous default probability.

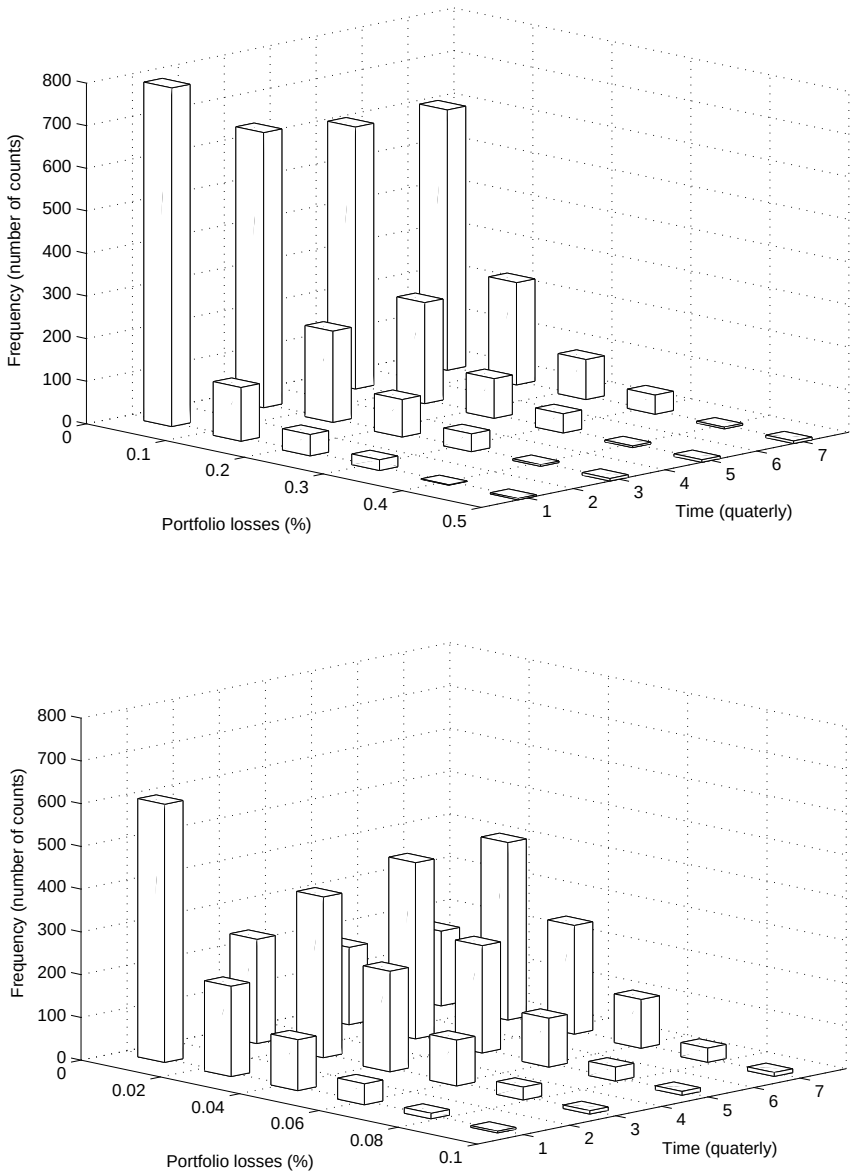


Figure 5.2: Evolution of portfolio losses in the Longstaff and Rajan (2007) framework for two different parameter sets. Notice that portfolio losses axes have different scales.

Impact of losses

The impact of a credit event in the total portfolio losses is given by the parameter γ_i in equation (5.5.3.1). The estimations given for a three factor model in Longstaff and Rajan (2007) suggest an economic interpretation of the different parameters γ in terms of percentage of losses in the portfolio: the market seems to be pricing three main credit events:

1. The default of a single obligor.
2. An industry-specific event that affects one sector of the economy (about 10 firms).
3. A wide crisis that affects the whole economy (about 75 firms).

As the authors pointed out, this interpretation is consistent with that mentioned by Duffie and Garlenau (2001) that considers default events in three categories: firm-specific events, industry events and recessions of the entire economy.

Again, we adopt here a different approach by randomizing the impact of losses after a credit event. Thus, we consider that losses are distributed as non-negative random variables. In more detail, we assume exponentially distributed variables with mean γ. Our pourpose is to relax the assumption of fixed impact, in order to allow for more flexibility to the total losses portfolio distribution.

5.4 Monte Carlo study

This section analyzes the different models posited in the previos section by means of simulations.[18] We start exploring the original model of Longstaff and Rajan (2007), and extend the study in order to consider two possible alternatives: in the first place by introducing jumps in the intensity process, as previously specified; secondly, by randomizing the losses or the jumps.

Before starting this simulation study, some general considerations must be taken into account. The simplest form of the model proposed by Longstaff and Rajan (2006) (one factor and constant loss size) needs a four-parameter vector $\Theta = (\alpha, \beta, \sigma, \gamma)$ including three parameters for the intensity process

[18] Our study comprises 5000 paths simulations with 1,250 steps each one. Initial parameters for the intensity process of expression (5.5.3.3) have been fixed to those of their long-run means. All computations have been carried on a 2.8 Mhz Pentium IV computer with 500 Mb of RAM.

and one for the loss impact. This means that any extension of the model implies adding more parameters; what is more, the different combinations of parameters that we must introduce to analyze the responses of the model increase the difficulties of interpretation.

On account of this, some restrictions on the study must be imposed, with the intention of extracting some reasonable conclusions about the parameters and their cross interrelationship. For this reason, we will proceed in a sequential form by detecting (in the early stages of the model) the set of parameters that could be of interest in subsequent extensions of the model.

As a general rule, our Monte Carlo study tries to identify the changes on the spread of tranches in two main directions:

1. Changes on the parameters of the intensity process

2. Changes on the losses size

In line with this idea, all the following tables mainly address variations on one of these alternative ways.

5.4.1 Longstaff and Rajan (2007) model

As a first step, we analyze the model presented in Longstaff and Rajan (2006) with a constant impact losses γ and intensity process (5.5.3.3) without jumps ($J = 0$). We start by using the parameters obtained therein due to the inexistence of similar studies, a point that will be discussed later when necessary.

We begin by simulating the one factor model. Table 1 displays the results we have obtained. The first column indicates the set of parameters for the intensity process under analysis. The second column displays the different impacts on the portfolio losses. Finally, the last column exhibits the spreads obtained for each of the different tranches.

Firstly, it can be seen that variations in the losses size, given by the parameter γ, seem to produce a high variability on the tranche spreads obtained. For instance, considering the lower impact on losses ($\gamma = 0.005$), we obtain the same values *independently* of the set of parameters used.[19] The influence of the volatility parameter σ seems to be irrelevant: likewise, those differences could be explained by the numerical implementation of the model.

[19]Differences in basis points of some spreads are explained by the variance of the simulation procedure used.

Table 5.1: Simulations for the one-factor model of Longstaff and Rajan (2007)

Parameters			Size losses	Tranches (basis points)				
α	β	σ	γ	0–3	3–7	7–10	10–15	15–30
0.50	0.60	0.05	0.005	94	0	0	0	0
			0.050	1,839	355	0	0	0
			0.100	1,852	1,852	1,079	0	0
			0.350	1,828	1,828	1,828	1,828	1,634
0.50	0.60	0.15	0.005	94	0	0	0	0
			0.050	1,843	355	0	0	0
			0.100	1,862	1,862	1,084	0	0
			0.350	1,842	1,842	1,842	1,842	1,634
1.50	0.60	0.15	0.005	102	0	0	0	0
			0.050	4,577	421	0	0	0
			0.100	4,550	4,550	1,704	0	0
			0.350	4,549	4,549	4,549	4,549	3,553
8.33	10.00	0.35	0.005	94	0	0	0	0
			0.050	1,860	357	0	0	0
			0.100	1,815	1,815	1,067	0	0
			0.350	1,828	1,828	1,828	1,828	1,627

Simulations for the one-factor model of Longstaff and Rajan (2007). First three columns display the diffusion parameters, respectively. Fourth column exhibits the size losses. Finally, the table displays the different tranches values in basis points from fifth to the end columns.

The increase of the parameter α, related with the actual mean arrival probability of credit events (α/β) produces large variations on tranche credit spreads, as is also displayed in Table 1: proportional increases in α, β parameters (from $\alpha = 1.50$, $\beta = 0.60$ to $\alpha = 8.33$, $\beta = 10.00$) keep the long-run mean of the intensity process constant. As a result, no changes are observed with respect to previous cases.

Table 2 displays the simulations for the two-factor version of Longstaff and Rajan (2007). The frst main column refers to parameters of the intensity process (5.5.3.3) for $i = 1, 2$. Values under the " Paramenters" named column must be intended as follows: values under the α column correspond to $\alpha_1 = 0.50$, $\alpha_2 = 0.02$; values under the β column correspond to $\beta_1 = 0.60$, $\beta_2 = 0.60$ and so on. Second and third main columns are those of Size losses and Tranches, in a similar way to those in Table 1.

As expected from observing the expression (5.5.2.3), the magnitude of losses are determinant when it comes to price the tranches: to a certain degree, lower γ results on lower prices for tranches. This last begins to be conditioned to the combination of losses of the two processes. For instance, considering low losses ($\gamma_1 = 0.005$, $\gamma_2 = 0.100$ or $\gamma_1 = 0.005$, $\gamma_2 = 0.035$) seems not to generate as a richiest variety of spreads as higher ones (for example, $\gamma_1 = 0.050$, $\gamma_2 = 0.050$).

Again, an increase in the α parameter (that controls the long-run mean of the process (5.5.3.3)), results on higher tranches spreads.

Finally, Table 3 shows the simulations for the three-factor model. Basically, this table reveals that the addition of a third factor enhances the spreads obtained for the one- and two-factor versions of the model. Particularly, in addition to bespoke intensity processes, the heterogeneity in the losses impact results in a diversity of spreads for all tranches.

To put it briefly, results of this section seem to address that changes in the σ (volatility) parameter seem not to produce changes in tranches values. By contrast, variations in a) the long-run mean and b) the impact of the portfolio sizes must be taken into account. This is the purpose of the subsequent sections.

5.4.2 Longstaff and Rajan (2007) model with jumps

Here, we extend the model of Longstaff and Rajan (2007) by adding jumps in the background processes that control the arrival of credit events. As previously cited, the underlying idea is to introduce discontinuities - jumps -in the paths of the intensity process that could be explained as non-marginal changes in the instantaneous probability of default. The arrival of unex-

Table 5.2: Simulations for the two-factor model of Longstaff and Rajan (2007)

Parameters						Size losses		Tranches (basis points)				
α		β		σ		γ_1	γ_2	0–3	3–7	7–10	10–15	15–30
0.50	0.02	0.60	0.60	0.15	0.20	0.005	0.100	110	0	0	0	0
						0.005	0.350	109	0	0	0	0
						0.050	0.050	1,890	436	61	0	0
						0.100	0.350	1,905	1,905	1,162	75	15
1.50	0.02	0.60	0.60	0.15	0.20	0.005	0.100	117	0	0	0	0
						0.005	0.350	117	0	0	0	0
						0.050	0.050	4,641	519	62	0	0
						0.100	0.350	4,637	4,637	1,849	77	15
0.50	0.20	0.60	0.60	0.15	0.20	0.005	0.100	185	0	0	0	0
						0.005	0.350	185	0	0	0	0
						0.050	0.050	2,507	1,099	493	0	0
						0.100	0.350	2,475	2,475	1,825	636	85

Simulations for the two-factor model of Longstaff and Rajan (2007). First block of columns displays the different sets of diffusion parameters used on the simulations. Second block of columns describes the magnitude of the size loses. Finally, the third block contains the different tranches values in basis points obtained.

Table 5.3: Simulations for the three-factor model of Longstaff and Rajan (2007).

	Size losses			Tranches (basis points)				
Set	γ_1	γ_2	γ_3	0–3	3–7	7–10	10–15	15–30
1	0.005	0.050	0.350	143	50	7	7	7
2	0.150	0.050	0.350	1,911	1,897	1,865	975	19
3	0.050	0.150	0.350	1,891	788	51	51	16
4	0.350	0.150	0.050	1,871	1,870	1,868	1,860	1,422

Simulations for the three-factor model of Longstaff and Rajan (2007). Parameters are fixed at $\alpha_1 = 0.50$, $\alpha_2 = 0.02$, $\alpha_3 = 0.001$, $\beta_1 = \beta_2 = \beta_3 = 0.60$, $\sigma_1 = 0.15$, $\sigma_2 = 0.20$ and $\sigma_3 = 0.15$.

pected information could cause these kind of effects.

To keep manageable the number of parameters under study, those which correspond to jumps in the intensity processes (the arrival intensity δ and the mean μ of the jump amplitude) will be considered equal for all the intensity process. All the tables in this and subsequent sections that contain these parameters must be intended in this form.

Tables 4 to 6 display the simulations for the one-, two- and three-factors of the Longstaff and Rajan (2007) model with jumps. By and large, same qualitative conclusions than aforementioned in the previous section arise: the addition of jumps results in a high value of tranche spreads. Basically, jumps introduce more volatility in the default process, which results in higher probability of default. As result of this, the more we increase the probability of a credit event, the more credit events we have. Another fact arises from these tables: considering the parameters under study, it seems that higher frequencies of relatively small jumps produce bigger impacts on the tranche spreads than lower frequencies with large leaps.

5.4.3 Longstaff and Rajan (2007) model with jumps and random losses

In view of the important role of the parameter γ in the total portfolio losses, we analyze now the consequences of randomizing the impact of losses in the Longstaff and Rajan (2007) model. We will also include jumps in the intensity process with the intention of studying simultaneously the joint

Table 5.4: Simulations for the one-factor model with jumps of Rajan and Longstaff (2006)

Jump Parameters		Size losses	Tranches (basis points)				
δ	μ	γ	0–3	3–7	7–10	10–15	15–30
0.02	0.10	0.005	98	0	0	0	0
		0.050	2,342	378	0	0	0
		0.100	2,336	2,336	1,243	0	0
		0.350	2,336	2,336	2,336	2,336	2,035
0.02	0.20	0.005	99	0	0	0	0
		0.050	2,653	388	0	0	0
		0.100	2,673	2,673	1,337	0	0
		0.350	2,633	2,633	2,633	2,633	2,260
0.10	0.10	0.005	100	0	0	0	0
		0.050	3,430	405	0	0	0
		0.100	3,421	3,421	1,509	0	0
		0.350	3,397	3,397	3,397	3,397	2,805

Simulations for the one-factor model with jumps of Longstaff and Rajan (2007). Diffusion parameters are fixed at $\alpha = 0.50$, $\beta = 0.60$ and $\sigma = 0.15$ for the three cases under study.

effect on the spreads.

Tables 7 to 9 show the simulations for the one-, two- and three-factors of the Longstaff and Rajan (2007) model with jumps in the background default process and random losses. The variables that model the size losses are exponentially distributed with mean γ.

By and large, we can observe that considering a random impact of credit events results in a richest diversity of spreads for all tranches. When it comes to comparing against the previous results for the Longstaff and Rajan (2007) model with constant γ, this is more evident. A possible explanation can be related to the fact that taking γ as random provides a wide range of different losses without necessity of considering combinations of fixed values for this

Table 5.5: Simulations for the two-factor model with jumps of Rajan and Longstaff (2006).

Jump param		Size losses		Tranches (basis points)				
δ	μ	γ_1	γ_2	0–3	3–7	7–10	10–15	15–30
0.02	0.10	0.005	0.100	1,068	937	907	0	0
		0.005	0.350	1,075	946	946	946	929
		0.050	0.050	2,737	1,330	631	0	0
		0.100	0.350	2,720	2,720	2,131	928	926
0.02	0.20	0.005	0.100	1,553	1,402	1,340	0	0
		0.005	0.350	1,556	1,407	1,407	1,407	1,373
		0.050	0.050	3,184	1,791	845	0	0
		0.100	0.350	3,145	3,145	2,623	1,370	1,367
0.10	0.10	0.005	0.100	2,575	2,387	2,227	0	0
		0.005	0.350	2,574	2,388	2,388	2,388	2,296
		0.050	0.050	4,218	2,749	1,166	0	0
		0.100	0.350	4,159	4,159	3,705	2,345	2,335

Simulations for the two-factor model with jumps of Rajan and Longstaff (2006). Parameters are fixed at $\alpha_1 = 0.50$, $\alpha_2 = 0.02$, $\beta_1 = \beta_2 = 0.60$, $\sigma_1 = 0.15$ and $\sigma_2 = 0.20$. For ease of exposition, jump parameters are equal for both processes.

parameter.

As revealed by the simulations in Table 1 for the mean value of size losses $\gamma = 0.005$, it is also important to notice that the magnitude of the impact in losses continues being important. Moreover, the effect of introducing γ as random produces an smoothness of tranches values with respect to previous cases, a point that will be discussed in the next section.

5.4.4 Sectional comparison

Finally, we study in detail the individual effects of adding jumps and randomizing the losses in the Longstaff and Rajan (2007) for one, two, and

Table 5.6: Simulations for the three-factor model with jumps of Rajan and Longstaff (2006).

	Size losses			Tranches (basis points)				
Set	γ_1	γ_2	γ_3	0–3	3–7	7–10	10–15	15–30
1	0.005	0.050	0.350	1,595	1,155	881	881	878
2	0.150	0.050	0.350	3,050	2,871	2,728	2,275	975
3	0.050	0.150	0.350	3,001	1,948	1,471	1,449	954
4	0.350	0.150	0.050	3,047	2,900	2,782	2,714	2,322

Simulations for the three-factor model with jumps of Rajan and Longstaff (2006). Parameters are fixed at $\alpha_1 = 0.50$, $\alpha_2 = 0.02$, $\alpha_3 = 0.001$, $\beta_1 = \beta_2 = \beta_3 = 0.60$, $\sigma_1 = 0.15$, $\sigma_2 = 0.20$ and $\sigma_3 = 0.15$. For ease of exposition, jump parameters are equal for the processes, $\delta = 0.02$ and $\mu = 0.10$.

three factors, respectively.

Figures 3 to 5 show respectively the different versions of the model under study. LR represents the Longstaff and Rajan (2007) model and it refers to its simplest version. Then, we display the "jumps" and "jumps plus random variable" versions of the model. Finally, we include the mean spreads cited in their article with the purpose of comparing the performance of the models. It is important to notice that we are interested in comparing the different models (one, two or three factors) to their extended versions, taking the empirical spreads *just* as an order reference.[20]

We begin the study for the one-factor version of Longstaff and Rajan (2007). Figure 3 displays the spreads obtained for various tranches under different models. As expected, the one-factor versions of Longstaff and Rajan (2007) with constant loss size produces spreads so far from mean values. Naturally, the version with jumps shows higher spreads, due to the fact that jumps in the default process generate higher probability of arrivals of default events. Finally, it is important to note that a single factor version with random losses can capture the spreads given in the market, attending at bars on Figure 1.[21]

[20] To put it in other words, there is no intention here of doing an empirical analysis of the model. An study on this direction overcomes the objectives of this chapter.

[21] How to intend the default process in this context can be a subject of further research.

Table 5.7: Simulations for the one-factor model with jumps and exponential random losses of Longstaff and Rajan (2007)

Diffusion Parameters			Jump Parameters		Size losses	Tranches (basis points)				
α	β	σ	δ	μ	γ	0–3	3–7	7–10	10–15	15–30
0.50	0.60	0.15	0.02	0.10	0.005	98	0	0	0	0
					0.050	934	269	104	39	7
					0.100	1,323	570	309	168	46
					0.350	1,938	1,328	981	732	392
0.50	0.60	0.15	0.02	0.20	0.005	96	0	0	0	0
					0.050	943	253	96	34	4
					0.100	1,434	595	311	169	51
					0.350	2,152	1,467	1,068	774	415
0.50	0.60	0.15	0.10	0.10	0.005	98	0	0	0	0
					0.050	1,079	279	107	43	6
					0.100	1,640	637	333	181	49
					0.350	2,645	1,671	1,188	847	431

Simulations for the one-factor model with jumps and exponential random losses of Longstaff and Rajan (2007). First block of columns displays the different sets of diffusion parameters used on the simulations. Second and third blocks of columns describe the jump parameters and the magnitude of the size loses, respectively. Finally, the fourth block contains the different tranches values in basis points obtained.

Table 5.8: Simulations for the two-factor model with jumps of Rajan and Longstaff (2006).

Jump Parameters		Size losses (mean)		Tranches (basis points)				
δ	μ	γ_1	γ_2	0–3	3–7	7–10	10–15	15–30
0.02	0.10	0.005	0.100	863	427	258	148	46
		0.005	0.350	1,006	731	600	492	304
		0.050	0.050	1,625	662	308	144	25
		0.100	0.350	2,190	1,485	1,110	823	451
0.02	0.20	0.005	0.100	1,157	529	302	169	51
		0.005	0.350	1,412	997	782	618	360
		0.050	0.050	1,968	774	347	157	27
		0.100	0.350	2,677	1,862	1,388	1,056	555
0.10	0.10	0.005	0.100	1,647	631	331	182	53
		0.005	0.350	2,201	1,448	1,087	803	419
		0.050	0.050	2,622	926	386	165	28
		0.100	0.350	3,716	2,662	1,978	1,416	675

Simulations for the two-factor model with jumps of Rajan and Longstaff (2006). Parameters are fixed at $\alpha_1 = 0.50$, $\alpha_2 = 0.02$, $\beta_1 = \beta_2 = 0.60$, $\sigma_1 = 0.15$ and $\sigma_2 = 0.20$. Jump parameters are equal for both processes. Parameter in Size losses correspond to the mean of an exponential distributed random variable.

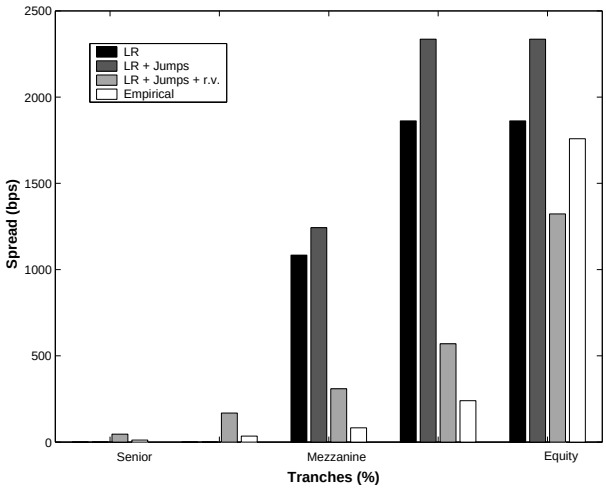


Figure 5.3: Comparison of spreads for different tranches and different versions of the Longstaff and Rajan (2007) model with one factor. Parameters are $\alpha_1 = 0.50$, $\beta_1 = 0.60$, $\sigma_1 = 0.15$, $\gamma_1 = 0.10$, $\delta_1 = 0.20$ and $\mu_1 = 0.10$. Finally, the γ_1 parameter corresponds to the mean of an exponential distributed random variable for the "random losses" version of the model.

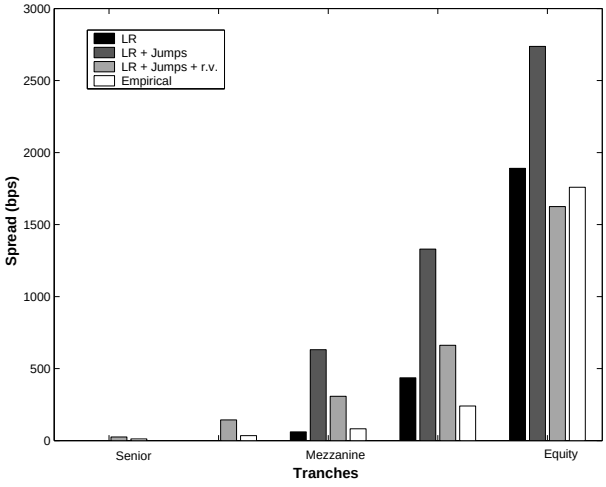


Figure 5.4: Comparison of spreads for different tranches and different versions of the Longstaff and Rajan (2007) model with two factors. Parameters are $\alpha_1 = 0.50$, $\alpha_2 = 0.02$, $\beta_1 = \beta_2 = 0.60$, $\sigma_1 = 0.15$, $\sigma_2 = 0.20$, $\gamma_1 = 0.05$, $\gamma_2 = 0.05$, $\delta_1 = \delta_2 = 0.20$ and $\mu_1 = \mu_2 = 0.10$. Finally, the parameters $\gamma_i,\ i = 1, 2$ correspond to the mean of an exponential distributed random variable for the "random losses" version of the model.

Table 5.9: Simulations for the three-factor model of Longstaff and Rajan (2007) with jumps and exponential random losses.

	Size losses			Tranches (basis points)				
Set	γ_1	γ_2	γ_3	0–3	3–7	7–10	10–15	15–30
1	0.005	0.050	0.350	1,404	939	733	580	347
2	0.150	0.050	0.350	2,610	1,929	1,471	1,141	645
3	0.050	0.150	0.350	2,345	1,590	1,232	982	598
4	0.350	0.150	0.050	2,772	2,272	1,863	1,468	830

Simulations for the three-factor model of Longstaff and Rajan (2007) with jumps and exponential random losses. Parameters are fixed at $\alpha_1 = 0.50$, $\alpha_2 = 0.02$, $\alpha_3 = 0.001$, $\beta_1 = \beta_2 = \beta_3 = 0.60$, $\sigma_1 = 0.15$, $\sigma_2 = 0.20$ and $\sigma_3 = 0.15$. For ease of exposition, jump parameters are equal for the three processes, $\delta = 0.02$ and $\mu = 0.10$. Finally, parameters in Size losses correspond to the mean of an exponential distributed random variable.

Figure 4 compares the spreads for alternative versions of a two- factor model in (5.5.3.2). Again, the model with jumps and constant γ generates the highest spreads in equity-mezzanine tranches. We observe here that the addition of a second factor results in spreads obtained by the constant size loss model closer to those obtained with random size losses for every tranches.

Finally, Figure 5 shows the spreads for the three-factor version of Longstaff and Rajan (2006) under study. As previously deduced in Figure 2, the addition of a third factor serves to provide spreads close to those observed empirically. The extensions that consider "jumps" and "jumps with random losses" seem to produce spreads far away from those observed in the market.

5.5 Conclusions

The development of markets where investors are able to trade their exposure to credit events has contributed to the appareance of a huge diversity of standardized financial assets. A wide range of these new products such as First-to-Default Swaps (FtD) or Credit Default Swap Indexes are examples

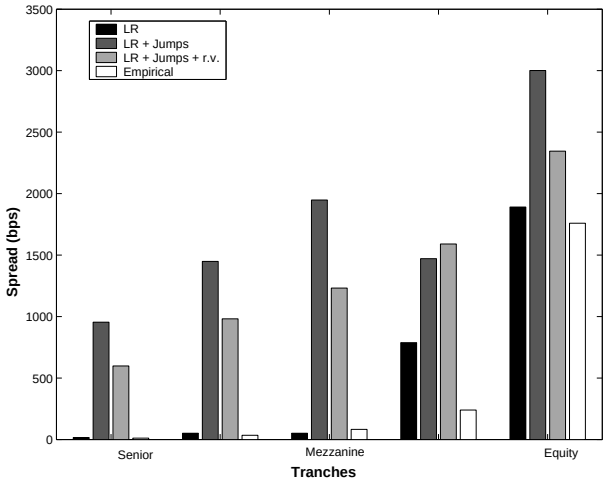


Figure 5.5: Comparison of spreads for different tranches and different versions of the Longstaff and Rajan (2007) model with three factors. Parameters are $\alpha_1 = 0.50$, $\alpha_2 = 0.02$, $\alpha_3 = 0.001$, $\beta_1 = \beta_2 = \beta_3 = 0.60$, $\sigma_1 = 0.15$, $\sigma_2 = 0.20$, $\sigma_3 = 0.15$, $\gamma_1 = 0.050$, $\gamma_2 = 0.15$, $\gamma_2 = 0.35$, $\delta_1 = \delta_2 = \delta_3 = 0.20$ and $\mu_1 = \mu_2 = \mu_3 = 0.10$. Finally, the parameters γ_i, $i = 1, 2, 3$ correspond to the mean of an exponential distributed random variable for the "random losses" version of the model.

of how the investors can reduce the hedging cost of portfolios including a large number of firms without necessity of buying insurance contracts on each individual firm.

This chapter has focused on these families of basket credit derivatives with special emphasis on Credit Default Swap Indexes, where the role of the default correlation (intended as the different relationship of dependence between defaults that can be given among firms) is crucial to price these assets.

Then, the aim of this chapter has been to study the pricing of standardized CDO tranches using the intensity based framework of clustered defaults developed by Longstaff and Rajan (2007). These authors assume that defaults can be bursted, that is, that joint defaults can happen simultaneously, an approach linked to the contagious scheme of Davis and Lo (1999). Then, this model contrasts to the individual approach of Duffie and Garlenau (2001), where the total number of defaulted firms is conditioned by the level of correlation among the obligors,

We have performed a Monte Carlo study to explore numerically the capability of this model to generate values for standardized CDOs tranches under two different circumstances, particularly those referred to the number

of factors used and the impact of credit events on the loss distribution.

We have also extended the original model in two directions: firstly, by adding jumps to the default process; then, by randomizing the impact of defaults on the loss distribution. To the best of our knowledge, no similar study has yet been reported.

Our results seem to suggest that a three-factor version of Longstaff and Rajan (2006) with constant losses impact is flexible enough to reproduce the spreads given by the market. In addition to this, the inclusion of jumps to the default process results in an high arrival of credit events, as corroborated by the high values spreads for equity tranches. Finally, the alternative of random losses can be helpful when dealing with one- and two-factor models, but it seems to be irrelevant in the case of three-factor models. Anyway, this last point can be developed as a subject for futher research.

Bibliography

[1] 2006. Bba credit derivatives report.

[2] Aase, K. and P. Guttorp 1987. Estimation in models for security prices. *Scandinavian Actuarial Journal* 6:211–224.

[3] Abramovitz, M. and I. Stegun 1971. *Handbook of Mathematical Functions*. New York : Dover Press.

[4] Acerbi, C., C. Nordio, and C. Sirtori. 2001. Expected shortfall as a tool for financial risk management.

[5] Altmann, T., T. Schmidt, and W. Stute 2007. Jump diffusion processes with shot noise effects and their applications to finance. *IJTAF*.

[6] Amin, K. I. 1993. Jump diffusion option valuation in discrete time. *Journal of Finance* 48(5):1833–1863.

[7] Andersen, L. and J. Andreasen 2000. Jump-diffusion processes: Volatility smile fitting and numerical methods of option pricing. *Review of Derivatives Research* 4:231–262.

[8] Andersen, L., J. Sidenius, and S. Basu 2003. All your hedges in one basket. *Risk* 67–72.

[9] Andersen, T. G., L. Benzoni, and J. Lund 2002. An empirical investigation of continuous-time equity return models. *Journal of Finance* 57(3):1239–1284.

[10] Artzner, P., F. Delbaen, J. Eber, and D. Heath 1999. Coherent measures of risk. *Mathematical Finance* 9(3):203–208.

[11] Aït-Sahalia, Y. 2004. Disentangling diffusion from jumps. *Journal of Financial Economics* 74:487–528.

[12] Ball, C. A. and W. N. Torous 1983. A simplified jump process for common stock returns. *Journal of Financial and Quantitative Analysis* 18(1):53–65.

[13] Ball, C. A. and W. N. Torous 1985. On jumps in common stock prices and their impact on call option pricing. *Journal of Finance* 40(1):155–173.

[14] Barndorff-Nielsen, O., E. Nicolato, and N. Shephard 2002. Some recent developments in stochastic volatility modelling. *Quantitative Finance* 2:11–23.

[15] Barone-Adesi, G. and A. Gigli. November 2002. Managing electricity risk.

[16] Beckers, S. 1981. A note on estimating the parameters of the jump-diffusion model of stock returns. *Journal of Financial and Quantitative Analysis* 36(1):127–140.

[17] Benito, F., A. León, and J. Nave 2007. Modeling the euro overnight rate. *Journal of Empirical Finance*.

[18] Bevan, S., R. Kullberg, and J. Rice 1979. An analysis of cell membrane noise. *The Annals of Statistics* 7(2):237–257.

[19] Bielecki, A. and F. Rutkowski 2002. *Credit Risk: Modelling, Valuation and Hedging*. Berlin and New York : Springer Finance.

[20] Black, F. and J. Cox 1976. Valuing corporate securities: Some effects of bond indenture provisions. *Journal of Finance* 31(2):351–367.

[21] Black, F. and M. Scholes 1973. The pricing of options and corporate liabilities. *Journal of Political Economy* 81(3):637–659.

[22] Blanco, R., S. Brennan, and I. Marsh 2006. An empirical analysis of the dynamic relation between investment-grade bonds and credit default swaps. *Journal of Finance* 40(5):2255–2281.

[23] Bondesson, L. 1988. *Shot-Noise Processes and Distributions, in Encyclopedia of Statistical Science* volume 8. New York : Wiley.

[24] Campbell, J., A. Lo, and A. MacKinlay 1997. *The Econometrics of Financial Markets*. Princeton, New Jersey : Princeton University Press.

[25] Carrasco, M., M. Chernov, J. Florens, and E. Ghysels. 2002. Efficient estimation of jump diffusions and general dynamic models with a continuum of moment conditions.

[26] Carrasco, M. and J. Florens. 2000. Estimation of a mixture via the empirical characteristic function.

[27] Casas, I. and J. Gao. 2004. Long-range dependent continuous-time financial econometric models: Theory and practice.

[28] Chacko, G. and L. Viceira 2003. Spectral gmm estimation of continuous-time processes. *Journal of Econometrics* 116(1):259–292.

[29] Chan, K., G. A. Karolyi, F. Longstaff, and A. Sanders 1992. An empirical comparison of alternative models of the short-term interest rate. *Journal of Finance* 47(2):1209–1227.

[30] Chan, T. 1999. Pricing contingent claims on stocks driven by levý processes. *The Annals of Applied Probability* 9(2):504–528.

[31] Cliff, M. March 2003. Gmm and minz program libraries for matlab.

[32] Cochrane, J. 2005. *Asset Pricing*. New Jersey : Princeton University Press.

[33] Cox, J., J. Ingersoll, and S. Ross 1985. A theory of the term structure of interest rates. *Econometrica* 53(2):385–407.

[34] Cox, J., S. Ross, and M. Rubinstein 1979. Option pricing: A simplified approach. *Journal of Financial Economics* 7:229–263.

[35] Daley, D. and D. Vere-Jones 1988. *An Introduction to the Theory of Point Processes*. Berlin and New York : Springer-Verlag.

[36] Das, S. R. and R. K. Sundaram 1999. Of smiles and smirks: A term structure perspective. *Journal of Financial and Quantitative Analysis* 34(2):211–239.

[37] Das, S. 2002. The surprise element: jumps in interest rates. *Journal of Econometrics* 106:27–65.

[38] Davis, M. and V. Lo. 1999. Infectious default.

[39] Delbaen, F. and W. Schachermayer 1997. Non-arbitrage and the fundamental theorem of asset pricing: Summary of main results. *Proceedings of Symposia in Applied Mathematics* 00.

[40] Dempster, A., N. Laird, and D. Rubin 1977. Maximum likelihood estimation from incomplete data via the em algorithm. *Journal of the Royal Statistical Society* 39:1–38.

[41] Duffee, G. 1999. Estimating the price of default risk. *The Review of Financial Studies* 12(1):197–226.

[42] Duffee, G. 2002. Term premia and interest rate forecasts in affine models. *Journal of Finance* 57(1):405–443.

[43] Duffee, G. and R. Stanton. March 2004. Estimation of dynamic term structure models.

[44] Duffie, D. and N. Garlenau 2001. Risk and valuation of collateralized debt obligations. *Financial Analysts Journal* 57(1):41–59.

[45] Duffie, D. and D. Lando 2001. Term structures of credit spreads with incomplete accounting information. *Econometrica* 69(3):633–664.

[46] Duffie, D. and J. Pan 1997. An overview of value at risk. *Journal of Derivatives* (Spring):7–49.

[47] Duffie, D., J. Pan, and K. Singleton 2000. Transform analysis and asset pricing for affine jump-diffusions. *Econometrica* 68:1343–1376.

[48] Duffie, D. and K. Singleton 1993. Simulated moments estimation of markov models of asset prices. *Econometrica* 61:929–952.

[49] Duffie, D. and K. Singleton 1999. Modeling term structures of defaultable bonds. *Review of Financial Studies* 12:687–720.

[50] Duffie, D. and K. Singleton 2003. *Credit Risk: Pricing, Measurement and Management*. Princeton, New Jersey : Princeton University Press.

[51] Elizalde, A. 2005. Credit risk models iv: Understanding and pricing cdos.

[52] Escribano, A., J. P. na, and P. Villaplana. June 2002. Modeling electricity prices: International evidence.

[53] Fan, Y. 1997. Goodness-of-fit tests for a multivariate distribution by the empirical characteristic function. *Journal of Multivariate Analysis* 62(1):36–63.

[54] Feller, W. 1971. *An Introduction to Probability Theory and Its Applications* volume 2. New York : Wiley Series in Probability and Statistics.

[55] Felsenheimer, J., P. Gisdakis, and M. Zaiser. 2004. Dj itraxx: Credit at its best!

[56] Feuerverger, A. and P. McDunnough 1981a. On the efficiency of empirical characteristic function procedures. *Journal of the Royal Statistical Society B* 43(1):20–27.

[57] Gallant, A. and G. Tauchen 1996. Which moments to match? *Econometric Theory* 12:657–681.

[58] Gao, J. 2004. Modelling long-range-dependent gaussian processes with application in continuous-time financial models. *Journal of Applied Probability* 41:467–482.

[59] Gibson, M. 2004. Understanding the risk of synthetic cdos.

[60] Giesecke, K. Credit Risk: Models and Management volume 2 chapter Credit Risk Modelling and Valuation: an Introduction. Riskbooks London 2004.

[61] Glasserman, P. and S. Suchintabandid 2007. Correlation expansions for cdo pricing. *Journal of Banking and Finance* 31(5):1375–1398.

[62] Gourieroux, C. and A. Monfort 1996. *Simulation-Based Econometric Methods*. New York : Oxford University Press.

[63] Gourieroux, C., A. Monfort, and E. Renault 1993. Indirect inference. *Journal of Applied Econometrics* 8:85–118.

[64] Green, W. H. 2003. *Econometric Analysis*. New Jersey : Pearson Education.

[65] Gregory, J. and J. Laurent 2004. In the core of correlation. *Risk* Octobre:87–91.

[66] Gregory, J. and J. Laurent 2005. Basket default swaps, cdo's and factor copulas. *Journal of Risk* 7(4):103–122.

[67] Gupton, G., C. Finger, and M. Bhatia. 1997. Creditmetrics. technical document.

[68] Hamilton, D., P. Varma, S. Ou, and R. Cantor. 2005. Default and recovery rates of corporate bond issuers, 1920-2004.

[69] Hamilton, J. 1994. *Time Series Analysis*. Princeton, New Jersey : Princeton University Press.

[70] Hannan, E. 1970. *Multiple Time Series*. New York : John Wiley and Sons.

[71] Hansen, L. 1982. Large sample properties of generalized method of moments estimators. *Econometrica* 50:1029–1054.

[72] Honoré, P. November 1998. Pitfalls in estimating jump-diffusion models.

[73] Hull, J. and A. White 2004. Valuation of a cdo and an n-th to default cds without monte carlo simulation. *Journal of Derivatives* Winter:8–23.

[74] Jakola, M. 2006. Credit default swap index options.

[75] Jarrow, R. and S. Turnbull 1995. Pricing derivatives on financial securities subject to credit risk. *Journal of Finance* 50(1):53–85.

[76] Jiang, G. November 1998. Jump-diffusion model exchange rate dynamics: estimation via indirect inference.

[77] Jiang, G. and J. Knight 2002. Estimation of continuous-time processes via the empirical characteristic function. *Journal of Business and Economic Statistics* 20(2):198–212.

[78] Jorion, P. 1988. On jump processes in the foreign exchange and stock markets. *Review of Financial Studies* 1(4):427–445.

[79] Klüppelberg, C. and T. Mikosch 1995. Explosive poisson shot noise processes with applications to risk reserves. *Bernoulli* 1(1/2):125–147.

[80] Knight, J. and S. Satchell 1997. The cumulant generating function estimation method. *Econometric Theory* 13(2):170–184.

[81] Lando, D. 2004. *Credit Risk Modeling: Theory and Applications*. Princeton, New Jersey : Princeton University Press.

[82] Leland, H. and K. Toft 1996. Optimal capital structure, endogenous bankruptcy and the term structure of credit spreads. *Journal of Finance* 51(3):987–1019.

[83] Lo, A. 1991. Long-term memory in stock market prices. *Econometrica* 59(5):1279–1313.

[84] Longstaff, F. A. and A. Rajan. 2007. An empirical analysis of the pricing of collateralized debt obligations.

[85] Lucas, A., P. Klaassen, P. Spreij, and S. Staetmans 2001. An analytic approach to credit risk of large corporate bond and loan portfolios. *Journal of Banking and Finance* 25(9):1635–1664.

[86] Lucia, J. and E. Schwartz 2001. Electricity prices and power derivatives. evidence from nordic power exchange. *Review of Derivatives Research* 5(1):5–50.

[87] Mardia, K., J. Kent, and J. Bibby 1979. *Multivariate Analysis*. London : Academic Press.

[88] Mashal, R. and M. Naldi 2002. Calculating portfolio loss. *Risk* August:82–86.

[89] McNeil, A., R. Frey, and P. Embrechts 2005. *Quantitative Risk Management*. New Jersey : Princeton University Press.

[90] Merton, R. 1974. On the pricing of corporate debt: The risk structure of interest rates. *Journal of Finance* 29(2):449–470.

[91] Merton, R. 1976b. Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics* 3:125–144.

[92] Moreno, M. and J. P. na. Forecasting Financial Markets: Exchange Rates, Interest Rates and Asset Management chapter On the Term Structure of Interbank Interest Rates: Jump-diffusion Processes and Option Pricing. John Wiley and Sons Ltd. 1996.

[93] Mortensen, A. 2006. Semi-analytical valuation of basket credit derivatives in intensity-based models. *Journal of Derivatives* Summer:8–26.

[94] Navas, J. 2003. Calculation of volatility in a jump-diffusion model. *Journal of Derivatives* Winter:66–72.

[95] Newey, W. and D. West 1987. A simple positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica* 55:703–708.

[96] Novales, A. 1993. *Econometría*. Madrid : McGraw-Hill.

[97] Parzen, E. 1962. *Stochastic Processes*. San Francisco : Holden Day Series in Probability and Statistics.

[98] Press, S. 1967. A compound events model for security prices. *Journal of Bussiness* 40:317–335.

[99] Priestley, M. 1981. *Spectral Analysis and Time Series*. New York : Academic Press.

[100] Rice, S. 1954. *Mathematical Analysis of Random Noise, from Selected Papers on Noise and Stochastic Processes*. New York : N. Wax, Dover Publications.

[101] Robinson, P. and J. V. Sanz 2006. Modified whittle estimation of multilateral models on a lattice. *Journal of Multivariate Analysis* 97(5):1090–1120.

[102] Ross, S. 1996. *Stochastic Processes*. Wiley : Princeton University Press.

[103] Samuelson, P. 1965. Rational theory of warrant pricing. *Industrial Management Review* 6:13–31.

[104] Schmidt, T. and W. Stute 2004. Credit risk - a survey. *Contemporary Mathematics* 336:75–115.

[105] Schönbucher, P. 2003. *Credit Derivatives Pricing Models: Models, Pricing, Implementation*. West Sussex, England : John Wiley and Sons Ltd.

[106] Schoutens, W. 2003. *Lévy Processes in Finance: Pricing Financial Derivatives*. : Wiley.

[107] Scott, L. 1997. Pricing stock options in a jump-diffusion model with stochastic volatility and interest rates: Applications of fourier inversion methods. *Mathematical Finance* 7(4):413–424.

[108] Shaw, W. April 2006. New methods for managing student's t distribution.

[109] Singleton, K. 2001. Estimation of affine asset pricing models using the empirical characteristic function. *Journal of Econometrics* 102:111–141.

[110] Singleton, K. 2006. *Empirical Dynamic Asset Pricing*. Princeton, New Jersey : Princeton University Press.

[111] Tavakoli, J. 2003. *Collateralized Debt Obligations and Structured Finance. New Developments in Cash and Synthetic Securitization*. Hoboken, New Jersey : John Wiley and Sons, Inc.

[112] Vasicek, O. 1977. An equilibrium characterization of the term structure. *Journal of Financial Economics* 5:177–188.

[113] Vasicek, O. 1991. Limiting loan loss probability distribution.

[114] Villaplana, P. March 2003. Pricing power derivatives: a two-factor jump-diffusion approach.

[115] Weron, R. November 2005. Heavy tails and electricity prices.

[116] Whittle, P. 1953. Estimation and information in stationary time series. *Arkiv for Matematik* 2:423–434.

[117] Yu, J. 2004. Empirical characteristic function estimation and its applications. *Econometric Reviews* 23(2):93–123.

[118] Zhou, H. January 2001. Finite sample properties of emm, gmm, qmle and mle for a square-root interest rate diffusion model.