Alma Mater Studiorum – Università di Bologna

DOTTORATO DI RICERCA IN

ECONOMIA

Ciclo XXIX

Settore Concorsuale di afferenza: 13/A1

Settore Scientifico disciplinare: SECS-P/01

ESSAYS IN FINANCIAL ECONOMETRICS

Francesca Lilla

Coordinatore Dottorato

Prof. Vincenzo Denicolò

Relatore

Prof. Sergio Pastorello

Esame finale anno 2017

Abstract

The first paper sheds light on the informational content of high frequency data and daily data. I assess the economic value of the two family models comparing their performance in forecasting asset volatility through the Value at Risk metric. In running the comparison this paper introduces two key assumptions: jumps in prices and leverage effect in volatility dynamics. Findings suggest that high frequency data models do not exhibit a superior performance over daily data models.

In the second paper, building on Majewski et al. (2015), I propose an affine-discrete time model, labeled VARG-J, which is characterized by a multifactor volatility specification. In the VARG-J model volatility experiences periods of extreme movements through a jump factor modeled as an Autoregressive Gamma Zero process. The estimation under historical measure is done by quasi-maximum likelihood and the Extended Kalman Filter. This strategy allows to filter out both volatility factors introducing a measurement equation that relates the Realized Volatility to latent volatility. The risk premia parameters are calibrated using call options written on S&P500 Index. The results clearly illustrate the important contribution of the jump factor in the pricing performance of options and the economic significance of the volatility jump risk premia.

In the third paper, I analyze whether there is empirical evidence of contagion at the bank level, measuring the direction and the size of contagion transmission between European markets. In order to understand and quantify the contagion transmission on banking market, I estimate the econometric model by Aït-Sahalia et al. (2015) in which contagion is defined as the within and between countries transmission of shocks and asset returns are directly modeled as a Hawkes jump diffusion process. The empirical analysis indicates that there is a clear evidence of contagion from Greece to European countries as well as self-contagion in all countries.

Contents

List of Figures

List of Tables

Chapter 1

High Frequency vs. Daily Resolution: the Economic Value of Forecasting Volatility Models

Forecasting volatility models typically rely on either daily or high frequency (HF) data and the choice between these two categories is not obvious. In particular, the latter allows to treat volatility as observable but they suffer from many limitations. HF data feature microstructure problem, such as the discreteness of the data, the properties of the trading mechanism and the existence of bid-ask spread. Moreover, these data are not always available and, even if they are, the asset's liquidity may be not sufficient to allow for frequent transactions. This paper considers different variants of these two family forecasting-volatility models, comparing their performance (in terms of Value at Risk, VaR) under the assumptions of jumps in prices and leverage effects for volatility. Findings suggest that daily-data models are preferred to HF-data models at 5% and 1% VaR level. Specifically, independently from the data frequency, allowing for jumps in price (or providing fat-tails) and leverage effects translates in more accurate VaR measure.

JEL-Classification: C58 C53 C22 C01 C13

Keywords: GARCH, DCS, jumps, leverage effect, high frequency data, realized variation, range estimator, VaR

1.1 Introduction

Modeling and forecasting volatility of asset returns are crucial for many applications, such as asset pricing model, risk management theory and portfolio allocation decisions. An earlier literature, including Engle (1982) and Bollerslev (1986) among others, has developed models of asset volatility dynamics in discrete time, known as heteroscedastic volatility models, i.e. ARCH-GARCH. Thanks to the availability of high frequency (HF) data, a new strand of literature has originated a new class of models based on the Realized Volatility (RV) estimator, therefore introducing a non-parametric measure of return volatility (see Andersen et al., 001a, Barndorff-Nielsen, 2002 and Andersen et al., 2012). As the main innovation, RV models provides an *ex-post* observation of volatility, at odds with the standard ARCH-GARCH approach, that treats volatility as a latent variable. Although forecasting-volatility models based on HF data are getting more and more popular in the literature, the choice between HF-data and daily-data models is yet not obvious, in particular from an applied standpoint. In particular, the former still suffer from various limitations, that can be addressed only at the cost of a heavy manipulation of the original data.

One of the main issues is the presence of the market microstructure noise, which prevents from getting a perfect estimate (at the limit) of the returns' variance (see Hansen and Lunde, 2006 and Aït-Sahalia et al., 2005, 2011). The market microstructure noise may originate from different sources, including the discreteness of the data, the properties of the trading mechanisms and the existence of a bid-ask spread. Regardless of the source, when return from assets are measured based on their transaction prices over very tiny time intervals, these measures are likely to be heavily affected by the noise and therefore brings little information on the volatility of the price process. Since the level of volatility is proportional to the time interval between two successive observations, as the time interval increases, the incidence of the noise remains constant, whereas the information about the "true" value of the volatility increases. Therefore, there is a trade-off between high frequency and accuracy, which has led authors to identify an optimal sampling frequency of 5 minutes[1].

[1]Since the best remedy for market microstructure noise depends on the properties of the noise, if data sampled at higher frequency, e.g. tick-by-tick, are used the noise term needs to be modeled and, as far as I know, there is no unified framework about how to deal with it. Aït-Sahalia et al. (2005) define a new estimator, Two Scales Realized Volatility (TSRV), which takes advantages of the rich information of tick-by-tick data and corrects the effects of microstructure noise on volatility estimation. The authors, instead of sampling over a longer time horizon and discarding observations, make use of all data and model the noise as an "observation error". But the microstructure noise modeling goes beyond the scope of this work.

HF data also features another inconvenient: they are not always available and, even if they are, the asset may be not liquid enough to be frequently traded. On the contrary, daily data are relatively simple to record and collect and are commonly easy-to-get.

This paper sheds light on the choice between HF-data and daily data models, by assessing the economic value of the two family models, based on a comparison of their performance in forecasting asset volatility. Following the risk management perspective, I use value at risk (VaR) as the econometric metric of volatility forecastability, as suggested by Christoffersen and Diebold (2000).

VaR is defined as the quantile of the conditional portfolio distribution, and is therefore quite intuitive as a measure: indeed, it is the most popular quantitative measure of the market risk associated with a portfolio of assets, and is generally adopted by banks and required by regulators all over the world[2].

In running the comparison between HF-data and daily data models, this paper introduces two key assumptions. Firstly, the data generating process for asset prices features discontinuities in its trajectories, *jumps*[3]. Secondly, volatility (i.e. the standard deviation of asset return) reacts differently to changes in asset return which have the same magnitude, but different sign, *leverage effect*. These two assumptions represent the main novelty of this paper since none of the previous studies on the economic value of different forecasting-volatility models has investigated the matter under both jumps in price and leverage effect combined together. Giot and Laurent (2004) compare the performance of a daily ARCH-type model with the performance of a model based on the daily RV in a VaR framework. The authors find that VaR specification based on RV does not really improve the performance of a VaR model estimated using daily returns. This paper underlines an important issue: in economics applications, it is important to

[2]Banks often construct VaR from *historical simulation* (HS-VaR): VaR is the percentile of the portfolio distribution obtained using historical asset prices and today weights. This procedure is characterized by a slow reaction to market conditions and for the inability to derive the term structure of VaR. The VaR term structure explains how risk measures vary across different investment horizons. In HS-VaR, for example, if T-day 1% VaR is calculated, the 1-day 1% VaR is simply scaled by $\sqrt{T}$. This relation is valid only if daily returns are i.i.d. realizations of a Normal distribution. We know that is not the case since returns present leptokurtosis and asymmetry. The main limit of HS-VaR is the substitution of the conditional return distribution with the unconditional counterpart. Risk Metrics and GARCH models represent improvements over HS-VaR measure. Both of them provide an explicit assumption about the DGP and the conditional variance but they have also important differences. In addition to the estimation method: GARCH conditional volatility is estimated by maximizing the log-likelihood function while the parameters used in Risk Metrics are chosen in an ad hoc fashion, they differ for the possibility to account for the term structure of VaR. This is because GARCH process allows for mean reversion in volatility while Risk Metrics does not, reproducing a flat term structure for VaR.

[3]A continuous price process is a restrictive assumption since it is not possible to distinguish between the dynamic originated from the two sources of variability, i.e. continuous and discontinuous movements with consequences on the return generating process

recognize and take into account the key features of the empirical data in order to choose a valid data generating process. Clements et al. (2008) evaluate quantile forecasts focusing exclusively on models based on RV in order to understand if the results presented for stock returns can be carried over exchange rates. According to the results in Clements et al. (2008) the distributional assumption for expected future returns is needed for computing quantile, irrespective of the frequency of data used. Brownlees and Gallo (2010) forecast VaR using different volatility measures based on ultra-high-frequency data using a two-step VaR prediction procedure. They find that using ultra-high-frequency observations, VaR predictive ability is considerably improved upon relative to a baseline GARCH but not so relative to the range. The reason is related to the microstructure noise issue which arises when ultra high-frequency data are used. Indeed I want to contribute to the existing literature focusing on the measurement and the efficient use of the information embedded in HF data with respect to the information content of daily observations. Assuming both jumps and leverage effects in the returns dynamics for both data categories, I provide a more balanced comparison than in the previous work.
In the choice of the model to use for the comparison, I consider the GARJI model of Maheu and McCurdy (2004), as the baseline for the daily data models. The latter is a mixed-GARCH jump model which allows for asymmetric responses to past innovations in asset returns: the news impact (resulting in jump innovations) may have a feedback effect on the expected volatility, in addition to the feedback effect associated with the normal error term. For the case of HF data, I consider models in which Realized Volatility (RV) is decomposed into continuous and discontinuous volatility components. The continuous component is captured by means of the bi-power variation (BV), introduced by Barndorff-Nielsen and Shephard (2004), whereas the discontinuous component (JV) is obtained as the difference between RV and BV at given point in time[4]. In Andersen et al. (2007), JV is obtained considering only jumps that are found to be significative, and neglecting the others[5]. Corsi et al. (2010) consider instead all jumps, stressing the importance to correct the positive bias in BV due to jumps classified as consecutive. In this paper, I consider both these approaches and make a comparison among them, finding evidence in favor of jump identification strategy of Corsi et al. (2010) when the leverage effect is introduced. To account for the leverage effect, I introduce in this class of models the hetero-

[4]As shown in Andersen et al. (2002), Andersen et al. (2007), RV is a consistent estimator for the quadratic variation, whereas BV represents a consistent estimator of the continuous volatility component, i.e. the so-called integrated volatility, in the presence of jumping prices.

[5]The authors with significant jumps refer to large value of $RV_t - BV_t$ while small positive values are treated both as part of continuous sample path variation or as measurement errors.

geneous structure proposed by Corsi and Renó (2009).

Throughout this paper, the GARJI-VaR measures are obtained by following Chiu et al. (2005), that is, by adjusting for skewness and fat tails in the specification of the conditional distribution of returns[6]. The HF-VaR measures, instead, are computed by assuming a conditional Gaussian distribution for asset returns: as shown in Andersen et al. (2010), returns standardized for the square root of RV are indeed approximatively Normal[7].

In order to assess the model's capability to forecast future volatility, I implement a backtesting procedure based on both the Christoffersen (1998) test and the Kupiec (1995) test. In addition to comparing the economic value of daily data and HF-data models, the analysis performed in this paper sheds light on three other issues. The first is represented by the economic value per se, i.e. out of the comparison, of the class of forecasting volatility models adopting HF-data. This is done by considering different specifications of this family models. I first run a comparison among them (based on their forecasting performances); then, I compare some of them with their variant, obtained by using the Range estimator (RA) of Parkinson (1980). The choice of this particular benchmark is motivated by the fact that the RA estimator is likely to deliver a measure of volatility which lies in the middle of the measure obtained from HF estimators and that obtained from daily data models[8]. My findings suggest that HF-data models which explicitly provide both jumps and leverage factors stand out from the others in term of forecasting capability.

The second by-product of my analysis is a quantitative assessment of the importance of the explicit jump component in the conditional distribution of asset returns [9]. This point is addressed in both the family models considered in this paper. Hence, I first compare the forecasting volatility performances of each HF-data model with and without a decomposition of

[6]The computation of VaR measure requires, in addition to the conditional volatility dynamics, the specification of the conditional distribution of returns.VaR is a conditional risk measure so an assumption on the conditional distribution of returns is needed. Conditional normality is an acceptable assumption (returns standardized by their conditional volatility could be approximately Gaussian even if the unconditional returns are not Gaussian) only if the volatility model is able to fatten conditionally Gaussian tails enough to match the unconditional distribution. If this is not the case another conditional distributional assumption is necessary.

[7]This result is confirmed by the standardized returns of the sample used in this paper. See Section 1.2.

[8]The RA estimator exploits information on the highest and the lowest price recorded in a given day for a particular asset. In this respect, it requires information on the intra-day activity (going beyond the simple closing price of the asset), but without relying on further information, that might be not readily available).

[9]The presence of a jump component is justified both at theoretical and empirical level. From a theoretical perspective, an explicit discontinuous volatility-component allows to have information on the market response to outside news, which is key for many applications. From an empirical standpoint, instead, it is very difficult to distinguish power-type tails from exponential-type tails, given that is not clear to what extent the return distribution is heavily tailed. In this regard, the jump component of a jump-diffusion model may be interpreted as the market response to outside news: when good or bad news arrive at a given point in time, the asset price changes according to the jump size (and the jump sign) and an extreme sources of variation is added to the idiosyncratic component.

the RV into the continuous and the discontinuous component. Then, I run a similar analysis for the case of the daily data models, considering the GARCH-t model, as well as the Beta-t model[10] proposed by Harvey and Luati (2014). According to my analysis, introducing an explicit, persistent jump component in the conditional return dynamics (together with an asymmetric response to bad and good news into conditional volatility dynamics) may help to forecast the ex-post volatility dynamics and obtain more accurate VaR measures, at least at the VaR level required by Basel accords (1%). For HF-data models, accounting for jumping prices does not seem to improve significantly the accuracy of the estimates.

The last issue of my analysis is related to the importance of leverage effect in forecasting volatility. The findings in this paper recommend the explicit introduction of a factor that generates the asymmetric volatility response to price movements in the forecasting model.

The rest of the paper is organized as follows. Section 1.2 summarizes the volatility measures and the forecasting models based on both HF and daily data. Section 1.3 and Section 1.4 show, respectively, the backtesting methods used to evaluate forecasting models accuracy and the empirical results. Section 2.5 concludes.

1.2 Volatility Measures and Forecasts

1.2.1 Estimates of volatility with High Frequency Data

The RV measure is an estimator for the total quadratic variation, namely, it converges in probability, as the sampling frequency increases, to the continuous volatility component if there are no jumps. Instead, it converges to the sum of continuous and discontinuous volatility components if at least one jump occurs. As explained in Andersen et al. (2012), it is possible to use the daily RV measures, the *ex-post* volatility observations, to construct the *ex-ante* volatility forecasts. This is possible simply by using standard ARMA time series tools but it is important to take into account the difference with GARCH-type forecasting. The fundamental difference is that in the former case the risk manager treats volatility as observed while in the latter framework volatility is inferred from past returns conditional on a specific model. The idea behind the RV is the following: even if prices are not available on continuous basis, prices are recorded at higher frequency than daily. Using these squared returns a daily RV could easily be computed. In this way the ex-post volatility is considered as observable at each point in

[10] Beta-*t* model, belongs to the general class of Dynamic Conditional Score (DCS) model. They are also known as Generalized Autoregressive Score (GAS) model proposed by Creal et al. (2013).

time.

More precisely, the RV on day t based on returns at the Δ intraday frequency is

$$RV_t(\Delta) \equiv \sum_{j=1}^{N(\Delta)} r_{t,j}^2$$

where $r_{t,j} = p_{t-1+j\Delta} - p_{t-1+(j-1)\Delta}$ and $p_{t-1+j\Delta}$ is the log-price at the end of the jth interval on day t and $N(\Delta)$ is the number of the observations available at day t recorded at Δ frequency. In the absence of microstructure noise, as $\Delta \to 0$ the RV estimator approaches the integrated variance of the underlying continuous-time stochastic volatility process on day t:

$$RV_t \longrightarrow_p IV_t \qquad \text{where} \qquad IV_t = \int_{t-1}^{t} \sigma^2(\tau)\, d\tau$$

Furthermore, in this paper I assume that the the underlying price process is characterized by discontinuities. Indeed, the previous convergence is not valid but the RV estimators approaches in probability to the sum of the integrated volatility and the variation due to jumps that occurred on day t:

$$RV_t \longrightarrow_p \int_{t-1}^{t} \sigma^2(\tau)\, d\tau \quad + \quad \sum_{j=1}^{\zeta_t} J_{t,j}^2$$

If jumps ($J_{t,j}$) are absent, the second term vanishes and the realized volatility consistently estimates the integrated volatility. A nonparametric estimate of the continuous volatility component is obtained by using the bipower variation (BV) measures:

$$BV_t \equiv \frac{\pi}{2} \frac{N(\Delta)}{N(\Delta)-1} \sum_{j=1}^{N(\Delta)-1} |r_{t,j}||r_{t,j+1}| \tag{1.1}$$

Furthermore, the contribution to the total return variation stemming from the jump component (JV_t) is consistently estimated by

$$RV_t - BV_t \longrightarrow_p \sum_{j=1}^{\zeta_t} J_{t,j}^2$$

Considering the suggestion of Barndorff-Nielsen and Shephard (2004) the empirical measurements are truncated at zero in order to ensure that all of the daily estimates are nonnegative:

$$JV_t = max\{RV_t - BV_t, 0\} \tag{1.2}$$

According to Andersen et al. (2007), this truncation reduces the problem of measurement error with fixed sampling frequency but it captures a large number of nonzero small positive values in the jump component series. These small positive values can be treated both as part of the continuous sample path variation process or as measurement errors

In order to identify statistically significant jumps, i.e. large values of $RV_t - BV_t$, the authors suggest the use of the following statistic:

$$Z_t = \frac{log(RV_t) - log(BV_t)}{\sqrt{N(\Delta)^{-1}(\mu_1^{-4} + 2\mu_1^{-2} - 5)TQ_t BV_t^{-2}}} \longrightarrow_d N(0,1) \tag{1.3}$$

where $\mu_1 = \sqrt{2/\pi}$. In the denominator appears the realized tripower variation (TQ) that is the estimator of the integrated quarticity as required for a standard deviation notion of scale:

$$TQ_t = N(\Delta)\mu_{4/3}^{-3} \sum_{j=3}^{N(\Delta)} |r_{t,j}|^{4/3} |r_{t,j+1}|^{4/3} |r_{t,j+2}|^{4/3}$$

where $\mu_{4/3} = 2^{2/3}\Gamma(7/6)\Gamma(1/2)$. The significant jumps and the continuos component are identified and estimated respectively as:

$$\begin{aligned} JV_t &= \mathbb{1}_{\{Z_t > \Phi_\alpha\}}(RV_t - BV_t) \\ CV_t &= RV_t - JV_t = \mathbb{1}_{\{Z_t \leq \Phi_\alpha\}} RV_t - \mathbb{1}_{\{Z_t > \Phi_\alpha\}} BV_t \end{aligned} \tag{1.4}$$

where $\mathbb{1}$ is the indicator function and Φ_α is the α quantile of the standard Normal cumulative distribution function.

Corsi et al. (2010) show that the nonparametric estimator BV can be strongly biased in finite sample because of the presence of consecutive jumps and they define a new nonparametric estimator, called Threshold Bipower Variation (TBV). In particular, TBV corrects for the positive bias of BV in the case of consecutive jumps:

$$TBV_t = \mu_1^{-2} \sum_{j=2}^{N(\Delta)} |r_{t,j}||r_{t,j+1)}| \mathbb{1}_{\{|r_{t,j}|^2 < \theta_j\}} \mathbb{1}_{\{|r_{t,j+1}|^2 < \theta_{j+1}\}}$$

where θ is strictly positive random threshold function equal to $\hat{V}_t c_\theta^2$, $\hat{V}_t$ is an auxiliary estimator and c_θ^2 is a scale-free constant that allows to change the threshold. The jump detection test

presented by Corsi et al. (2010) is the following:

$$\texttt{C-Tz} = N(\Delta)^{-1/2} \frac{(RV_t - TBV_t)RV_t^{-1}}{\sqrt{(\frac{\pi^2}{4} + \pi - 5)max\{1, \frac{TTriPV_t}{TBV_t^2}\}}} \longrightarrow_d N(0,1) \tag{1.5}$$

where $TTriPV$ is a quarticity estimator which is obtained by multiplying the TBV by $\mu_{4/3}^{-3}$. Also in this case the jumps and the continuos component are identified and estimated respectively as:

$$\begin{aligned} JV_t &= \mathbb{1}_{\{\texttt{C-Tz}_t > \Phi_\alpha\}}(RV_t - TBV_t) \\ CV_t &= RV_t - JV_t = \mathbb{1}_{\{\texttt{C-Tz}_t \leq \Phi_\alpha\}} RV_t - \mathbb{1}_{\{\texttt{C-Tz}_t > \Phi_\alpha\}} TBV_t \end{aligned} \tag{1.6}$$

The other measure chosen in this work is the Range volatility (RA) presented by Parkinson (1980):

$$RA_t = \frac{1}{4\log 2}(\log(H_t) - \log(L_t))^2 \tag{1.7}$$

This estimator is constructed by taking the highest price (H) and the lowest price (L) for each day as summary of the intraday activity, i.e. the full path process. Its major empirical advantage is that for many assets these informations are ready available. Alizadeh et al. (2002), it is affected by a much lower measurement error than the RV estimator, it is more robust to microstructure noise in a stochastic volatility framework and it allows to extract efficiently latent volatility. On the one hand, the RA estimator contains informations comparable to those embedded in RV. On the other hand, RA is easy to compute also for those assets that are not frequently traded. Indeed, this estimator has advantages typical of both HF data and daily observations.

1.2.2 Forecasting volatility using High Frequency Data

In the literature, there is no consensus if jumps help to forecast volatility. In this sense, this work can be useful in order to understand if allowing for an explicit jump component is important to forecast volatility, independently of the sampling frequency of the price process. Moreover, if different sampling frequencies (daily and 5-minutes) are considered then a discrimination between the two kinds of data used, can be done.

For all forecasting models that I am going to describe in this section, I define a log specification both for inducing normality and for ensuring positivity of volatility forecasts [11]. The natural

[11] Volatility forecasts at each time is obtained by applying the exponential transformation.

starting point in forecasting volatility is to use an Autoregressive (AR) specification[12]. The first model for both RV and RA is the AR model. In particular, an AR(8) model is identified for both RV measure and for Range estimator[13]. The AR specification is easy to implement but it does not capture the volatility long-range dependence due to the slowly decaying autocorrelation of returns. As an alternative, it is possible to use the Heterogenous Autoregressive model proposed by Corsi (2009). This model can be seen as an approximation of long memory model with an important advantage: it is easier to implement than the pure long-memory model (see Andersen et al., 2007, Corsi and Renó, 2009). Indeed, the second forecasting model for both volatility measures is the Heterogeneous Autoregressive model (HAR). The aggregate measures for the daily, weekly and monthly realized volatility are computed as sum of past realized volatilities over different horizons:

$$RV_t^{(N)} = \frac{1}{N} RV_t + \cdots + RV_{t-N+1} \tag{1.8}$$

where N is typically equal to 1, 5 or 22 according to if the time scale is daily, weekly or monthly. Then, HAR-RV becomes:

$$\log RV_{t+h} = \beta_0 + \beta_1 \log RV_{t+h-1} + \beta_2 \log RV_{t+h-1}^{(5)} + \beta_3 \log RV_{t+h-1}^{(22)} + \epsilon_t \tag{1.9}$$

where ϵ_t is IID zero mean and finite variance noise [14]

Moreover, as suggested in Corsi and Renó (2009), the heterogeneous structure applies also to leverage effect. As a consequence, volatility forecasts are obtained by considering asymmetric responses of realized volatility to previous daily, weekly and monthly negative returns. The past aggregated negative returns are constructed as:

$$l_t^{(N)} = \frac{1}{N}(r_t + \cdots + r_{t-N+1}) \mathbb{1}_{\{(r_t+\cdots+r_{t-N+1})<0\}} \tag{1.10}$$

[12]It is also possible to use an ARMA model to forecast volatility in order to consider some measurement errors since the empirical sampling is not done in continuous time.

[13]The identification procedure for the order of both AR models is done by exploiting the sample autocorrelation and the sample partial autocorrelation function, by running both AIC and BIC information criteria and significance of single parameters. Then I check the properties of the residuals: they are normal and the Ljung Box test does not reject the null of no autocorrelation at any significance level.

[14]Corsi and Renó (2009) model the dynamic of the latent quadratic variation, call it $\tilde{\sigma}_t$. Suppose that $\hat{V}_t$ is a generic unbiased estimator of $\tilde{\sigma}_t$ and $log(\tilde{\sigma}_t) = log(\hat{V}_t) + \omega_t$ where ω_t is a zero mean and finite variance measurement error. Then ϵ_t is independent from ω_t.

Then the L-HAR model is defined as:

$$\begin{aligned}\log RV_{t+h} =&\beta_0 + \beta_1 \log RV_{t+h-1} + \beta_2 \log RV^{(5)}_{t+h-1} + \beta_3 \log RV^{(22)}_{t+h-1} + \\ &\beta_4 l_{t+h-1} + \beta_5 l^{(5)}_{t+h-1} + \beta_6 l^{(22)}_{t+h-1} + \epsilon_t\end{aligned} \tag{1.11}$$

The explanatory variables of the HAR-RV model can be decomposed into continuous and jump components, in this way the forecasting model obtained is:

$$\begin{aligned}\log RV_{t+h} =&\beta_0 + \beta_1 \log CV_{t+h-1} + \beta_2 \log CV^{(5)}_{t+h-1} + \beta_3 \log CV^{(22)}_{t+h-1} + \\ &\beta_4 \log(1 + JV_{t+h-1}) + \beta_5 \log(1 + JV^{(5)}_{t+h-1}) + \beta_6 \log(1 + JV^{(22)}_{t+h-1}) + \epsilon_t\end{aligned} \tag{1.12}$$

Depending on how the jump component is detected three different forecasted realized volatility are obtained. First, the HAR-Jumps is obtained according to (1.2) and for the continuous component to (1.1). Second, the HAR-CV-JV model is obtained following Andersen et al. (2007), namely according to (1.4). The last model, HAR-C-J is defined according to (1.6) following the estimation strategy presented in Corsi and Renó (2009).

If a cascade leverage structure is considered as in (1.10) then the forecasting volatility model becomes:

$$\begin{aligned}\log RV_{t+h} =&\beta_0 + \beta_1 \log CV_{t+h-1} + \beta_2 \log CV^{(5)}_{t+h-1} + \beta_3 \log CV^{(22)}_{t+h-1} + \\ &\beta_4 \log(1 + JV_{t+h-1}) + \beta_5 \log(1 + JV^{(5)}_{t+h-1}) + \beta_6 \log(1 + JV^{(22)}_{t+h-1}) + \\ &\beta_7 l_{t+h-1} + \beta_8 l^{(5)}_{t+h-1} + \beta_9 l^{(22)}_{t+h-1} + \epsilon_t\end{aligned} \tag{1.13}$$

As before, according to the estimators used for the volatility components, I obtain the LHAR-Jumps, LHAR-CV-JV and LHAR-C-J models.

In order to asses the forecast ability of the RA, I extend the idea of the heterogeneity in the time horizons of investors in the financial markets and I define two different forecasting models, in addition to the AR(8) model:

$$\log RA_{t+h} = \beta_0 + \beta_1 \log RA_{t+h-1} + \beta_2 \log RA^{(5)}_{t+h-1} + \beta_3 \log RA^{(22)}_{t+h-1} + \epsilon_t \tag{1.14}$$

called Range-HAR and

$$\begin{aligned}\log RA_{t+h} =&\beta_0 + \beta_1 \log RA_{t+h-1} + \beta_2 \log RA^{(5)}_{t+h-1} + \beta_3 \log RA^{(22)}_{t+h-1} + \\ &\beta_4 l_{t+h-1} + \beta_5 l^{(5)}_{t+h-1} + \beta_6 l^{(22)}_{t+h-1} + \epsilon_t\end{aligned} \tag{1.15}$$

called Range-L-HAR.

1.2.3 Forecasting volatility using daily data

The first specification for the continuous volatility component is the GARJI model:

$$
\begin{aligned}
R_t &= \mu + \sigma_t z_t + \sum_{i=1}^{N_t} X_t^{(i)} && (1.16)\\
\lambda_t &= \lambda_0 + \rho\lambda_{t-1} + \gamma\xi_{t-1} && (1.17)\\
\sigma_t^2 &= \gamma + g(\Lambda, \mathcal{F}_{t-1})\epsilon_{t-1}^2 + \beta\sigma_{t-1}^2 && (1.18)\\
g(\Lambda, \mathcal{F}_{t-1}) &= exp(\alpha + \alpha_j E(N_t|\mathcal{F}_{t-1}) && (1.19)\\
&+ \mathbb{1}_{\{\epsilon_{t-1}<0\}}[\alpha_a + \alpha_{a,j} E(N_t|\mathcal{F}_{t-1})])
\end{aligned}
$$

where $\epsilon_t = \epsilon_{1,t} + \epsilon_{2,t} = \sigma_t z_t + \sum_{i=1}^{N_t} X_t^{(i)}$, $z_t \sim \mathcal{N}(0,1)$, $N_t \sim Poisson(\lambda_t)$, $X_t^{(j)} \sim \mathcal{N}(\mu, \omega^2)$ and $\xi_{t-1} = E[N_{t-1}|\mathcal{F}_{t-1}) - \lambda_{t-1}$.

As explained in Maheu and McCurdy (2004), the last equation allows for the introduction of a differential impact if past news are deemed good or bad. If past news are business as usual, in the sense that no jumps occurred, and are positive, then the impact on current volatility will be $exp(\alpha)\epsilon_{t-1}^2$. If no jump takes place but news are bad, the volatility impact becomes $exp(\alpha + \alpha_a)\epsilon_{t-1}^2$. If a jump takes place, with good news, the impact is $exp(\alpha + \alpha_j)\epsilon_{t-1}^2$. If a jump takes place, with bad news, then the impact becomes $exp(\alpha + \alpha_j + \alpha_a + \alpha_{a,j})\epsilon_{t-1}^2$.

The arrival rate of jumps is assumed to follow a non homogeneous Poisson process while jump size is described by a Normal distribution. In this way, the single impact of extraordinary news on volatility is identified through the combination of parameters in $g(\Lambda, \mathcal{F}_{t-1})$. The idea of the authors is the following: the conditional variance of returns is a combination of a smoothly evolving continuous-state GARCH component and a discrete jump component. In addition previous realization of both innovations, $\epsilon_{1,t}$ and $\epsilon_{2,t}$ affect expected volatility through the GARCH component of the conditional variance. This feedback is important because once return innovations are realized, there may be strategic or liquidity tradings related to the propagation of the news which are further sources of volatility clustering[15]. With this model it is possible to allow for several asymmetric responses to past returns innovations and then obtain a richer characterization of volatility dynamics, especially with respect to events in the tail of

[15] A source of jumps to price can be important and unusual news, such as earnings surprise (result as an extreme movement in price) while less extreme movements in price can be due to typical news events, such as liquidity trading and strategic trading.

the distribution (jumps).

In particular $E[N_{t-1}|\mathcal{F}_{t-1})$ is the ex-post assessment of the expected number of jumps that occurred from $t-2$ to $t-1$ and it is equal to $\sum_{j=0}^{\infty} jP(N_{t-1}=j|\mathcal{F}_{t-1})$. Therefore ξ_{t-1} is the change in the econometrician's conditional forecast on N_{t-1} as the information set is updated, it is the difference between the expected value and the actual one. As shown by Maheu and McCurdy (2004) this expression may be inferred using Bayes' formula:

$$P(N_t=j|\mathcal{F}_{t-1}) = \frac{f(R_t|N_t=j,\mathcal{F}_{t-1})P(N_t=j|\mathcal{F}_{t-1})}{f(R_t|\mathcal{F}_{t-1})} \quad \text{for} \quad j=0,1,2,\ldots \tag{1.20}$$

Indeed, conditional on knowing λ_t, σ_t, and the number of jumps that took place over a time interval, $N_t=j$, the density of R_t in terms of observable is Normal:

$$f(R_t|\mathcal{F}_{t-1}) = \sum_{j=0}^{\infty} f(R_t|N_t=j,\mathcal{F}_{t-1}) \times P(N_t=j|\mathcal{F}_{t-1}) \tag{1.21}$$

where

$$f(R_t|N_t=j,\mathcal{F}_{t-1}) = \frac{1}{\sqrt{2\pi(\sigma_t^2+j\delta^2)}} \exp\left(-\frac{(R_t-\mu+\theta\lambda_t-\theta j)^2}{2(\sigma_t^2+j\delta^2)}\right) \tag{1.22}$$

Naturally the likelihood function is defined starting from (1.22), where $\tilde{\theta}$ is the vector of the parameters of interest, i.e. $\tilde{\theta}=(\gamma,\rho,\theta,\delta^2,\alpha,\alpha_j,\alpha_a,\alpha_{aj},\omega,\beta,\lambda_0,\mu)$:

$$\mathcal{L}(R_t|N_t=j,\mathcal{F}_{t-1};\tilde{\theta}) = \prod_{t=1}^{T} f(R_t|N_t=j,\mathcal{F}_{t-1}) \tag{1.23}$$

and the log-likelihood is:

$$l(R_t|N_t=j,\mathcal{F}_{t-1};\tilde{\theta}) = \sum_{t=1}^{T} \log f(R_t|N_t=j,\mathcal{F}_{t-1}) \tag{1.24}$$

The maximum number of jumps in each day in the filter (1.20) is set equal to 10. This is becasue, as suggested in Maheu and McCurdy (2004), the conditional Poisson distribution has almost zero probability in the tails for values of $N_t \geq 10$.

In order to isolate the role of jumps, I estimate a nested version of the GARJI model, i.e. ARJI, which is obtained by imposing $\alpha_j=\alpha_a=\alpha_{a,j}=0$.

In addition, I consider the GARCH-t model and Beta-t-GARCH model for conditional volatility. The aim is to understand if the ARJI model can provide a better fit to the empirical distribu-

tion of the data and a better quantile forecast with respect to volatility specifications based on fat tails, such as *t*-Student. In particular, Beta-*t*-GARCH presents a more sophisticated volatility specification with respect to GARCH-*t* model. The former consists of an observation driven model based on the idea that the specification of the conditional volatility as a linear combination of squared observations is taken for granted but, as a consequence, it responds too much to extreme observations and the effect is slow to dissipate. Harvey and Luati (2014) define a class of models (DCS) in which the observations are generated by a conditional heavy-tailed distribution with time-varying scale parameters and where the dynamics are driven by the score of the conditional distribution. In this way, Beta-*t*-GARCH counts the innovation outliers but also the additive outliers.

1.3 Computing and comparing VaR forecasts

The VaR is defined as the $100\alpha\%$ quantile of the distribution of returns. The probability that the return of a portfolio over a t holding period will fall below the VaR is equal to $100\alpha\%$.
The predicted VaRs are based on the predicted volatility and they depend on the assumption on the conditional density of daily returns. The one day-ahead VaR prediction at time $t+1$ conditional on the information set at time t is:

$$\widehat{VaR}_{t+1|t} = \sqrt{\widehat{\sigma}^2_{t+1|t}} F_t^{-1}(\alpha) \tag{1.25}$$

In (1.25) $\widehat{\sigma}^2_{t+1|t}$ is the returns variance, estimated in both parametric and non-parametric models, $F_t^{-1}(\alpha)$ is the inverse of the cumulative distribution of daily returns while α indicates the degree of significance level. In the case of HF data $\widehat{\sigma}^2_{t+1|t}$ is equal to $\widehat{RV}_t$ or $\widehat{RA}_t$ estimated as explained in Section 1.2.2 while for GARJI model the returns variance is not simply the modified GARCH dynamic but it also consist of the variance due to jumps (Hung et al., 2008):

$$\widehat{VaR}_{t+1|t} = \sqrt{\widehat{\sigma}^2_{t+1|t} + (\widehat{\theta}^2_t + \widehat{\delta}^2_t)\widehat{\lambda}_t} \widetilde{F}_t^{-1}(\alpha) \tag{1.26}$$

where $\widetilde{F}_t^{-1}(\alpha) = F_t^{-1}(\alpha) + \frac{1}{6}((F_t^{-1}(\alpha))^2 - 1)Sk(R_t|t\mathcal{F}_{t-1})$ and $Sk(R_t|t\mathcal{F}_{t-1})$ is the conditional return skewness computed after estimating the model. Once obtained VaR forecasts, I assess the relative performance of the models through the violation[16] rate and the quality of the esti-

[16] In the testing literature exception is used instead of violation because the former is referred, as I explain later, to a loss function. The loss function changes according to the test applied and the motivation behind the testing

mates by applying backtesting methods[17].

A violation occurs when a realized return is greater than the estimated ones (VaR). The violation rate is defined as the total number of violations divided by the total number of one period-forecasts[18]The tests used in this paper are the Unconditional Coverage (L_{UC}) and Conditional Coverage (L_{CC}) tests suggested respectively by Kupiec (1995) and Christoffersen (1998). The $L_{UC}$and L_{CC} are the most popular tests among practitioners and academics. This is because they are very simple to implement and because they are incorporated in the Basel accords requirements [19]. These two motivations represent also the reason why both tests are used also in the academic literature. The L_{UC} and the L_{CC}tests assess the *adequacy* of the model by considering the number of VaR exceptions, i.e. days when returns exceed VaR estimates. If the number of exceptions is less than the selected significance level would indicate, the system overestimates risk; on the contrary too many exceptions signal underestimation of risk. In particular, the first test examines whether the frequency of exceptions over some specified time interval is in line with the selected significance level. A good VaR model produces not only the "correct" amount of exceptions but also exceptions that are independent each other and, in turn, not clustered over time. The test of conditional coverage takes into account for the number of exceptions and when the exceptions occur.

The tick loss function considered is defined as Binary loss function (BLF) which counts the number of exceptions, that are verified when the loss is larger than the forecasted VaR:

$$BLF_{t+1} = \begin{cases} 1 & \text{if} \quad R_{t+1} < \widehat{VaR}_{t+1|t} \\ 0 & \text{if} \quad R_{t+1} \geq \widehat{VaR}_{t+1|t} \end{cases} \tag{1.27}$$

where $\widehat{VaR}_{t+1|t}$ is the estimated VaR at time t that refers to the period $t+1$.

The Likelihood Ratio test of unconditional coverage tests the null hypothesis that the true

strategies.

[17]The backtesting tests give the possibility to interpret the results and then the quality of the forecasting model choose in inferential terms.

[18]As well explained in Gençay et al. (2003) at$q$ th quantile, the model predictions are expected to underpredict the realized return $\alpha = (1-q)$ percent of the time. A high number of exceptions implies that the model excessively underestimates the realized return. If the exception ratio at the q *th* quantile is greater than α percent, this implies excessive underprediction of the realized return. If the number of exceptions is less than α percent at the q *th* quantile, there is excessive overprediction of the realized return by the underlying model.

[19]See **?** for a review on VaR forecasting and evaluation through backtesting.

probability of occurrence of an exception over a given period is equal to α:

$$H_0: \quad p = \alpha$$
$$H_1: \quad p \neq \alpha$$

where $\widehat{p} = \frac{n_0}{n_1+n_0}$ is the unconditional coverage (the empirical coverage rate) or the failure rate and n_0 and n_1 denote, respectively, the number of exceptions observed in the sample size and the number of non-exceptions.
The unconditional test statistic is given by:

$$LR_{UC} = -2\log\left(\frac{(1-\alpha)^{n_1}\alpha^{n_0}}{(1-\widehat{p})^{n_1}\widehat{p}^{n_0}}\right) \sim \chi^2(1) \tag{1.28}$$

So, under the null hypothesis the significance level used to forecast VaRs and the empirical coverage rate are equal. The test of conditional coverage proposed by Christoffersen (1998) is an extended version of the previous one taking into consideration whether the probability of an exception on any day depends on the exception occurrence in the previous day. The loss function in constructed as in (1.27) and the log-likelihood testing framework is as in (1.28) including a separate statistic for the independence of exceptions. Define the number of days when outcome j occurs given that outcome i occurred on the previous day as n_{ij} and the probability of observing an exception conditional on outcome i of the previous day as π_i. Summarizing:

$$\pi_0 = \frac{n_{01}}{n_{00}+n_{01}} \qquad \pi_1 = \frac{n_{11}}{n_{10}+n_{11}} \qquad \pi = \frac{n_{01}+n_{11}}{n_{00}+n_{01}+n_{10}+n_{11}} \tag{1.29}$$

The independence test statistic is given by:

$$LR_{IND} = -2\log\left(\frac{(1-\pi)^{n_{00}+n_{10}}\pi^{n_{01}+n_{11}}}{(1-\pi_0)^{n_{00}}\pi_0^{n_{01}}(1-\pi_1)^{n_{10}}\pi_1^{n_{11}}}\right) \tag{1.30}$$

Under the null hypothesis the first two probabilities in (1.29) are equal, i.e. the exceptions do not occur in cluster. Summing the statistics (1.28) and (1.30) the conditional coverage statistic is obtained, i.e. $LR_{CC} = LR_{UC} + LR_{IND}$ and it is distributed as a χ^2 with two degrees of freedom since two is the number of possible outcomes in the sequence in (1.27). In order to avoid the possibility that the models considered passing the joint test but fail either the coverage or the independence test I choose to run LR_{CC} and also its decomposition in LR_{UC} and LR_{IND}.

1.4 Data and Empirical results

1.4.1 Data

In order to assess the informational content of HF and daily data, I use S&P 500 index from 5 Jan.1996 to 30 Dec.2005 for both samples.

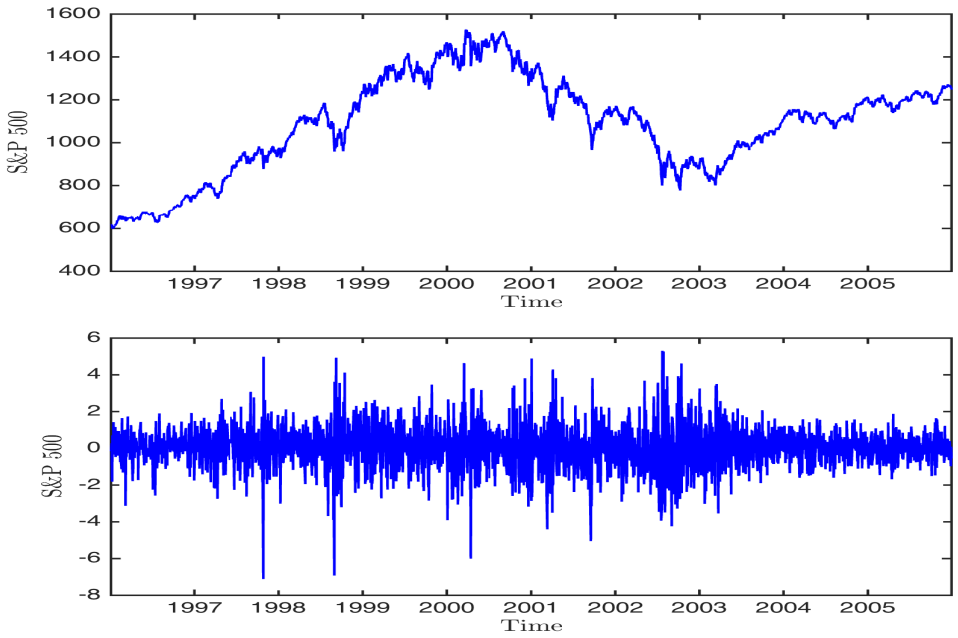

Figure 1.1: Top: daily S&P 500 index from 5 Jan.1996 to 30 Dec.2005. The horizontal axis corresponds to time while the vertical axis displays the value of the index. Bottom: daily S&P 500 percentage returns calculated by $r_t = log(p_t/p_{t-1})$, where p_t is the value of the index at time t.

The total number of trading days is equal to 2516 which coincides with the number of daily returns. In the top panel of Figure 1.1 the level of the S&P 500 index is presented. The corresponding daily returns are displayed in the bottom panel of Figure 1.1. Given the literature on the effects of microstructure noise of estimates of RV and the forecast performance of RV models based on different sampling frequency, I use 5-minutes data for a total of 197,689 observations. I compute 5-minutes intraday returns as the log-difference of the closing prices in two subsequent periods of time. The daily returns are computed taking the last closing prices in each trading day. The range volatility at each date is calculated as scaled log difference between the highest and the lowest price in a trading day. Table 1.1 reports the descriptive statistics of S&P 500 index for RA_t, RV_t and its decomposition in BV_t and JV_t. In particular

Table 1.1: Summary Statistics. The rows report the sample mean, standard deviation, skewnwss, kurtosis, sample minimum and maximum for the daily returns (R_t), the standardized daily returns ($R_t/\sqrt{RV_t}$) the daily realized volatility (RV_t), the daily bipower variation (BV_t), the daily jump component (JV_t) and the daily range estimator (RA_t). Returns are expressed in percentage.

	R_t	$R_t/\sqrt{RV_t}$	RV_t	BV_t	JV_t	RA_t
Mean	0.0279	0.1378	0.8250	7.93E-05	3.15E-06	9.70E-05
St. Dev.	1.1520	1.3138	1.0097	9.85E-05	1.00E-05	1.52E-04
Skewness	-0.0951	0.253	4.8721	4.8786	19.9283	7.1671
Kurtosis	5.9165	2.8505	39.1013	39.3401	659.3967	84.1687
Min	-7.1127	-3.6092	0.0281	0.0281	0	0.0206
Max	5.3080	4.7161	11.890	11.890	3.6200	25.931

JV_t is computed as $max\{RV_t - BV_t, 0\}$[20]. A number of interesting features are founded. Firstly, returns exhibit negative asymmetry and leptokurtosis. As shown in Andersen et al. (2007) the daily returns standardized with respect to the square root of the ex-post realized volatility are closed to Gaussian. In fact its mean and asymmetry are close to zero, its variance is close to one while its kurtosis is near to 3. This result is clear from Figure 1.2 in which the empirical density distribution is plotted with the normal density distribution for $R_t/\sqrt{RV_t}$. Moreover if I compare RV_t and BV_t the latter is less noisy than the former, considering the role of jumps. Finally, jump process does not show any Gaussian feature [21].

Figure 1.3 shows the plot of RV_t, BV_t, JV_t and RA_t estimators. It is evident that RV_t, BV_t and JV_t follow a similar pattern and the latter tends to be higher when RV_t is higher. JV_t exhibits a relatively small degree of persistence as consequence of the clustering effect. Not surprisingly, RA_t follows the same pattern of RV_t since both of them are ex-post volatility measures.

Estimation results based on daily data

Table 1.2, provides parameter estimates for both the GARJI and ARJI model applied to the S&P500. The parameter estimates are presented separating the diffusion component from the jump component. First, both parameters ρ and γ are significantly different from zero. The former represents the persistence of the arrival process of jumps that is quite high for both models

[20] The summary statistics of the continuous and disontinuous components computed according to Andersen et al. (2007) and Corsi et al. (2010) are not reported because are very similar to those presented in Table 1.1.

[21] In particular, jumps computed according to (1.6) exhibit a higher mean with respect to those computed according to (1.4), given that the former exploits the possibility of consecutive jumps.

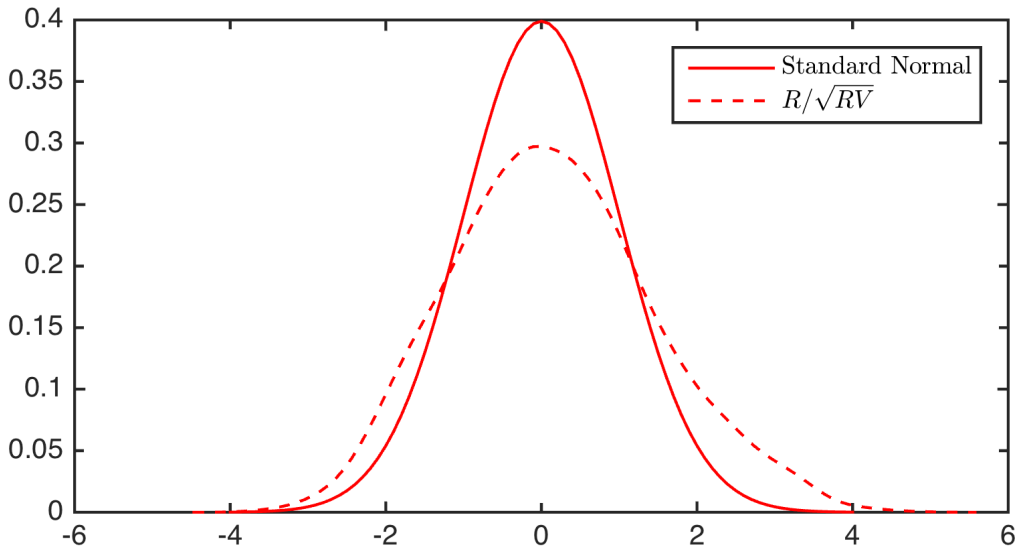

Figure 1.2: Standardized log-returns distribution of the S&P 500 index.The standard normal distribution (solid line) is compared with the standardized log-returns distribution (dashed line).

implying the presence of jump clustering. The latter, γ, measures the change in the conditional forecast of the number of jumps due to the last day information. The significance of these two parameters suggests that the arrival process of jumps can deviate from its unconditional mean. The implied unconditional jump intensity is 0.8727 while the average variance due to jumps is equal to 0.5516: the index is volatile. This result is confirmed by the average proportion of conditional variance explained by jumps which is equal to 0.3068, jumps explained almost the 23% of the total returns variance. Moreover the jump size mean θ is negative for both model and the most interesting feature is that it affects conditional skewness and conditional kurtosis. The sign of θ indicates that large negative return realizations due to jumps are associated with an immediate increase in the variance explaining the contemporaneous leverage effect: when jumps are realized they tend to have a negative effect on returns. In particular the average conditional skewness is equal to -0.2766 while the average conditional kurtosis is equal to 3.2814. Furthermore the feedback coefficient $g(\Lambda, \mathcal{F}_{t-1})$ tends to be smaller when at least one jump occurs because the total innovation is larger after jumps . Considering the first column of Table 1.2, the feedback coefficient associated with good news and no jump is equal to 0.0005 and it increases if one jump occurs, i.e. 0.0010. If no jumps occur and if news are bad the coefficient is equal to 0.0411; it is equal to 0.0348 in case of bad news if one jump occurs. These results provide evidence for the asymmetric effect of good and bad news and

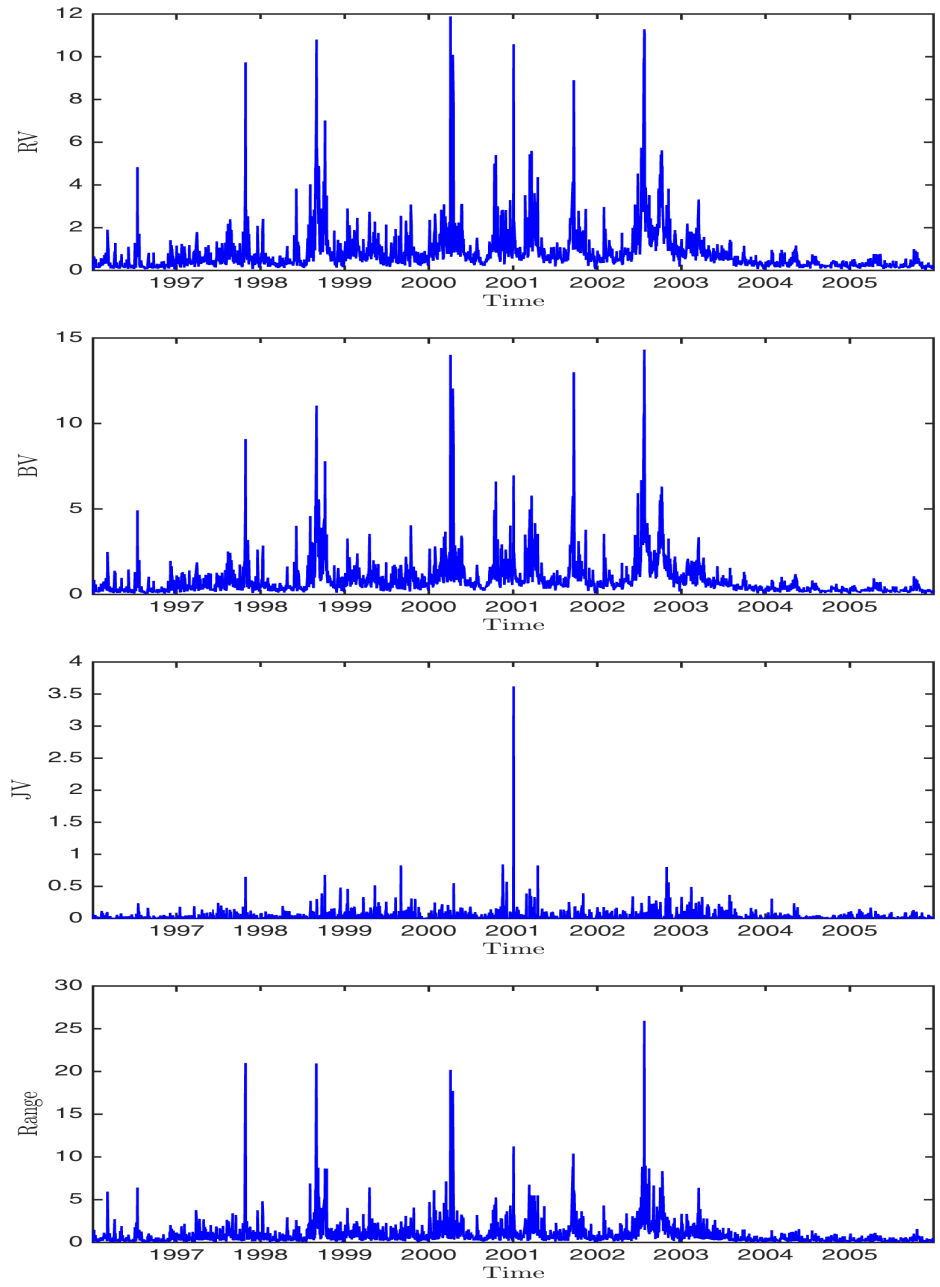

Figure 1.3: Top: RV_t computed using 5- minutes data from 5 Jan.1996 to 30 Dec.2005. Second: BV_t computed using 5- minutes data from 5 Jan.1996 to 30 Dec.2005. Third: $JV_t = max\{RV_t - BV_t, 0\}$ is computed using 5- minutes data from 5 Jan.1996 to 30 Dec.2005. Bottom: Range estimator computed using daily data from 5 Jan.1996 to 30 Dec.2005. Time is on the horizontal axis.

they show that the asymmetry associated to bad news is more important in the absence of jumps, namely for normal innovations. In fact the difference between the coefficient estimates for both good and bad news in the case of no jumps and one jump are quite similar. This means that news associated with jump innovations is incorporated more quickly into current

prices. The second column of Table 1.2 presents the estimated parameters for the model with $\alpha_j = \alpha_a = \alpha_{a,j} = 0$. With this specification and through the LR test it is possible to understand if the asymmetric effect of good versus bad news is statistically significant: the asymmetric news effect is statistically significant.

Table 1.2: GARJI and ARJI models estimates. ARJI model is obtained assuming $\alpha_j = \alpha_a = \alpha_{a,j} = 0$. Standard errors are in parenthesis.

Process	Parameters	S&P 500	
		GARJI	ARJI
Diffusion			
	μ	0.0106 (1.9839)	0.0153 (2.1142)
	ω	0.0036 (0.0005)	0.0034 (0.0005)
	α	-7.7048 (0.4332)	-4.7623 (0.3063)
	α_j	0.8096 (0.7538)	-
	α_a	4.5131 (0.4213)	-
	$\alpha_{a,j}$	-0.9776 (0.7204)	-
	β	0.9696 (0.0002)	0.9787 (0.0000)
Jump			
	λ_0	0.0211 (0.0039)	0.0229 (0.0052)
	ρ	0.9758 (0.0025)	0.9757 (0.0030)
	γ	0.5262 (0.0501)	0.4792 (0.0641)
	θ	-0.9895 (0.3985)	-0.9793 (0.4501)
	δ^2	0.0005 (0.0000)	0.0000 (0.0000)
Log-likelihood		-3570.8	-3574.4

Estimation results based on high frequency data

All the estimates presented in Table 1.3, Table 1.4 and Table 1.5 are computed employing the OLS method over the entire sample period, i.e. from 5 Jan. 1996 to 30 Dec. 2005, for the S&P500 index. Table 1.3 and Table 1.4 show the results for the models presented in Section 1.2.2 for models based on RV, its decomposition in BV and JV and the cascade structure for the leverage effect.

The coefficients of the continuous component expressed as daily, weekly and monthly measures, respectively β_1,β_2 and β_3 are significants in all models. Moreover, jump components appear to be fundamental to forecast one step ahead volatility; the predictive power is larger for those specifications that allow for RV decomposed in its continuous and discontinuous components, regardless the identified method used for jump magnitude. Furthermore, the estimates

Table 1.3: Estimation of models based on high frequency data: AR(8), HAR, L-HAR, HARC-Jumps, LHARC-Jumps. The coefficients refer to models presented in Section 1.2.2. Standard errors are in parenthesis.

Parameter	AR(8)		HAR		L-HAR		HARC-Jumps		LHARC-Jumps	
β_0	-0,0615	(0,0130)	-0,1086	(0,0125)	-0,3916	(0,0474)	-0,0670	(0,0260)	-0,3658	(0,0552)
β_1	0,3877	(0,0195)	0,4020	(0,0189)	0,2763	(0,0199)	0,4047	(0,0189)	0,2799	(0,0199)
β_2	0,1670	(0,0211)	0,3537	(0,0317)	0,3257	(0,0336)	0,3452	(0,0318)	0,3218	(0,0335)
β_3	0,0693	(0,0212)	0,1735	(0,0314)	0,2029	(0,0372)	0,1691	(0,0331)	0,1891	(0,0387)
β_4	0,0902	(0,0213)	-	-	-0,2374	(0,0152)	-0,0886	(0,1554)	0,0009	(0,1477)
β_5	0,0831	(0,0213)	-	-	-0,1962	(0,0422)	-0,0828	(0,3146)	-0,1692	(0,2987)
β_6	0,0332	(0,0213)	-	-	-0,0840	(0,0907)	0,1147	(0,6055)	0,4391	(0,5776)
β_7	0,0567	(0,0210)	-	-	-	-	-	-	-0,2349	(0,0152)
β_8	0,0316	(0,0195)	-	-	-	-	-	-	-0,1914	(0,0422)
β_9	-	-	-	-	-	-	-	-	-0,0862	(0,0911)
Obs.	2494		2494		2494		2494		2494	
R^2	0,6463		0,6444		0,6744		0,6459		0,6752	
Adj. R^2	0,6452		0,6439		0,6736		0,6450		0,6740	

for the aggregate leverage variables are negatives (as expected) and significant. Moreover, the predictive power increases adding the cascade structure for the leverage regressors.

Table 1.4: Estimation of models based on high frequency data: HAR-CV-JV,LHAR-CV-JV, HAR-C-J, LHAR-C-J. The coefficients refer to models presented in Section 1.2.2. Standard errors are in parenthesis.

Parameter	HAR-CV-JV		LHAR-CV-JV		HAR-C-J		LHAR-C-J	
β_0	-0,0666	(0,0244)	-0,3671	(0,0549)	-0,0482	(0,0182)	-0,3225	(0,0512)
β_1	0,4040	(0,0188)	0,2792	(0,0198)	0,4085	(0,0185)	0,2906	(0,0195)
β_2	0,3450	(0,0317)	0,3219	(0,0334)	0,3115	(0,0310)	0,2942	(0,0327)
β_3	0,1711	(0,0327)	0,1902	(0,0385)	0,1942	(0,0313)	0,2153	(0,0371)
β_4	-0,0938	(0,1545)	0,0059	(0,1468)	-0,1345	(0,1004)	-0,1242	(0,0954)
β_5	-0,0676	(0,3115)	-0,1511	(0,2958)	0,4000	(0,1733)	0,2607	(0,1649)
β_6	0,0435	(0,5948)	0,4172	(0,5686)	-0,4836	(0,2960)	-0,2292	(0,2827)
β_7	-	-	-0,2351	(0,0152)	-	-	-0,2310	(0,0151)
β_8	-	-	-0,1909	(0,0422)	-	-	-0,1888	(0,0419)
β_9	-	-	-0,0869	(0,0913)	-	-	-0,0604	(0,0908)
Obs.	2494		2494		2494		2494	
R^2	0,6458		0,6752		0,6497		0,6781	
Adj. R^2	0,6450		0,6740		0,6489		0,6769	

This finding confirms the different reaction of daily volatility to negative returns. The estimates of the forecasting models based on the Range estimator are reported in Table 1.5. The coefficients of the HAR specification are statistically significant; these results imply a heterogeneous structure also for RA volatility measure. The highest predictive power is recorded for the L-HAR model. Indeed also, in this case, the heterogenous structure in the leverage effect has an important role in predicting future volatility.

Table 1.5: Estimation of models based on Range estimator: AR(8), HAR, L-HAR. The coefficients refer to models presented in Section 1.2.2. Standard errors are in parenthesis.

Parameter	AR(8)		HAR		L-HAR	
β_0	-0,0937	(0,0212)	-0,2694	(0,0194)	-0,8699	(0,0724)
β_1	0,1094	(0,0204)	0,0993	(0,0205)	-0,0455	(0,0224)
β_2	0,2054	(0,0205)	0,4580	(0,0422)	0,2970	(0,0484)
β_3	0,1212	(0,0209)	0,3255	(0,0463)	0,3486	(0,0566)
β_4	0,0918	(0,0209)	-	-	-0,3069	(0,0260)
β_5	0,1002	(0,0209)	-	-	-0,4060	(0,0707)
β_6	0,0791	(0,0209)	-	-	-0,4148	(0,1479)
β_7	0,0573	(0,0205)	-	-	-	-
β_8	0,0938	(0,0204)	-	-	-	-
β_9	-	-	-	-	-	-
Obs.	2494		2494		2494	
$R\hat{2}$	0,3964		0,3914		0,4337	
Adj. $R\hat{2}$	0,3945		0,3907		0,4323	

1.4.2 VaR accuracy results

To assess the model's capability of predicting future volatility, I report the results of the Kupiec (1995) and the Christoffersen (1998) tests described in Section 1.3. Both tests address the accuracy of VaR models and their results interpretation give insights into volatility models usefulness to risk managers and supervisory authorities. The tests are computed for both models based on HF data and on daily data. In evaluating models performance, the available sample is divided into two subsamples. The in-sample period is equal to 1677 observation, around 2/3 of the total sample, while the out-of-sample period is around 1/3 of the total sample, equal to 839 observations. A rolling window procedure is used to implement the backtesting procedure and, in turn, to choose among different specifications. After estimating the alternative

VaR models, the one-day-ahead VaR estimate is computing using the in-sample period. Then the in-sample period is moved forward by one period and the estimation is run again. This procedure is repeated step by step for the remaining 839 days, until the end of the sample. For both tests the expected number of exceedances is chosen equal to 5% and 1% level[22]. Table 1.6 and Table 1.7 shows the VaR accuracy results at both 5% and 1% level, respectively, for all models presented in Section 1.2.2 and Section 1.2.3. The economic value per se of the HF-data forecasting models is assessed looking at the first part of Table 1.6 and Table 1.7: All models that allow for explicit jumps and leverage components do not reject the null at 1% while LHAR-C-J (jumps specified according to Corsi et al., 2010) is the only model that does not reject the null conditional coverage at both αs level. In fact, for this model, the average number of violations for the VaR at 5% level is the closest to the true probability of occurrence of an exception over one day.

Instead, looking at the accuracy of daily data models, GARCH-*t* and Beta-*t*-GARCH do not reject the null of conditional coverage at 5%, while all models pass the L_{CC} test at 1% level. Comparing this last result with the accuracy of the LHAR-C-J model, both GARCH-*t* and Beta-*t*-GARCH provide an average number of violations closer to the theoretical one. AR(8) provides accurate VaR measures if the Range estimator is used to proxy the latent volatility. Even if the statistical significance of all βs parameters in both Range HAR and L-HAR models give insight on the possibility to extend the heterogeneous structure to such forecasting models (see Table 1.5), these models do not pass accuracy tests at both considered level.

Indeed, the VaR forecasts according to both L_{UC} and L_{CC} are more accurate for daily data than HF-data models.

Furthermore, allowing for an explicit jump component improves over HF-based VaR performance at 1% level. No matter what the jump identification strategy is chosen, all models (HARC-Jumps, HAR-CV-JV, and HAR-C-J) do not reject the null of unconditional and conditional coverage at 1% significance level. At odds, the null is rejected for VaR computed at 5% level. For what concerns daily-data models, accounting for an explicit jump component (GARJI, ARJI) or supposing a fat-tails distribution for log-returns gives the same VaR accuracy at 1% in terms of L_{UC} and L_{CC}. Allowing for an explicit jump factor in the conditional log-returns distribution provides more accurate VaR measure, in addition to important infor-

[22]Both tests are also implemented to 10% level and the results are shown in the A.1. The quantile required by Basel accords is 1%. Financial institutions, recently, has implemented stress tests which require VaR forecasts for level smaller than 1%.

Table 1.6: VaR accurancy at 5%. The first column shows the model chosen in order to compute the VaR forecasts. H is the average number of violations computed for each model. VaR is the average VaR forecasts. LR_{UC}, LR_{CC} and LR_{IND} represent the pvalue associated to the Kupiec (1995) and Christoffersen (1998) tests. All tests are evaluated at 1% significance level.

Model	H	VaR	LR_{UC}	LR_{IND}	LR_{CC}
AR(8)	0.074	-1.276	0.003	0.493	0.009
HAR	0.075	-1.267	0.002	0.134	0.002
L-HAR	0.073	-1.268	0.004	0.096	0.004
HARC-Jumps	0.076	-1.261	0.001	0.597	0.004
LHARC-Jumps	0.074	-1.266	0.003	0.114	0.003
HAR-CV-JV	0.076	-1.260	0.001	0.597	0.004
LHAR-CV-JV	0.074	-1.265	0.003	0.114	0.003
HAR-C-J	0.075	-1.257	0.002	0.544	0.006
LHAR-C-J	0.072	-1.261	0.007	0.399	0.018
GARJI	0.029	-1.646	0.002	0.030	0.001
ARGJI	0.024	-1.673	0.000	0.087	0.000
GARCH-t	0.032	-2.221	0.012	0.279	0.023
Beta-t-GARCH	0.032	-2.195	0.012	0.279	0.023
Range AR(8)	0.070	-1.255	0.010	0.664	0.034
Range HAR	0.075	-1.233	0.002	0.055	0.001
Range L-HAR	0.098	-1.135	0.000	0.453	0.000

mation about the market response to outside news.

Another focus of this paper is represented by the leverage effect. Looking at Table 1.6 and Table 1.7 , leverage effect has an important role in improving volatility forecasts and, in turn, VaR accuracy. In fact, at least for daily data and HF-data, models that allow for an asymmetric volatility response to price movements, do not reject the null of conditional coverage, passing both Kupiec (1995) and Christoffersen (1998) tests at 1% level. Surprisingly, L-HAR model at 1% level generates the same proportion of hits (13%) of LHAR-CV-JV and LHARC-Jumps, involving an equal value for the L_{CC} statistic. This means that adding jumps as an explanatory variable in the forecasting volatility model does not improve over VaR accuracy if a leverage component is considered.

A slightly different result is registered for the HAR-C-J and the L-HAR-C-J models, underlying a superior ability of jump identification strategy proposed by Corsi et al. (2010).

Table 1.7: VaR accurancy at 1%. The first column shows the model chosen in order to compute the VaR forecasts. H is the average number of violations computed for each model. VaR is the average VaR forecasts. LR_{UC}, LR_{CC} and LR_{IND} represent the pvalue associated to the Kupiec (1995) and Christoffersen (1998) tests. All tests are evaluated at 1% significance level.

Model	H	VaR	LR_{UC}	LR_{IND}	LR_{CC}
AR(8)	0.017	-1.805	0.075	0.490	0.162
HAR	0.019	-1.791	0.019	0.034	0.007
L-HAR	0.013	-1.794	0.385	0.129	0.217
HARC-Jumps	0.018	-1.783	0.039	0.025	0.010
LHARC-Jumps	0.013	-1.790	0.385	0.129	0.217
HAR-CV-JV	0.018	-1.782	0.039	0.025	0.010
LHAR-CV-JV	0.013	-1.790	0.385	0.129	0.217
HAR-C-J	0.017	-1.778	0.075	0.019	0.013
LHAR-C-J	0.015	-1.783	0.138	0.191	0.142
GARJI	0.004	-2.423	0.031	0.883	0.098
ARGJI	0.002	-2.486	0.008	0.922	0.029
GARCH-t	0.008	-3.449	0.622	0.731	0.835
Beta-t-GARCH	0.010	-3.409	0.894	0.695	0.918
Range AR(8)	0.020	-1.775	0.009	0.350	0.021
Range HAR	0.023	-1.744	0.002	0.070	0.001
Range L-HAR	0.029	-1.606	0.000	0.180	0.000

Summing up, daily-data models are preferred to HF-data models when the VaR is required at 5% level[23]. At 1% VaR level, all daily data models pass the Kupiec (1995) and the Christoffersen (1998) tests, at odd of HF-data models. For this data category, only the more sophisticated volatility forecasting models give accurate VaR forecasts. Finally, both jumps and leverage effect are important factors in order to obtain reliable VaR measures.

1.5 Conclusion

This paper assesses the economic value of different forecasting volatility models, in terms of informational content embedded in the HF observations and daily data. In order to do so,

[23]From Table A.1 in the A.1 only the AR(8) model passes all accuracy tests. This result can be interpreted in favor of more sophisticated forecasting models when the α level required is less conservative.

I compare the performance of HF-data and daily data models in a VaR framework. Two key assumptions are introduced: jumps in price and leverage effect in volatility dynamics.
Specifically, I consider various specifications of HF-data models for volatility forecast, which differs along three main dimensions: different time-horizons for investors, separation of continuous and discontinuous volatility components and, finally, a cascade dynamic for the leverage effect. I also consider different variants of the daily data models, in form of GARJI models either with or without an asymmetric effect of news on volatility, as well as in form of two fat-tails models, namely the GARCH-*t* and the Beta-*t* GARCH models. All these models are compared with a correspondent and equivalent model, based on the Range volatility measure; the latter is expected to estimate a level of volatility which is intermediate with respect to those measured by HF-data and daily data models. This analysis highlights important issues. First, it stresses the importance of the sampling frequency for data needed in economic applications such as the VaR measurement. Second, it emphasizes the strict relationship between VaR measures and the type of model used to forecast volatility. In sum, daily-data models are preferred to HF-data models at 5% and 1% VaR level.
The accuracy of the VaR measure significantly improves when introducing both an explicit jump component and a fat-tails distribution in forecasting volatility models. Specifically, independently from the data frequency, allowing for jumps in price (or providing fat-tails) and leverage effects translates in more accurate VaR measure. However, introducing jumps allows risk managers to have relevant information on the market reaction to outside news.

Appendix A

Additional Results of Chapter 1

A.1 VaR accurancy at 10% level

Table A.1: VaR accurancy at 10% level. The first column shows the model chosen in order to compute the VaR forecasts. H is the average number of violations computed for each model. VaR is the average VaR forecasts. LR_{UC}, LR_{CC} and LR_{IND} represent the pvalue associated to the Kupiec (1995) and Christoffersen (1998) tests. All tests are evaluated at 1% significance level.

Model	H	VaR	LR_{UC}	LR_{IND}	LR_{CC}
AR(8)	0.132	-1.276	0.003	0.670	0.010
HAR	0.139	-1.267	0.000	0.817	0.001
L-HAR	0.136	-1.268	0.001	0.881	0.004
HARC-Jumps	0.142	-1.261	0.000	0.945	0.001
LHARC-Jumps	0.138	-1.266	0.000	0.987	0.002
HAR-CV-JV	0.142	-1.260	0.000	0.726	0.001
LHAR-CV-JV	0.137	-1.265	0.001	0.949	0.003
HAR-C-J	0.147	-1.257	0.000	0.979	0.000
LHAR-C-J	0.141	-1.261	0.000	0.913	0.001
GARJI	0.070	-1.646	0.003	0.165	0.004
ARGJI	0.068	-1.673	0.001	0.279	0.003
GARCH-t	0.052	-2.221	0.000	0.029	0.000
Beta-t-GARCH	0.052	-2.195	0.000	0.029	0.000
Range AR(8)	0.143	-1.255	0.000	0.559	0.000
Range HAR	0.145	-1.233	0.000	0.621	0.000
Range L-HAR	0.166	-1.135	0.000	0.439	0.000

Chapter 2

Option Pricing with High Frequency Estimates of Continuous and Discontinuous Volatility Components

Building on Majewski et al. (2015), I propose an affine discrete-time model based on high frequency measure which is characterized by a multi-factor volatility specification, labeled VARG-J model. The latent volatility process is modeled as the sum of two independent factors: a diffusive and a jump one. The former involves small changes while the latter allows volatility to experience periods of extreme movements. The estimation under historical measure is done via Extended Kalman Filter. This strategy allows to filter out both volatility factors introducing a measurement equation that relates the Realized Volatility (RV) to latent volatility. In this way it is possible to take into account three RV-related problems: microstructure noise, measurement errors and the overnight effect. The change of measure is performed adopting an exponentially affine stochastic discount factor which preserves all the analytical results in order to obtain closed-form option pricing formula. An empirical analysis of S&P 500 index options show that this model makes out-of-the-money and in-the-money options more expensive proving superior ability to capture the volatility smile with respect to some benchmark models. A superior performance is registered also for at-the-money options.

JEL-Classification: C13, G12, G13

Keywords: Volatility Jumps, ARG-Zero, Realized Volatility, High Frequency, Option Pricing

2.1 Introduction

The specification of the volatility[1] dynamics of the underlying asset represents a key point in option pricing. Since the stock market crash on October 19, 1987, extraordinary deviations of stock index option prices from the benchmark Black-Scholes (BS henceforth) model have been registered. According to the BS assumptions, all option prices on the same underlying security with different expiration date and different exercise prices should have the same implied volatility. On the contrary, looking at financial data, implied volatilities computed on traded options change across various strike prices, displaying a U-shaped curve [2].

Attempts in the financial literature to reconcile the theory with the data have mostly centered around two approaches: jump diffusion model (see for example Merton, 1976,Kou, 2002) and stochastic volatility model (see for example Duan et al., 1995, Heston, 1993 and Heston and Nandi, 2000). In the former stock prices are allowed to jump, i.e. the data generating process for asset prices features discontinuities in its trajectories. Asset prices, being observable quantities, signal the presence and magnitude of jumps, providing a simplified estimation for such model. On the contrary, the estimation of stochastic volatility model is more challenging since volatility is latent and unobservable.

Recently, the literature has mainly focused on models that incorporate both stochastic volatility and jumps in returns since they are able to generate fat tails in the return distribution. Bakshi et al. (1997) compare competing option pricing models in order to understand how each model characteristic improves option pricing. They find that a model with stochastic volatility and random jumps in price represents a significant alternative to BS model but it requires highly implausible levels of volatility variation. Bates (2000) examines two possible explanations for the deviations implicit in option prices since the crash of 1987. The two competing hypothesis are stochastic volatility that incorporates leverage effects (i.e. time-varying volatility inversely related to market returns) and stochastic volatility with jumps in price (i.e. crash fear). Bates (2000) presents strong evidence against the model with both stochastic volatility and jumps since, even if the square root diffusion hypothesis guarantees many desirable features, it cannot account for the large implicit volatility shocks observed in the S&P 500 options market after

[1] I use the term volatility for both measures of variance and standard deviation. It is clear from the contest to which I refer to.

[2] It is demonstrated that excess of kurtosis (skewness) in the underlying asset return distribution is the main source of volatility smile (smirk) in option prices. Allowing for fatter tails makes extreme observations more likely than the value predicted by the BS model. The consequence is that the value of away-from-the-money options increases relative to at-the-money options.

the 1987 crash. Pan (2002) sheds some light on how various types of uncertainty, i.e. volatility and jump risks, are priced in the financial market. The author underlines the stochastic volatility with jumps in price misspecification: the volatility of volatility cannot increase fast enough during volatile markets. Therefore, both stochastic volatility and jumps in price do not generate the level of kurtosis (skewness) implied by the volatility smile (smirk) observed in the data.

Adding discontinuities (or jumps) in the stochastic volatility process, represents a possible solution to this issue, providing the so-called "double-jump" specification. Duffie et al. (2000) provide jumps in volatility as the explanation for the high volatility of volatility in the stochastic volatility model with jumps only in price. Moreover, the authors suggest that jumps in volatility may attenuate the overpricing of out-of-the-money call options. Eraker et al. (2003) provide empirical evidence supporting the presence and the importance of jumps in volatility. Jumps in volatility reduce the misspecification of stochastic volatility model with jumps in price and they play an important role in generating rapid bursts in volatility. Furthermore, models with jumps in volatility result in a significant increase in implied volatility for deep-in-the-money or deep-out-of-money options. Besides the cross-sectional impact on option prices, an important issue is related to the type of changes through which the volatility evolves over time.

Traditionally, stochastic volatility models have assumed that the spot variance is continuous. Todorov and Tauchen (2011) make nonparametric inference regarding the activity level of stock market volatility and find that market volatility is a very vibrant process, i.e. it involves many small changes as well as occasional big moves. The presence of big moves justifies the use of jumps in volatility modeling. More recenty, Caporin et al. (2015) focus on the estimation of the volatility jump component in a discrete time setting in order to understand if the introduction of discontinuities in the volatility process could be an interesting way for describing rapid and large volatility increments. They find a positive probability of jumps in volatility and, especially when the level of volatility is high, the jump component represents a relevant part of the estimated conditional volatility.

Thanks to the availability of HF data, Christensen et al. (2014) examine the price evolution at the finest tick-by-tick resolution and show that the traditional measures of jump variation (Andersen et al., 2007, Corsi et al., 2010) tend to erroneously assign a burst of volatility to jumps in price. In particular, jumps in price account for a very small proportion, about 1%, of the quadratic price variation. Moreover, Christensen et al. (2014) provide a detailed selection of

literature results reporting estimates of the jump variation component, expressed as a fraction of total return variation. The results can be summarized as follows: the jump component decreases in magnitude as the sampling frequency increases. In tick-by-tick data, most of the jumps identified at lower frequency vanish and what is left are high volatility episodes. The consequence is that what a stochastic volatility model with jumps in price identifies as genuine price jumps are periods of heightened volatility.

I contribute to this strand of the financial econometrics literature by proposing an affine discrete-time model based on a HF volatility measure, labeled as VARG-J, which is characterized by a multifactor volatility specification. In the VARG-J model volatility experiences periods of extreme movements through a jump factor modeled as an Autoregressive Gamma Zero (ARG_0) process. Surprisingly, to the best of my knowledge, this paper is the first attempt at introducing discontinuities in the volatility dynamics in an option pricing framework.

This work is mostly related to the class of Realized Volatility (RV) option pricing models, introduced by Corsi et al. (2013) . Corsi et al. (2013) propose the HARGL-RV model in which the RV dynamics follows the HAR process by Corsi (2009) with a daily binary variable that accounts for leverage effect. Majewski et al. (2015) introduce a heterogeneous parabolic structure for leverage in the HARGL-RV, defining the LHARG-RV model. More recently, Alitab et al. (2015) design a more advanced version, labeled as J-LHARG, where the volatility is positively affected by price jumps. All these RV-based models focus on continuous time stochastic volatility and assume that volatility is perfectly observable through an intra-day volatility measure. From an econometric perspective, the introduction of volatility jumps requires an additional latent state variable.

In this work, the latent volatility is equal to the sum of two independent random variables which account for continuous and discontinuous factors of volatility. The continuous component is modeled as an ARG process with an autoregressive structure for the non-central parameter. The jump factor is modeled as an ARG_0 process obtained by setting to zero the shape parameter. Monfort et al. (2014b) build this process in order to explain the zero lower bound (ZLB) on bond yield which has become a common situation in several countries after the 2008 financial crisis. The point-mass at zero that characterizes the conditional distribution given the past of the Zero-ARG process can be exploited in order to describe the dynamics of the volatility jump component.

In order to obtain the complete specification of the VARG-J model under the $\mathbb{P}$ measure, I derive the analytical formula of the Moment Generating Function (MGF) of the log-returns. Fol-

lowing the approach of Majewski et al. (2015), the change of measure is performed adopting a stochastic discount factor (SDF) which incorporates risk premia associated with the sources of risk concerning the two volatility factors, beyond the risk related to shocks in returns. Thanks to the SDF exponential affine-form, I prove that the risk neutral ($\mathbb{Q}$) dynamics is from the same family as the historical $\mathbb{P}$ dynamics, namely the $\mathbb{Q}$ model is still a VARG-J model. Furthermore, thanks to the recursive expression for the MGF, there is a one-to-one mapping between the set of parameters under $\mathbb{P}$ and the set of parameters under $\mathbb{Q}$. A great advantage of VARG-J is represented by the model estimation. I estimate the $\mathbb{P}$ parameters via pseudo maximum likelihood with the Extended Kalman Filter. This estimation procedure allows to filter out both volatility factors, by exploring information on the latent volatility through the realized volatility estimator and, at the same time, by controlling for measurement errors and overnight effects. To be more precise, I assume that the total volatility is approssimated by the ex-post RV volatility estimator (by Andersen and Bollerslev, 1998), computed at 5 minutes frequency, as suggested by Liu et al. (2015) but I introduce a measurement equation for the total volatility in order to correct the RV bias due to measurement errors and overnight effects. The analysis under $\mathbb{P}$ is done using spot price of S&P 500 index sampled at 5 minutes frequency from 4 Jan. 1996 to 30 Dec. 2005. I empirically assess the option pricing performance of the model using Plain Vanilla Call options written on S&P500 Index for the same time span. The results clearly illustrates the important contribution of the jump factor in the pricing performance of the S&P500 Index options and the economic significance of the volatility jump risk premia.
The paper is organized as follows. Section 2.2 introduces the model under both the historical and the risk-neutral probability measure. Section 2.3 discusses the estimation under the historical measure and the competing models. Section 2.4 discusses the calibration under the risk-neutral measure and the pricing performance. Section 2.5 concludes.

2.2 The VARG-J Model

2.2.1 The asset retun process

Given the result reported in Christensen et al. (2014), I assume a price process free of jumps. Specifically, I consider the following discrete-time stochastic volatility model for daily log-returns:

$$y_{t+1} := \log\left(\frac{S_{t+1}}{S_t}\right) = r^y_{t+1} + \lambda^y f_{t+1} + z_{t+1} \tag{2.1}$$

where r^y_{t+1} denotes the risk-free rate at time $t+1$, assumed to be exogenous, and where λ^y is the market price of risk. Moreover, the latent factor f_{t+1} denotes the true volatility and the innovation, z_{t+1}, represents a heteroskedastic Gaussian Innovation:

$$\begin{aligned} z_{t+1} &= \sqrt{f_{t+1}}\epsilon^y_{t+1} \\ \epsilon^y_{t+1} &\sim \mathcal{IID}\mathcal{N}(0,1) \end{aligned} \tag{2.2}$$

This framework allows volatility to experience periods of extreme movements in addition to the diffusive ones, thanks to the multifactor volatility specification. More precisely, I assume that the true volatility is given by the sum of two independent factors[3]:

$$f_{t+1} = f_{1,t+1} + f_{2,t+1}. \tag{2.3}$$

In (2.3) $f_{1,t+1}$ is the diffusive or continuous volatility factor while $f_{2,t+1}$ is a burst factor of the volatility dynamics (henceforth volatility jump). Given the information set at time t, denoted $\mathcal{F}_t$, the former follows an Extended Autoregressive Gamma (EARG) process:

$$f_{1,t+1}|\mathcal{F}_t \sim \gamma_v(\theta_{1t}, \mu_1) \tag{2.4}$$

This process is coherent with strictly positive volatitiliy values, it can describe the dynamics of volatility process and is very tractable from an analytical point of view. The volatility jump component, $f_{2,t+1}$, conditionally on $\mathcal{F}_t$ and $f_{1,t+1}$, evolves according to an Autoregressive Gamma Zero (ARG_0) process:

$$f_{2,t+1}|f_{1,t+1}, \mathcal{F}_t \sim \gamma_0(\theta_{2t}, \mu_2) \tag{2.5}$$

This process is, instead, consistent with non-negative volatility jumps and is able to accomodate extended periods of zero or close-to-zero values. These features make the ARG_0 a suitable process to model the volatility jump factor. Furthermore, the ARG_0 is a particular case of the Extended ARG of Monfort et al. (2014a) and it is obtained by extending it to a zero shape parameters. Past information enters both volatility factor dynamics, through the non-central

[3] In this paper, I focus my attention on the application of the simplest version of the VARG-J model: both non-central parameters have an autoregressive structure of order one. I plan to relax this specification and allow for a richer characterization of volatility in future work.

parameters:

$$\boldsymbol{\theta}_t = \begin{pmatrix} \theta_{1t} \\ \theta_{2t} \end{pmatrix} = \begin{pmatrix} d_1 + \beta_1 f_{1,t} \\ d_2 + \beta_2 f_{2,t} \end{pmatrix} \tag{2.6}$$

The dynamics in Equation (2.1) differs from that employed by Corsi et al. (2013) and Majewski et al. (2015) since I do not assume that the spot volatility (the total f_t) coincides with some realized measure, i.e. an estimator of asset return volatility constructed using high frequency data. In particular, realized volatility aims at estimating the quadratic variation (or the integrated variance in the case of jumps in price) of the price process over a some time interval. Consistently with the absence of jumps in price, the quadratic variation of the log-price process over period $t+1$ is:

$$\begin{aligned} QV_{t+1} &= p \lim_{n\to\infty} \sum_{j=1}^{n} r^2_{t+\frac{j}{n}} \\ r_{t+\frac{j}{n}} &= p_{t+\frac{j}{n}} - p_{t+\frac{j-1}{n}} \end{aligned} \tag{2.7}$$

where$p_{t+j/n}$ is the log-price at the end of the jth interval on day t and n is the number of the observations available at day t recorded at frequency $\frac{1}{n}$. The simplest realized variance estimator is the empirical analog of QV, which is:

$$RV_{t+1} = \sum_{j=1}^{n} r^2_{t+\frac{j}{n}}. \tag{2.8}$$

Realized measures are used by practitioners because they contain information on the true latent volatility process. These estimators, however, are affected by microstructure noise, measurement errors and overnight effects.

The first contaminates the price measurement and is due, among others, to price discreteness and rounding, to the properties of the trading mechanism, to the existence of bid-ask spreads and to data recording mistakes. The consequence is that microstructure noise induces serial autocorrelation in the observed returns, which biases the realized measure computed at high frequency. The literature offers a wide range of alternative realized measures in order to capture the microstructure noise effect on high frequency returns. Liu et al. (2015) compare a large collection of realized measures across a range of assets and conclude that it is very rare to find a realized measure that significantly outperforms 5-minutes RV estimator. Moreover the authors, in a panel investigation, explain differences in performace, in terms of microstructure noise and market conditons and find that the gains from using a more sophisticated realized

measure rather that RV 5-minutes fall when the microstructure noise is large and in periods of high volatility.

In addiction to microstructure noise, RV estimators are affected by a discretization error due to the unavailability of a continuous price record and a downward bias due to the unavailability of overnight intraday returns.

Therefore, in this paper, I introduce the following measurement equation which relates the observed 5-minutes RV (hencefoth RV) to the latent process f_t:

$$RV_{t+1} = \eta_0 + \eta_1 f_{t+1} + \epsilon^{rv}_{t+1}, \qquad \epsilon^{rv}_{t+1} \sim \mathcal{IID}\mathcal{N}(0, \sigma^2) \tag{2.9}$$

where ϵ^{rv}_{t+1} is independent from ϵ^{y}_{t+1} in equation (2.1). According to equation (2.9), ϵ^{rv}_{t+1} models the measurement erros in RV while η_1 corrects the RV bias due to the overnight effect.

Equations (2.1) and (2.4)-(2.5) completely characterise the Vector Autoregressive Gamma with Jumps (VARG-J) model. A great advantage of the VARGJ model is that it satisfies the affine property. The following propositions are directly derived from the theoretical results presented in Majewski et al. (2015) and are proposed according to the same scheme.

Proposition 1. *Under* $\mathbb{P}$*, the MGF for VARG-J model has the following form*

$$\varphi^{\mathbb{P}}_{t,T,z} = \mathbb{E}^{\mathbb{P}}[e^{z y_{t,T}} | \mathcal{F}_t] = exp\left(a_t + \mathbf{b}_t' \mathbf{f}_t\right) \tag{2.10}$$

where

$$a_s = a_{s+1} + z r_s - \nu W_{1,s+1} + d_1 V_{1,s+1} + d_2 V_{2,s+1} \tag{2.11}$$

$$\mathbf{b}_s' = (V_{1,s+1}, V_{2,s+1})\boldsymbol{\beta} \tag{2.12}$$

$$\tag{2.13}$$

with

$$x_{h,s+1} = x_h(z, \mathbf{b}_{s+1}) = \mathbf{b}_{s+1}' + z\lambda + \frac{z^2}{2}, \quad h = 1, 2$$

subject to the initial conditions:

$$a_T = 0, \quad \mathbf{b}_T' = 0$$

The functions V and W are defined as follows:

$$\begin{aligned} V_{h,s+1} &= V_h(x_{h,s+1}, \mu_h) = \frac{x_{h,s+1}\mu_h}{1 - x_{h,s+1}\mu_h}, \quad h = 1, 2 \\ W_{1,s+1} &= W_1(x_{1,s+1}, \mu_1) = \log\left[1 - x_{1,s+1}\mu_1\right] \end{aligned}$$

Proof. *See B.1*

The parameters under historical measure ($\mathbb{P}$) are given by

$$\boldsymbol{\psi} = [\lambda, \nu, \mu_1, \mu_2, d_1, d_2, \beta_1, \beta_2, \sigma^2] \tag{2.14}$$

Apart from λ, all of them are assumed to be nonnegative. The estimation startegy used to estimate the vector $\boldsymbol{\psi}$ is presented in the Section 2.3.1

2.2.2 Risk-neutral dynamics

In the previous section, I characterized the behaviour of the underlying asset under the $\mathbb{P}$ measure. In this section, I introduce an assumption on the Stochastic Discount Factor (SDF) that allows to transform the distribution under $\mathbb{P}$ to a distribution under the risk neutral ($\mathbb{Q}$) probability measure and, therefore, to compute option prices. In specifying the SDF, I follow Majewski et al. (2015):[4]

$$M_{t,t+1} = \frac{\exp(-\delta_2 y_{t+1} - \delta_{11} f_{1,t+1} - \delta_{12} f_{2,t+1})}{\mathbb{E}^{\mathbb{P}}[\exp(-\delta_2 y_{t+1} - \delta_{11} f_{1,t+1} - \delta_{12} f_{2,t+1})|\mathcal{F}_t]} \tag{2.15}$$

This SDF is a very flexible specification since it identifies two risk premia, i.e. δ_{11} and δ_{12} in addition to the usual equity premium, i.e. δ_2.

More precisely, δ_{11} compensates for the continuous volatility while δ_{12} compensates for the discontinuous source of risk[5]. Also in this Section, the propositions are directly derived from the theoretical results presented in Majewski et al. (2015) and are proposed according to the same scheme.

[4] Corsi et al. (2013) introduce a SDF involving both the log-return and Realized Volatility, applying a modified version of the standar discrete-time exponential affine SDF applied in Gourieroux and Monfort (2007). Majewski et al. (2015) present a more general and flexible version.

[5] Many authors (see Gagliardini et al., 2011,Christoffersen et al., 2013, Corsi et al., 2013 and Majewski et al., 2015) recognized the importance of variance-dependent risk premia in SDF in reconciling the time series properties of asset returns with the cross-section of option prices.

Proposition 2. *Under the model specification in (2.1) and (2.4)-(2.5) with the SDF specified as in (2.15), the VARG-J satisfies the no-arbitrage condition if and only if*

$$\delta_2 = \lambda + \frac{1}{2} \tag{2.16}$$

Proof. *See B.2*

Given the result in Preposition 2 and the market incompleteness, δ_{11} and δ_{12} are free parameters to be calibrated while δ_2 is considered as fixed.
The SDF in (2.15) belongs to the family of the exponential-affine factors. Thanks to this characteristic, it is possible to compute analogous recursion under $\mathbb{Q}$ and provide a one-to-one mapping between the set of parameters under $\mathbb{P}$ and the set of parameters under $\mathbb{Q}$.

Proposition 3. *Under the risk-neutral measure $\mathbb{Q}$ the latent volatility still follows a VARG-J process with parameters*

$$d_1^{\mathbb{Q}} = \frac{d_1}{1-y_1^* \mu_1} \quad d_2^{\mathbb{Q}} = \frac{d_2}{1-y_2^* \mu_2} \quad \beta_1^{\mathbb{Q}} = \frac{\beta_1}{1-y_1^* \mu_1}$$

$$\beta_2^{\mathbb{Q}} = \frac{\beta_2}{1-y_2^* \mu_2} \quad \mu_1^{\mathbb{Q}} = \frac{\mu_1}{1-y_1^* \mu_1} \quad \mu_2^{\mathbb{Q}} = \frac{\mu_2}{1-y_2^* \mu_2}$$

$$\nu^{\mathbb{Q}} = \nu$$

where $y_h^* = -\delta_{1h} - \delta_2 \lambda + \frac{\delta_2^2}{2}$ for $h = 1, 2$.

Proof. *B.3*

The result proved in Proposition 3 and the analytical tractability of the VARG-J process simplify the computation of the risk-neutral MGF. The latter is obtained starting from the MGF under $\mathbb{P}$ and substituting the parameters under $\mathbb{P}$ with those under $\mathbb{Q}$.

Corollary 4. *Under $\mathbb{Q}$ the MGF for the VARG-J model has the same form as in (1) with equity risk premium*
$\lambda^{\mathbb{Q}} = -\frac{1}{2}$ and $\mathbf{d}^{\mathbb{Q}}, \boldsymbol{\beta}^{\mathbb{Q}}, \mu_1^{\mathbb{Q}}, \mu_2^{\mathbb{Q}}, \mu^{\mathbb{Q}}$ as in (2.17).

Therefore, f_t is still a VARG-J process under $\mathbb{Q}$ measure and the two risk premia δ_{11}andδ_{12} are the only parameters to be calibrated on option prices, as explained in Section 2.4.2. Once the value of δ_{11}andδ_{12} are calibrated, all the parameters can be computed in closed-form following the Corollary 4.

2.3 Estimation under $\mathbb{P}$ and statistical properties

2.3.1 VARG-J estimation methodology

The parameters under $\mathbb{P}$ are estimated via pseudo-maximum likelihood with the Extended Kalman filter. The two transition equations can be specified using the first two conditional moments:

$$\begin{aligned} E(f_{jt+1}|\mathcal{F}_t) &= \mu_j(d_j + \beta_j f_{jt}) + \mu_j \nu_j \qquad \text{for} \quad j = 1,2 \\ V(f_{jt+1}|\mathcal{F}_t) &= 2\mu_j^2(d_j + \beta_j f_{jt})) + \mu_j^2 \nu_j \qquad \text{for} \quad j = 1,2 \end{aligned} \tag{2.17}$$

Clearly, ν_2 is equal to zero since f_{2t} is distributed as a γ_0 as in equation (2.5). Both volatility factors are stationary if $\mu_j \beta_j < 1$. To this purpose, I formulate the VARG-J model in state space form. For the measurement equations, I consider two type of observables: the return and the RV. The first measurement equation is directly obtained from the daily log-returns dynamics. The second measurement equation relates the observed RV to the latent f_t according to equation (2.9). Specifically, the transition equations are directly derived from Equation (2.17):

$$\begin{aligned} \begin{pmatrix} f_{1,t+1} \\ f_{2,t+1} \end{pmatrix} &= \begin{pmatrix} \mu_1(\nu + d_1) \\ \mu_2 d_2 \end{pmatrix} + \begin{pmatrix} \mu_1\beta_1 & 0 \\ 0 & \mu_2\beta_2 \end{pmatrix} \begin{pmatrix} f_{1,t} \\ f_{2,t} \end{pmatrix} + \begin{pmatrix} v_{t+1}^1 \\ v_{t+1}^2 \end{pmatrix} \\ \mathbf{Q}_t &= \begin{pmatrix} \sqrt{\mu_1^2(\nu + 2d_1 + 2\beta_1 f_{1,t})} & 0 \\ 0 & \sqrt{\mu_2^2(2d_2 + 2\beta_2 f_{2,t})} \end{pmatrix} \end{aligned} \tag{2.18}$$

while the two measurement equations read:

$$\begin{aligned} \begin{pmatrix} y_t \\ RV_t \end{pmatrix} &= \begin{pmatrix} r_t^y \\ \eta_0 \end{pmatrix} + \begin{pmatrix} \lambda^y & \lambda^y \\ \eta_1 & \eta_1 \end{pmatrix} \begin{pmatrix} f_{1,t} \\ f_{2,t} \end{pmatrix} + \begin{pmatrix} w_t^1 \\ w_t^2 \end{pmatrix} \\ \mathbf{R}_t &= \begin{pmatrix} \sqrt{f_{1,t} + f_{2,t}} & 0 \\ 0 & \sigma \end{pmatrix} \end{aligned} \tag{2.19}$$

To estimate the model, I use second order pseudo-maximum likelihood with the Extended Kalman filter, since both latent factors and the returns in the first measurement equation are conditionally heteroskedastic. In order to do so, the first two conditional moments are correctly specified and $\mathbf{v}_{t+1}$ is approximated by a bivariate uncorrelated Gaussian white noise. In particular, the true log-likelihoods derived from the conditional noncentral Gamma distributions are replaced by the log-likelihoods obtained from Gaussian distributions.

2.3.2 Alternative models

The VARG-J is a discrete time model in which the volatility dynamics is given by the sum of two independent factors, i.e. diffusive and jump. The first natural alternative is represented by the ARG model. It is obtained by setting the volatility jump factor equal to zero, namely the true volatility involves a sequence of small increments, and by assuming that the unobservable total variation of the log-price process is estimated by some realized volatility measure. Specifically the discrete-time stochastic volatility model for daily log-returns is:

$$\begin{aligned} y_{t+1} &= r_{t+1} + \lambda f_{t+1} + \sqrt{f_{t+1}}\epsilon_{t+1}, \quad \epsilon_{t+1} \sim \mathcal{IID}\mathcal{N}(0,1) \\ f_{t+1} &= RV_{t+1} \end{aligned} \tag{2.20}$$

The RV measure is modeled as the autoregressive gamma process of order one by Gouriéroux and Jasiak (2006):

$$RV_{t+1}|\mathcal{F}_t \sim \gamma_\omega(\alpha + \beta_d RV_t, \theta) \tag{2.21}$$

The introduction of the discontinuous factor in the volatility dynamics generates additional persistence in the volatility process which plays an important role in explaining long-term part of the inplied volatility (IV) surface. Thus, the second natural alternative is the HARG model by Corsi et al. (2013). In this model the volatility jump factor is set equal to zero and at each point in time volatility is perfectly observable. The difference with the VARG-J model lies in the assumption on the volatility measure. In particular, the conditional mean of the RV is modeled according the HAR specification of Corsi (2009) and RV_{t+1} follows a noncentral gamma transition distribution:

$$RV_{t+1}|\mathcal{F}_t \sim \gamma_\omega(\alpha + \beta_d RV_t^d + \beta_w RV_t^w + \beta_m RV_t^m, \theta) \tag{2.22}$$

where

$$\begin{aligned} RV_t^d &= RV_t \\ RV_t^w &= \frac{1}{4}\sum_{i=1}^{4} RV_{t-i} \\ RV_t^m &= \frac{1}{17}\sum_{i=5}^{21} RV_{t-i} \end{aligned}$$

are the short-term (daily), medium-term (weekly) and the long-term volatility factors, respec-

tively[6].

A comparison between the ARG and the VARG-J models illustrates the importance of a multifactor specification, i.e. the addition of a volatility jump factor, in stochastic volatility option pricing models. This comparison is possibly unfair since the latent volatility is not perfectly observed in the VARG-J, at odds with the ARG model. Majewski et al. (2015) concludes that the use of RV helps in pricing short term options, especially at-the-money. On the contrary other factors, as the persistence in volatility, have a prominent impact on pricing long-term options. Therefore, looking at the pricing performance on long-term options far from at-the-money should give insights on the role of jumps in volatility process. Additional intuitions on the jump volatility factor can be pbtained by comparing the VARG-J with the HARG model. Both of them should improve on the pricing of long-term options with differences in terms of moneyness. The reason is that both HARG and VARG-J models generate persistent volatility dynamics but through different mechanisms. The former allows memory persistence by using an heterogeneous structure for the RV conditional mean. In the latter model the jump volatility factor, generates persistence and bursts in the volatility process. This last feature should have consequences on pricing deep-in-the-money and deep-out-of-the-money options. In this case the comparison can be considered more balanced since, as Corsi et al. (2013) point out, the heterogeneous structure helps in smoothing the noise affecting the RV measure. Thanks to the use of the RV as a proxy for the latent volatility process, the parameters of both ARG and HARG models are stimated via the Maximum Likelihood Estimator (MLE) on intra-day historical data, using the available formula of the conditional transition density provided by Gouriéroux and Jasiak (2006). As shown in Corsi et al. (2013), the market price of risk λ in the log return equation is estimated rewriting the equation (2.1) as:

$$\frac{y_{t+1} - r_{t+1}}{\sqrt{RV_{t+1}}} = \lambda \sqrt{RV_{t+1}} + \epsilon_{t+1} \quad \text{with} \quad \epsilon_{t+1} \sim \mathcal{IID}\mathcal{N}(0,1) \tag{2.23}$$

and applying the ordinary least squares (OLS).

2.3.3 Results

In this Section I report the estimation results for the VARG-J model and for both competiting models[7]. Daily return and RV time series are obtained from spot prices of the S&P

[6]Clearly, the ARG model can be nested in the HARG model by setting $\beta_w = \beta_m = 0$.

[7]For identification purpose the diffusive-intercept coefficients (α and d_1) are imposed equal to zero in all models.

500 index sampled at 5 minutes frequency for a total of 2517 daily observations from January 5,1996 to December 30, 2005. Since the daily risk free rate is assumed to be exogenous, I proxy it with the FED Fund rate for the same time period of log-returns.

Table 2.1 shows the estimated parameters, and the relative standard errors for the VARGJ

Table 2.1: Estimates of parameter under the historical measure and standard errors (in parenthesis) for the VARG-J, ARG and HARG model. The parameters reported in the first column are estimated via pseudo maximum likelihood with Extended Kalman filter of order two. The parameter of both ARG and HARG models are estimated using Maximum Likelihood. The historical data for all models are daily RV computed on 5-minutes data for the S&P 500 from January 5,1996 to December 30, 2005.

Parameter	VARG-J	Parameter	ARG	HARG
λ	-0,0010 (0,0002)	λ	2,1368 (2,8952)	2,1368 (2,8952)
μ_1	2,93E-05 (4,41E-06)	θ	1,69E-05 (1,46E-09)	1,40E-05 (2,34E-09)
β_1	2,74E+04 (5,10E+03)	β_d	3,99E+04 (87,608)	3,06E+04 (1,53E+03)
μ_2	0,0002 (0,0001)	β_w	-	1,72E+04 (2,15E+03)
ν	0,8085 (0,1358)	β_m	-	6,91E+03 (1,86E+03)
d_2	5,29E-08 (1,40E-07)	ω	1,5875 (0,0084)	1,3757 (0,0270)
β_2	4,19E+03 1,47E+03			
σ	5.22e-06 (3.76e-06)			
η_0	7.05e-06 (1.14e-06)			
η_1	0,4898 (0,0514)			
Persistence f_1	0,8021	Log-likelihood	-14997	-14772
Persistence f_2	0,9431	Persistence	0,6746	0,7674

model and the alternative models presented in the Section 2.3.2. Table 2.1 also reports the log-likelihood values for only the last family of models in order to compare their performance According to the estimates, all VARG-J coefficients are statistically significant with the only

exception of d_2. It is important to note that d_2 must be different from zero preventing the possibility that $f_{2t} = 0$ becomes an absorbing state. In fact, if d_2 is equal to zero once the jump component reaches zero, it remains at zero for all the next periods. From Table 2.1, the parameter η_1 is smaller than 1 and significant. This means that the approach presented in this paper allows to adjust the downward bias due to overnight effect.

The sensitivity of $f_{1,t}$ on the conditional mean of $f_{1,t}$ is smaller than that of $f_{2,t}$ on its conditional mean, which means that the jump factor is more persistent than the diffusive one. The ARG and HARG models' coefficients are all significant. The parameters of the HARG model show a decreasing impact of the past lags on the present value of the RV, in line with the literature.

An interesting difference is in the λ estimates across the two models categories [8]. For the VARG-J model λ is negative and significant, implying that the distribution of returns is negatively skewed. In this paper the conditional distribution of returns is non- Gaussian since the volatility process is not observable. The negative sign for λ in the VARG-J model, hence, is coherent with the argument of a risk premium related to skewness and negative tails events (leverage effect) rather than to volatility per se. In fact, investors fear large negative drops of their wealth induced by neagtive events, which contribute to the negative skewness of the distribution.

The conditional distribution of log-returns is Gaussian in the (H)ARG framework, since RV is observed at each point in time. For these models the risk premium is positive and not significant, leading to a sligthly different interpretation. Namely, from Equation(2.23), more volatilile stocks are more profitable. As said before, the Extended Kalman Filter estimation strategy allows to filter out the time series of both volatility factors. Figure 2.1 shows the updated[9] values estimated by the Extended Kalman Filter procedure. From this figure f_{1t} is the diffusive volatility factor that describes the asset's idyosincratic changes while $f_{2,t}$ represents the discontinuous component which provides bursts in the volatility process.

The Extended Kalman Filter procedure allows to take into account measurement errors, i.e. microstructure noise and overnight effects, in the RV measure. Figure 2.2 shows the RV and the updated values for f_t. The VARG-J modelbetter fits volatility changes, measured *ex-post* by RV.

[8] Note that the risk premium estimate is the same for both ARG and HARG models since it is estimated using the same regression formula on the same data.

[9] At each point in time the current value of the state variable (volatility) is updated on the basis of the observation of the returns and RV.

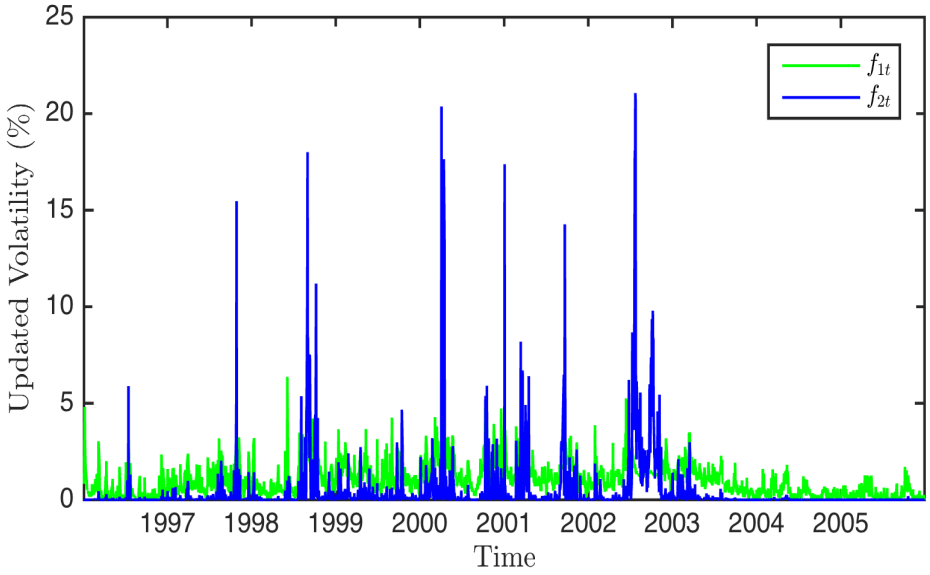

Figure 2.1: This figure shows the updated time series of the diffusive ($f_{1,t}$, green line) and the jump ($f_{2,t}$, blue line) factors which are obtained applying the estimation procedure based on the Extended Kalman filter. The sample consists of S&P 500 index data from January 5, 1996 to December 30, 2005.

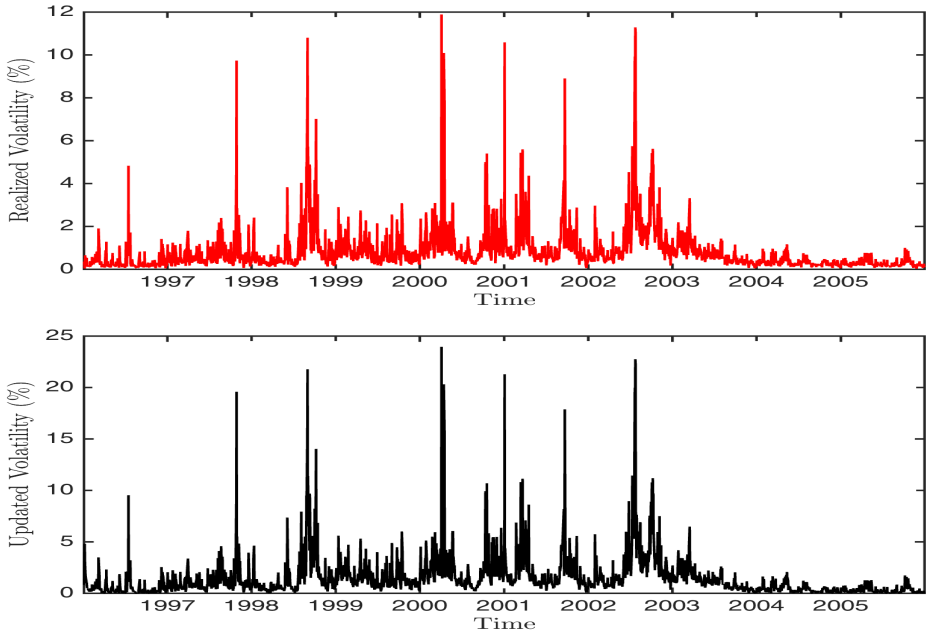

Figure 2.2: This figure shows the comparison between RV_t (top figure) and the updated time series of f_t (bottom figure). The sample consists of S&P 500 index data from January 5, 1996 to December 30, 2005.

2.4 Empirical analysis under $\mathbb{Q}$

2.4.1 Data and stylized facts

The option dataset consists of daily European call options written on the S&P 500 index. The observations for the option prices range from January 5, 1996 to December 30, 2005. I only use Wednesday options data[10] yielding to a total of 30061 observations. As it is customary in the literature (see Barone-Adesi et al., 2008, Corsi et al., 2013,Majewski et al., 2015), I filter out options with time to maturity less than 10 days or more than 360 days, implied volatility larger than 70% or prices less that 0.05. To perform the analysis, I split the options into different categories according to time to maturity and moneyness. Moneyness (m) is defined as S_t/K, i.e. the underlying index level divided by the option strike price. A call option is said to be deep out-of-the-money (DOTM hereafter and therein) if m$\leq$ 0.94, OTM if 0.94 $<$ m$\leq$ 0.97, ATM if 0.97 $<$m$\leq$ 1.03, DITM if 1.03 $<$m $\leq$ 1.06 and ITM if m $>$ 1.06. Using time to maturity (T), options are classified in four categories: short maturity if T$\leq$50, medium-short maturity if 50 $<$ T $\leq$ 90, medium-long if 90 $<$ T $\leq$ 160, and long maturity if T $>$ 160. Table 2.2 reports descriptive statistics for the options classified by the moneyness and maturity definitions given above. From Panel A, the DOTM call options are heavly traded: they represent an insurance against an increase of the underlying price. From Panel B and Panel C, it is possible to observe that the data show a volatility smile/smirk and, in particular the volatility smirk is pronunced for long maturity options. Hence, ITM (DITM) calls are more expensive compared to OTM (DOTM) calls. This implies that the risk-neutral distribution of the log-return is far from Gaussian, suggesting that a richer characterization of the volatility process is needed to explain these features both qualitatively and quantitatively. In this paper I try to assess this feature by increasing the excess kurtosis in the log-return distribution.

2.4.2 The calibration of risk premia

Given the parameters under $\mathbb{P}$ measure obtained via the procedure described in the Section 2.3.1, the risk premia parameters in (2.15) need to be calibrated in order to derive the risk-neutral dynamics. Specifically, δ_2 is determined by the no-arbitrage condition in the Proposition 2, and δ_{11} and δ_{12} are calibrated on observed option prices. The purpose of the calibration

[10]The first motivation for using Wednesday data is that Wednesday is the day of the week least likely to be a holiday. Therefore, it is less likely than other days to be affected day-of-the-week effects (see Bakshi et al., 1997, Christoffersen et al., 2008). Moreover this procedure is customary in the literature.

Table 2.2: Summary statistics for the S&P 500 index option data. The observations refer to each Wednesday during the January 5, 1996 to December 30, 2005. Panel A shows the number of option contracts sorted by moneyness and maturity. Panel B shows the average option prices sorted by moneyness and maturity. Panel C shows the average implied volatilities sorted by moneyness and maturity. Implied volatilities are calculated using the Black & Scholes formula. T refers to the number of days to maturity while m represents the moneyness defined as the underlying index level divided by the option strike price.

Moneyness	$T \leq 50$	$50 < T \leq 90$	$90 < T \leq 160$	$T > 160$	All
Panel A: Number of Contracts					
$m \leq 0.94$	2391	2294	1496	2650	8831
$0.94 < m \leq 0.97$	2246	1393	595	656	4890
$0.97 < m \leq 1$	2671	1732	637	809	5849
$1 < m \leq 1.03$	2169	1160	475	633	4437
$1.03 < m \leq 1.06$	1240	546	274	285	2345
$m > 1.06$	1631	853	602	623	3709
All	12348	7978	4079	5656	30061
Panel B: Average Option Prices					
$m \leq 0.94$	0.9188	3.2835	6.9051	17.044	7.3858
$0.94 < m \leq 0.97$	3.8515	11.450	23.346	46.767	14.145
$0.97 < m \leq 1$	13.009	24.920	39.597	65.467	26.687
$1 < m \leq 1.03$	29.895	42.033	55.952	81.544	43.226
$1.03 < m \leq 1.06$	54.066	65.985	78.103	103.23	65.625
$m > 1.06$	131.48	148.11	169.64	195.13	152.19
All	31.740	34.817	48.920	58.595	39.940
Panel C: Average Implied Volatility					
$m \leq 0.94$	0.1726	0.1555	0.1498	0.1474	0.1568
$0.94 < m \leq 0.97$	0.1378	0.1410	0.1490	0.1556	0.1425
$0.97 < m \leq 1$	0.1454	0.1503	0.1593	0.1611	0.1505
$1 < m \leq 1.03$	0.1594	0.1619	0.1644	0.1664	0.1616
$1.03 < m \leq 1.06$	0.1800	0.1785	0.1758	0.1735	0.1784
$m > 1.06$	0.2762	0.2312	0.2199	0.2034	0.2445
All	0.1725	0.1624	0.1650	0.1599	0.1664

is the selection of risk premia such that the model implied unconditional volatility under the risk-neutral measure matches the unconditional risk-neutral volatility. Since it is not possible to directly observe the latter, I follow the same strategy used in Corsi et al. (2013): the market-observed IV is used as an instrument to be matched with the model-generated IV, since both depend on the volatility under Q measure. The two risk premia, δ_{11} and δ_{12}, are calibrated by

minimizing the loss function which measures the distance between the model generated IV for two options and the market IV corresponding to the same two options in the sample[11]:

$$f(\delta_{11}, \delta_{12}) = \sqrt{\frac{1}{2}\left[(IV_1^{mkt} - IV_1^{mod})^2 + (IV_2^{mkt} - IV_2^{mod})^2\right]} \tag{2.24}$$

where $IV_{i,t}^{mkt}$ is the market IV of the option i, $IV_{i,t}^{mod}$ is the IV computed from the model, for $i = 1, 2$. In order to deal with the risk premia calibration for the VARG-J model, I randomly select two ATM options from the entire sample, observed in two different days. For the competitor models the motivation behind the calibration is the same but, since the risk premium to be calibrated is only δ_{11}, I select only one option from the sample. To avoid the problem of a possible unfair comparison among models, I randomly choose one of the two options[12] used to minimize the objective function in (2.24). Then, I proceed pricing options, first mapping the parameters of the model estimated under $\mathbb{P}$ into the parameters under $\mathbb{Q}$ according to Proposition 3; second approximating option prices by COS method by Fang and Oosterlee (2008) using the MGF formula in Proposition 1. Finally, I compute the relative model IVs. As expected[13], both δ_{11} and δ_{12} are negative and equal to -0,8197 and -0,2386, respectively. These values ensure that the persistence under $\mathbb{P}$ is lower than the persistence under $\mathbb{Q}$ or, in other words, that the investors have a lower conditional mean under the historical than under the risk-neutral distribution.

2.4.3 Option pricing results

As is customary in the literature, I analyze the option pricing performances of each model in terms of Root Mean Square Error on the percentage IV:

$$RMSE_{IV} = \sqrt{\frac{1}{2}\left[(IV_1^{mkt} - IV_1^{mod})^2 + (IV_2^{mkt} - IV_2^{mod})^2\right]} \tag{2.25}$$

where $IV_{i,t}^{mkt}$ is the market IV of the option i, $IV_{i,t}^{mod}$ is the IV computed from the model, for $i = 1, 2$. For completness I report the performance results also for Root Mean Square Error on

[11] A better choice would be to explore the entire volatility surface, i.e. to use all the options in the sample, but in this case, the calibration becomes difficult and time-demanding. For this reason, I decide to use only two options.

[12] I also calibrate both the ARG and HARG model on the other option and the results are the same, in terms of pricing performance.

[13] Options are volatility-sensitive investments. They typically pay off in adverse states of nature, i.e. when the marginal utility of wealth is high. This means that such investment are negative-beta and, in turn, are characterized by negative risk premia.

option prices:

$$RMSE_P = \sqrt{\frac{1}{2}\left[(P_1^{mkt} - P_1^{mod})^2 + (P_2^{mkt} - P_2^{mod})^2\right]} \tag{2.26}$$

$P_{i,t}^{mkt}$ is the market price of the option i, $P_{i,t}^{mod}$ is the price computed from the model, for $i = 1, 2$. The former metric represents an intuitive weighting of options across strikes and maturities. The latter gives more weigth to options with high intrinsic value (DITM) and time value (longer maturity) but has the advantage of interpreting RMSE as $ errors. Table 2.3 reports the global option pricing performance on S&P 500 call option from January 5, 1996 to December 30, 2005. The first row shows the absolute RMSE_{IV} and the RMSE_P for the VARG-J model, while the remaining rows display the VARG-J relative performance with respect to other models. In particular, I compute the ratio between the RMSE_{IV} (RMSE_P) of the VARG-J and that of each competiting model. A value less than one indicates an outperformance of the model set as numerator, which for all the results presented here is the VARG-J model.

At first sight, the VARG-J model outperforms all competitors, both via RMSE_{IV} and RMSE_P.

Table 2.3: Global option pricing performance. The first row shows the implied volatility root mean square error and the price root mean square error. RMSE_{IV} and RMSE_P are both expressed in percentage. The second and the thir rows show the RMSE_{IV} and RMSE_P of the competitor models relative to the VARG-J. A ratio smaller than 1 indicates an outperformance of the VARG-J model. I use the parameter estimates from Table 2.1 and S&P 500 Call options from January 5, 1996 to December 30, 2005.

Model	RMSE_{IV}	RMSE_P
VARGJ	6.2924	0.6648
VARGJ/ARG	0.8661	0.5992
VARGJ/HARG	0.9649	0.6389

Specifically, looking at the RMSE_{IV} (RMSE_P), the VARG-J model improvement is about 14% (40%) over ARG and about 4% (36%) over HARG model. In order to get a deeper understanding of the VARG-J pricing performance, Table 2.4 (Table 2.5) reports the results in terms of RMSE_{IV} (RMSE_P) disaggregated for different maturities and moneyness. Looking at Table 2.4, Panel A shows that the VARG-J model involves some degree of underpricing for DITM call (DOTM put) options. This is because the VARG-J model does not take into account the leverage effect. Consequently, the VARG-J model attaches more probability to extreme events but is not completely able to match the smirk (negative skewness of returns) observed in the S&P

500 options (see Table 2.2). Moreover, the largest underpricing is registered for short-time to maturity options. This could be interpreted in favour of an option pricing model that allows for jumps in price. This additional factor should also attenuate also the underpricing for DITM call options. In fact, negative skewness can arise either because of negative correlation between stock index and volatility or through negative-mean jumps in price. I leave the possibility to allow for these feature to future research.

Panel B compares the performance of the VARG-J and ARG models. Overall, the VARG-J model overperforms the ARG for all moneyness and maturity, with the exceptions of DOTM short-maturity and DITM options for which the performance is quite similar between the models. The improvements on the ARG are evident especially for long-term (T>90) options. For the reason explained in the Section 2.3.2, the measurement equation in (2.9), improves the VARG-J pricing over ARG for short maturity options. This confirms the importance of jumps in volatility process. The ratio between RMSE_{IV} of VARG-J and HARG model is displayed in Panel C of Table 2.4. The advantage provided of the volatility jump factor is strong for long-term options: the improvement varies from about 25% to about 45%. So, allowing for a factor that on the one hand generates persistence and, on the other hand, generates bursts in the volatility process results in a better pricing with respect to an heterogeneous structure for the RV conditional mean. It is interesting to note that the VARG-J model outperforms the HARG model also for ATM options. In fact, in Corsi et al. (2013) and in Majewski et al. (2015) the pricing performance for all HAR-RV specifications is similar for this region. The reasonable explanation for the VARG-J underperformance in Panel C of Table 2.4 is that in order to explain the volatility smirk for short-time options an option pricing model should include also a transient price jump component. In fact, if prices jump, the RV estimator measures also price-jumps contribution to QV, inflating the estimated historical volatility over short time. Therefore, adding jumps in volatility generates volatility bursts, resulting in a fatter-tailed return distribution, making OTM and ITM options more expensive, especially long-maturity, and closing the gap between implied volatility and that directly observed from the market. All these results are in line with those reported in Table 2.5.

I now provide some more insight into the improved performance by analyzing the differences across models along two dimensions: the spot volatility level and the term structure. The VARG-J model should better match the volatility features in periods of higher volatility. For this reason I study the differences across models in both dimensions in different volatility regimes. In order to do so, I study at-the-money options in three different sample periods

Table 2.4: Option pricing performance via the percentage implied volatility root mean square error ($RMSE_{IV}$). The Panel A shows the $RMSE_{IV}$ of the VARG-J model sorted by moneyness and maturity. Panel B shows the $RMSE_{IV}$ of the ARG model relative to the VARG-J sorted by moneyness and maturity. Panel C shows the $RMSE_{IV}$ of the HARG model relative to the VARG-J sorted by moneyness and maturity. A ratio smaller than 1 indicates an outperformance of the VARG-J model. I use the parameter estimates from Table 2.1 and S&P 500 Call options from January 5, 1996 to December 30, 2005. T refers to the number of days to matutiry while m represents the moneyness defined as the underlying index level divided by the option strike price.

Moneyness	T≤50	50<T≤90	90<T≤160	T>160
Panel A: VARGJ Implied Volatility RMSE				
m≤0.94	7.1135	5.8068	4.8609	4.4883
0.94<m≤0.97	5.6552	4.9842	4.2245	4.1220
0.97<m≤1	4.6264	4.3465	3.7574	3.7756
1<m≤1.03	4.0896	3.9320	3.6435	3.7939
1.03<m≤1.06	4.6424	4.1310	3.7545	3.6637
m>1.06	15.570	9.9653	7.5613	7.5643
Panel B: VARGJ/ARG Implied Volatility RMSE				
m≤0.94	1.1848	0.8495	0.6589	0.5767
0.94<m≤0.97	0.8528	0.6492	0.5758	0.5651
0.97<m≤1	0.7676	0.6302	0.5763	0.5557
1<m≤1.03	0.7394	0.6446	0.5896	0.5826
1.03<m≤1.06	0.9011	0.7679	0.6664	0.6050
m>1.06	1.1488	1.1255	0.9567	1.0451
Panel C: VARGJ/HARG Implied Volatility RMSE				
m≤0.94	1.1663	1.0004	0.6908	0.5540
0.94<m≤0.97	1.2042	0.8638	0.6573	0.5652
0.97<m≤1	1.1133	0.8702	0.6697	0.5534
1<m≤1.03	1.0999	0.9222	0.7010	0.5857
1.03<m≤1.06	1.1961	1.1177	0.7995	0.6057
m>1.06	1.1392	1.2127	0.9863	1.0953

characterized by different level of volatility. Volatility regimes are identified from the implied volatility index (VIX) dynamics from January 5,1996 to December 30, 2005, displayed in the Figure 2.3. The first sample period is from January 5, 1996 until October 2, 1998 with an average VIX value of 15,60% and it represents a low volatility period. The second period is from

Table 2.5: Option pricing performance via the percentage price root mean square error ($RMSE_P$). The Panel A shows the $RMSE_P$ of the VARG-J model sorted by moneyness and maturity. Panel B shows the $RMSE_P$ of the ARG model relative to the VARG-J sorted by moneyness and maturity. Panel C shows the $RMSE_P$ of the HARG model relative to the VARG-J sorted by moneyness and maturity. A ratio smaller than 1 indicates an outperformance of the VARG-J model. I use the parameter estimates from Table 2.1 and S&P 500 Call options from January 5, 1996 to December 30, 2005. T refers to the number of days to matutiry while m represents the moneyness defined as the underlying index level divided by the option strike price.

Moneyness	T≤50	50<T≤90	90<T≤160	T>160
Panel A: VARGJ Price RMSE				
m≤0.94	0.2801	0.4441	0.6308	0.9909
0.94<m≤0.97	0.4015	0.7520	0.9128	1.2694
0.97<m≤1	0.4781	0.8102	0.9014	1.2068
1<m≤1.03	0.4184	0.6788	0.8588	1.2224
1.03<m≤1.06	0.3549	0.5680	0.8026	1.0551
m>1.06	0.2423	0.3772	0.5922	0.8743
Panel B: VARGJ/ARG Price RMSE				
m≤0.94	1.3588	0.6709	0.5202	0.4926
0.94<m≤0.97	0.7974	0.6156	0.5503	0.5604
0.97<m≤1	0.7213	0.6216	0.5784	0.5790
1<m≤1.03	0.7113	0.6252	0.6185	0.6383
1.03<m≤1.06	0.8950	0.7580	0.7215	0.6523
m>1.06	0.9925	0.9943	0.9173	0.8830
Panel C: VARGJ/HARG Price RMSE				
m≤0.94	1.5242	0.8285	0.5458	0.4632
0.94<m≤0.97	1.1435	0.8287	0.6151	0.5479
0.97<m≤1	1.0532	0.8536	0.6581	0.5625
1<m≤1.03	1.0616	0.8861	0.7199	0.6295
1.03<m≤1.06	1.1979	1.1149	0.8479	0.6382
m>1.06	1.1036	1.2833	1.0844	0.8735

October 5, 1998 to July 06, 2001. The average VIX value is 26,06% and I therefore define it as a high volatility regime. The third period is April 12, 2004 to December 30, 2005 and, since the average VIX is 18,27%, an intermediate volatility period.

I investigate the first dimension by comparing the differences over time between implied

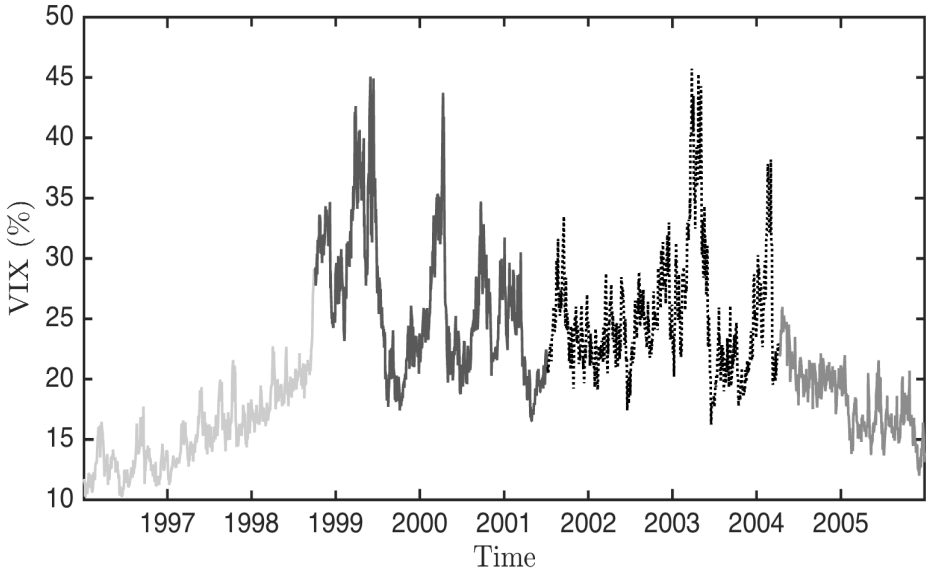

Figure 2.3: Volatility regimes. Plot of VIX index from January 5, 1996 to December 30, 2005. I identify three different volatility regimes: low volatility from January 5, 1996 to October 2, 1998, high volatility from October 5, 1998 to July 06, 2001 and medium volatility from April 12, 2004 to December 30, 2005.

volatility from the data and the models. Figure 2.4 displays the dynamics of the average weekly ATM implied volatility bias computed as the difference between the average observed market implied volatility and the average implied volatility obtained from the model. The definition of ATM options is the same used above (moneyness between 0.97 and 1.03). The ARG model has the same bias profile of the HARG model but it presents a greater bias in magnitude. The HARG model presents a negative bias both during low volatility and high volatility regimes. The VARG-J model has a negative bias during low volatility period and a positive bias during high volatility period. In 2004-2005, an intermediate volatility period, both models present a negative bias. In comparison, the VARG-J model shows a smaller bias in all volatility regimes. In particular, during the high volatility period, the average VARG-J bias is smaller, in absolute term, than the HARG bias (-0.0131 vs. 0.0050). This suggests that the VARG-J model is better able to capture the dynamics of market volatility, especially, as expected, during high volatility regime.

Figure 2.5 analyses the second dimension of interest: the volatility term structure. At first sight the VARG-J model outperforms both competitors during all three volatility regimes. In particular, it performs better in the high volatility period. In the VARG-J model the spot volatility has a larger impact on implied volatility at short maturities. This feature is more evident dur-

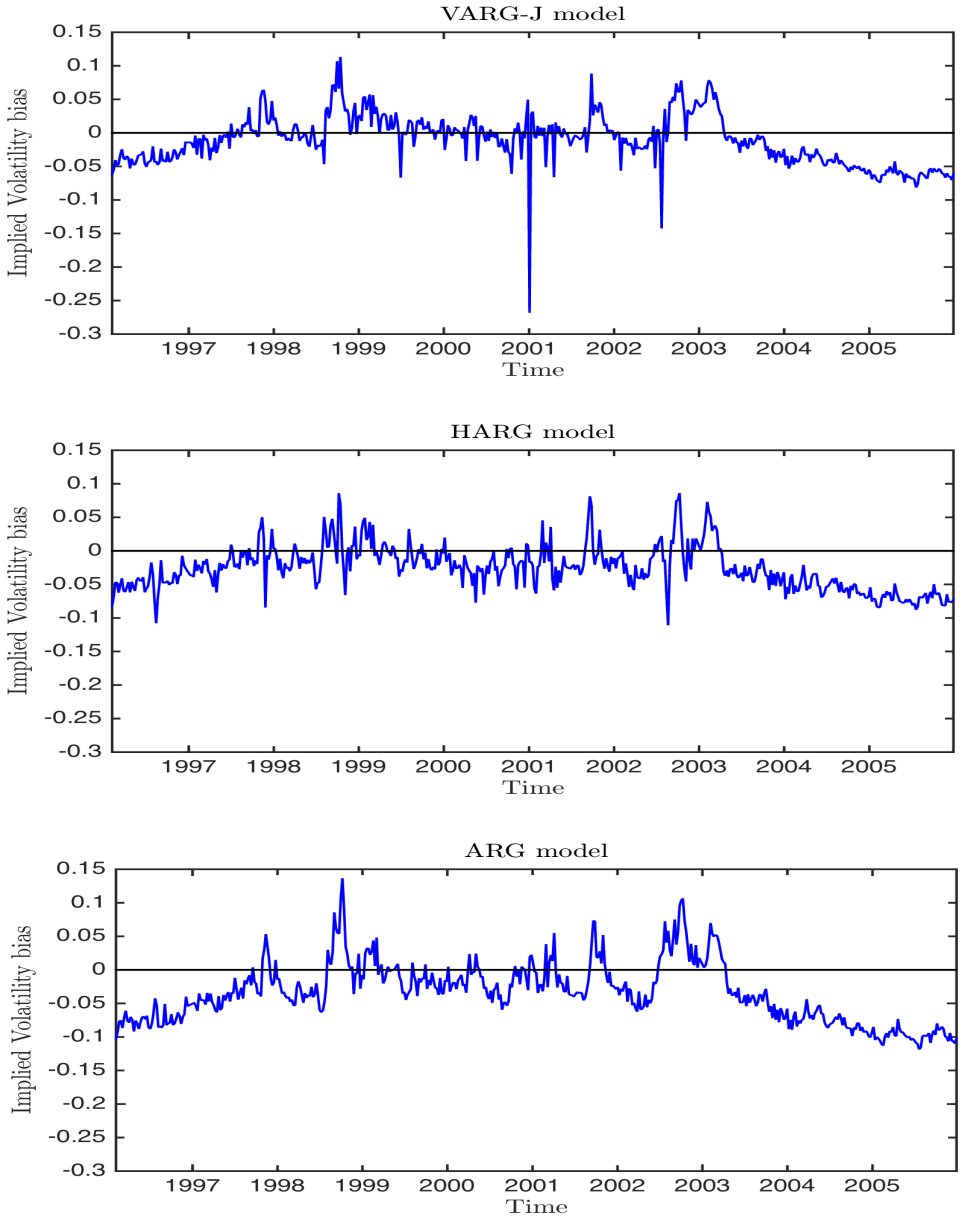

Figure 2.4: Weekly implied volatility bias for at-the-money options. On each Wednesday I compute the B&S implied volatility for at-the-money option contracts using parameters estimates from Table 2.1. Options with moneyness between 0.97 and 1.03 are considered at-the-money.

ing low and high volatility periods, in which the VARG-J model captures a smaller bias with respect to the other models. The jump volatility factor, allowing for bursts in the volatility dynamics, captures the variation in longer-maturity implied volatilities. Moreover, since the jump factor can experience periods of zero values, the VARG-J model is able to adequately capture short-maturity implied volatilities.

The same conclusions can be drawn looking at Figure 2.6 in which I plot the difference

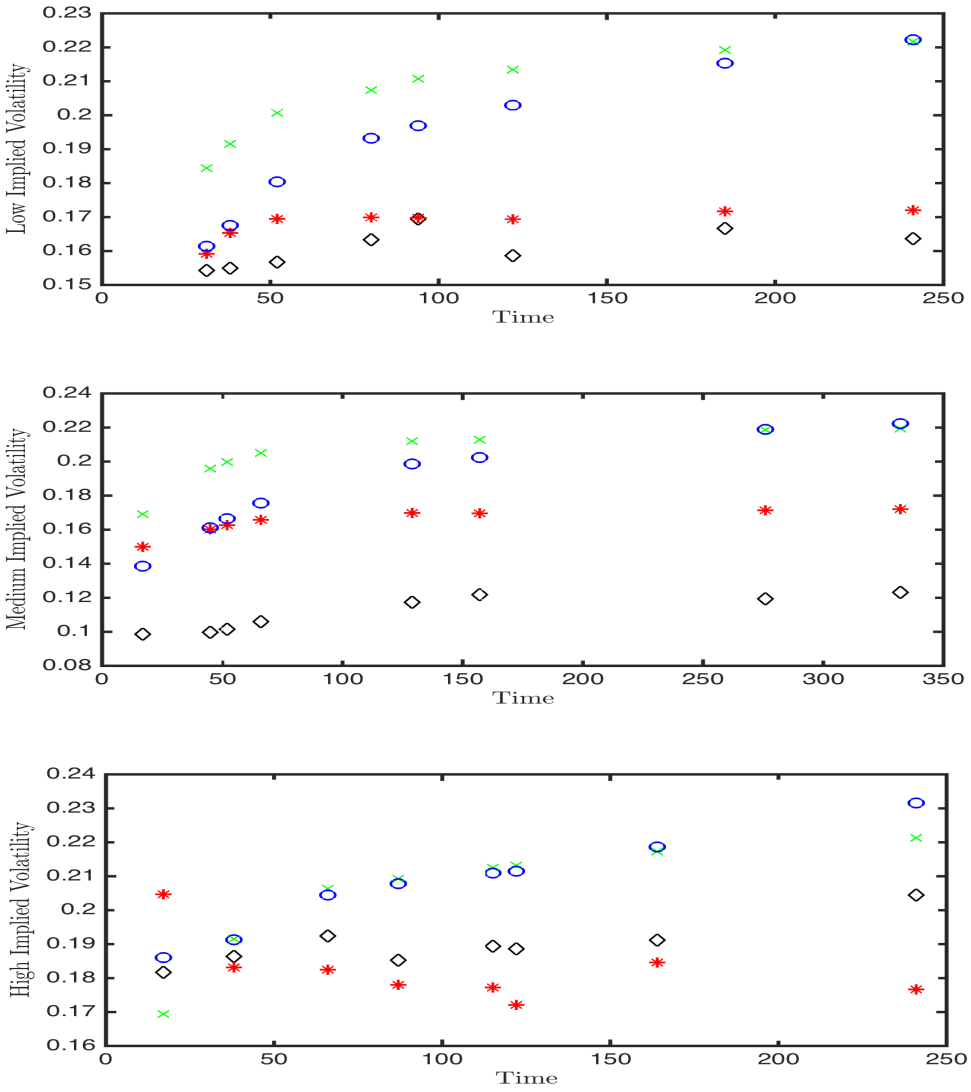

Figure 2.5: Implied volatility term structure for at-the-money options. Options with moneyness between 0.97 and 1.03 are considered at-the-money. I compute the B&S implied volatility for at-the-money option contracts using parameters estimates from Table 2.1 for three different volatility regimes: low, medium and high. In each panel, a "$\diamond$" represents the market implied volatility, a "$*$" represents the VARG-J model, a "$\circ$"represents the HARG model and a "$\times$" represents the ARG model. The time to maturity is on the horizontal axis.

between the average IV of ATM long maturity options and the level. Options with moneyness between 0.97 and 1.03 and maturity greater than 120 days are considered ATM long maturity. By looking at the dynamics of the IV term structure, the VARG-J model performs well and it is possible to clearly identify the improvement due to the introduction of a jump factor in the volatility dynamics. I conclude that the VARG-J model is able to capture both dimensions of the implied volatility surface thanks to the jump factor that generates a higher persistence and

bursts in the volatility dynamics.

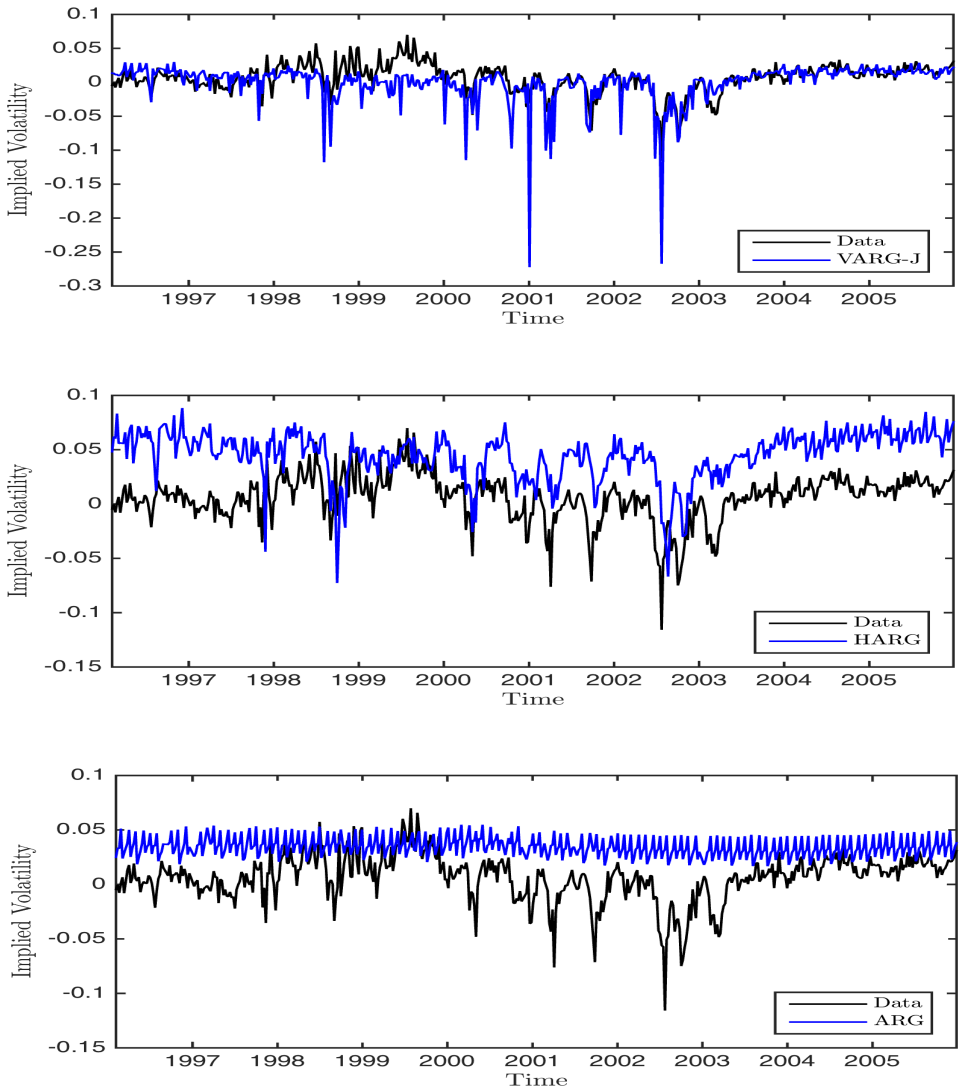

Figure 2.6: Term structure of the implied volatility surface. The slope of the volatility surface is computed as the difference between the average at-the-money long maturity options and level. Options with moneyness between 0.97 and 1.03 are considered at-the-money while options with maturity greater than 120 days are considered long maturity options. In each panel the black line represents the average market implied volatility while the blue line represents the average model implied volatility.

2.5 Conclusion and directions of future work

In this paper I propose an affine discrete-time model which features a multi-factor volatility specification and exploits information about the latent volatility process using high frequency data. In the VARG-J model volatility experiences periods of extreme movements through a

volatility jump factor that is modeled as an ARG_0 process, in addition to a continuous component, modeled as an EARG. Whitin this framework, I show how to analytically characterize the MGF under both $\mathbb{P}$ and $\mathbb{Q}$; the change of measure is performed adopting a SDF which incorporates risk premia associated to the sources of risk concerning the two volatility factors, beyond the risk related to shocks in returns. The MGF allows to obtain closed-form option pricing formulas in the VARG-J model. A great advantage of VARG-J is represented by the model estimation strategy. I estimate the model's parameters under the historical measure via pseudo-maximum likelihood with the Extended Kalman Filter. This estimation procedure allows to filter both volatility factors, exploiting the information on the latent volatility in the RV estimates and, at the same time, by controlling for measurement errors and overnight effect. I choose the 5-minute RV estimator in order to reduce at minimum the impact of microstructure noise and I introduce a measurement equation for the total volatility in order to correct the RV bias due to measurement erros and overnight effects.

The results show a good pricing performance for all moneyness and maturities. The highest improvement is registered for D-OTM and ITM options at medium and long maturity. For these option categories the VARG-J model outperforms both competiting models and provides an improvement also for ATM options. The multi-factor volatility specification presented in this paper is able to generate a fatter-tailed return distribution and bursts in volatility. In this way OTM and ITM options are more expensive and the gap between implied volatility and that directly observed from the market is reduced. Moreover, the VARG-J model is able to capture both dimensions of the implied volatility surface thanks to the jump factor that generates bursts in the volatility dynamics and allows for a higher persistence. Specifically, the VARG-J model fits the data better showing a small bias in all volatility regimes and it captures the dynamics of market implied volatility especially during high volatility periods. Moreover, by looking at the term structure of the implied volatility surface, the VARG-J model adequately captures the slope in the implied volatility surface along the time dimension. The model presented here matches the stylized fact that shocks on the current conditional volatility impact on the conditional variance forecast up to one year in the future. This has an impact on the price of at-the-money options at different maturities.

The simplified model specification presented in this paper does not fit some important retun stylized facts. Nevertheless, I think that the performance results underline the importance of the jump factor in the volatility dynamics. Therefore, a number of extensions to the VARG-J model are warranted. First, price jumps are needed to decrease the ?gap? between model

implied volatility and that directly observed from the market. According to the specification, the introduction of jumps in price should improve the pricing over different options categories and explain different features of volatility surface. Second, a leverage effect component should be introduced in order to take into account the asymmetry in the volatility smile. Finally I believe that introducing a dependence between volatility components should leads to precise and accurate measurement and forecast of the unobservable asset volatility improving pricing performance.

Appendix B

Technical Details of Chapter 2

B.1 MGF for VARG-J model under $\mathbb{P}$ measure

The Assumption 1 in Majewski et al. (2015) is the following:

$$\begin{aligned}
&\mathbb{E}^{\mathbb{P}}\left[\exp(zy_{t+1}+\mathbf{b}'\mathbf{f}_{t+1}+\mathbf{c}'\mathbf{l}_{t+1})|\mathcal{F}_t,\mathcal{L}_t\right] \\
&= \exp\left[\mathcal{A}_t(z,\mathbf{b},\mathbf{c})+\sum_{i=1}^{p}\mathcal{B}_i(z,\mathbf{b},\mathbf{c})'\cdot\mathbf{f}_{t+1-i}+\sum_{j=1}^{q}\mathcal{C}_j(z,\mathbf{b},\mathbf{c})'\cdot\mathbf{l}_{t+1-j}\right]
\end{aligned} \tag{B.1}$$

for some functions $\mathcal{A}:\mathbb{R}\times\mathbb{R}^k\times\mathbb{R}^k\to\mathbb{R}, \mathcal{B}_i:\times\mathbb{R}^k\times\mathbb{R}^k\to\mathbb{R}^k$ and $\mathcal{C}_i:\times\mathbb{R}^k\times\mathbb{R}^k\to\mathbb{R}^k$, where $\mathbf{b},\mathbf{c}\in\mathbb{R}^k$ and $\cdot$ stands for the scalar product in $\mathbb{R}^k$.

For the setup in Section 2.2.1, assumption (B.1) is satisfied imposing $\mathbf{l}_t=0$ for $t=1,\dots,T$ and $p=1$[1].

To check that and to derive the expressions for $\mathcal{A}_t(z,\mathbf{b},\mathbf{c})$ and $\mathcal{B}_i(z,\mathbf{b},\mathbf{c})$:

$$\begin{aligned}
&\mathbb{E}^{\mathbb{P}}\left[\exp(zy_{t+1}+\mathbf{b}'\mathbf{f}_{t+1}|\mathcal{F}_t\right] \\
&= \mathbb{E}^{\mathbb{P}}\left[\exp(zr_t+z\lambda f_{t+1}+z\sqrt{f_{t+1}}\epsilon_{t+1}+b_1f_{1,t+1}+b_2f_{2,t+1}|\mathcal{F}_t\right] \\
&= e^{zr}\mathbb{E}^{\mathbb{P}}\left[\exp[(b_1+z\lambda)f_{1,t+1}+(b_2+z\lambda)f_{2,t+1}+z\sqrt{f_{t+1}}\epsilon_{t+1}]|\mathcal{F}_t\right] \\
&= e^{zr}\mathbb{E}^{\mathbb{P}}\left\{\exp\left[(b_1+z\lambda)f_{1,t+1}+(b_2+z\lambda)f_{2,t+1}\right]\right. \\
&\quad\left.\times\mathbb{E}^{\mathbb{P}}\left[\exp\left[z\sqrt{f_{t+1}}\epsilon_{t+1}\right]\middle|f_{1,t+1},f_{2,t+1},\mathcal{F}_t\right]\middle|\mathcal{F}_t\right\}
\end{aligned}$$

To compute the inner expectation I now use the following property: if $Z\sim\mathcal{N}(0,1)$ and

[1] Richer characterization of volatility dynamics are left for future work.

$Y = aZ$, then

$$\mathbb{E}\{\exp[xY]\} = \exp\left[\frac{1}{2}(xa)^2\right]$$

Hence:

$$\begin{aligned}
&\mathbb{E}^{\mathbb{P}}\left[\exp(zy_{t+1} + \mathbf{b}'\mathbf{f}_{t+1})|\mathcal{F}_t\right] \\
&= e^{zr}\mathbb{E}^{\mathbb{P}}\left\{\exp\left[(b_1 + z\lambda)f_{1,t+1} + (b_2 + z\lambda)f_{2,t+1} + \frac{z^2}{2}f_{t+1}\ \right]\middle|\ \mathcal{F}_t\right\} \\
&= e^{zr}\mathbb{E}^{\mathbb{P}}\left\{\exp\left[(b_1 + z\lambda)f_{1,t+1} + (b_2 + z\lambda)f_{2,t+1} + \frac{z^2}{2}f_{1,t+1} + \frac{z^2}{2}f_{2,t+1}\ \right]\middle|\ \mathcal{F}_t\right\} \\
&= e^{zr}\mathbb{E}^{\mathbb{P}}\left\{\exp\left[x_1(z,\mathbf{b})f_{1,t+1} + x_2(z,\mathbf{b})f_{2,t+1}\right]|\ \mathcal{F}_t\right\}
\end{aligned}$$

where:

$$x_1(z,\mathbf{b}) = b_1 + z\lambda + \frac{z^2}{2} \tag{B.2}$$

$$x_2(z,\mathbf{b}) = b_2 + z\lambda + \frac{z^2}{2} \tag{B.3}$$

In what follows I will sometimes simplify the notation using x_1 (resp. x_2) instead of $x_1(z,\mathbf{b})$ (resp. $x_2(x,\mathbf{b})$). I now use the following assumption of the noncentral Gamma-zero distribution: if $Z \sim \gamma_0(\theta,\mu)$, then

$$\mathbb{E}[\exp(xZ)] = \exp\left[\frac{x\mu}{1-x\mu}\,\theta\right].$$

I get:

$$\begin{aligned}
&\mathbb{E}^{\mathbb{P}}\left[\exp(zy_{t+1} + \mathbf{b}'\mathbf{f}_{t+1})|\mathcal{F}_t\right] \\
&= e^{zr}\mathbb{E}^{\mathbb{P}}\left\{\exp[x_1(z,\mathbf{b})f_{1,t+1}]\mathbb{E}^{\mathbb{P}}\left[\exp\left(x_2(z,\mathbf{b})f_{2,t+1}\right)|\ f_{1,t+1},\mathcal{F}_t\right]\middle|\ \mathcal{F}_t\right\} \\
&= e^{zr\frac{x_2\mu_2}{1-x_2\mu_2}\theta_{2t}}\mathbb{E}^{\mathbb{P}}\left\{\exp\left[x_1 f_{1,t+1}\right]|\ \mathcal{F}_t\right\} \\
&= \mathrm{e}^{zr+V_2(x_2,\mu_2)\theta_{2t}}\mathbb{E}^{\mathbb{P}}\left\{\exp\left[x_1 f_{1,t+1}\right]|\ \mathcal{F}_t\right\}
\end{aligned}$$

where:

$$V_2[x_2,\mu_2] = \frac{x_2(z,\mathbf{b})\mu_2}{1 - x_2(z,\mathbf{b})\mu_2} \tag{B.4}$$

I now use the following property of the noncentral Gamma distribution: if $Z \sim \gamma_\nu(\theta,\mu)$, then

$$\mathbb{E}[\exp(xZ)] = \exp\left[\frac{x\mu}{1-x\mu}\,\theta - \nu\log(1-x\mu)\right].$$

I get:

$$\begin{aligned}
&\mathbb{E}^{\mathbb{P}}\left[\exp(zy_{t+1}+\mathbf{b}'\mathbf{f}_{t+1})|\mathcal{F}_t\right] \\
&= \exp\left\{zr+V_2(x_2,\mu_2)\theta_{2t}-\nu\log(1-x_1\mu_1)+\frac{x_1\mu_1}{1-x_1\mu_1}\theta_{1t}\right\} \\
&= \exp\left\{zr-\nu W_1(x_1,\mu_1)+V_1(x_1,\mu_1)\theta_{1t}+V_2(x_2,\mu_2)\theta_{2t}\right\} \qquad \text{(B.5)}
\end{aligned}$$

$$W_1[x_1,\mu_1] = \log[1-x_1(z,\mathbf{b})\mu_1] \qquad \text{(B.6)}$$

$$V_1[x_1,\mu_1] = \frac{x_2(z,\mathbf{b})\mu_2}{1-x_1(z,\mathbf{b})\mu_1} \qquad \text{(B.7)}$$

Substituting (2.6) in (B.5) and collecting terms, it is easy to check that Assumption B.1 is satisfied, with:

$$\mathcal{A}(z,\mathbf{b},\mathbf{c})=\mathcal{A}(z,\mathbf{b}) = zr-\nu W_1(x_1,\mu_1)+V_1(x_1,\mu_1)d_1+V_2(x_2,\mu_2)d_2 \qquad \text{(B.8)}$$

$$\mathcal{B}_i(z,\mathbf{b},\mathbf{c})'=\mathcal{B}(z,\mathbf{b})' = [V_1(x_1,\mu_1)\ ,\ V_2(x_2,\mu_2)]\,\boldsymbol{\beta} \qquad \text{(B.9)}$$

$$\mathcal{C}_j(z,\mathbf{b},\mathbf{c})'=\mathcal{C}(z,\mathbf{b})' = 0 \qquad \text{(B.10)}$$

where $\mathbf{d}=(d_1,d_2)'$ and $\boldsymbol{\beta}=(\beta_1,\beta_2)'$.
As explained in Majewski et al. (2015) the above expressions represent the main ingredient to compute the MGF of the log-return $y_{t,T}=\sum_{i=t+1}^{T}y_i$ for VARG-J model

$$\varphi^{\mathbb{P}}_{t,T,z}=\mathbb{E}^{\mathbb{P}}[e^{zy_{t,T}}|\mathcal{F}_t]=exp\left(a_t+\mathbf{b}_t'\mathbf{f}_t\right)$$

where

$$a_s = a_{s+1}+zr-\nu W_{1,s+1}+d_1V_{1,s+1}+d_2V_{2,s+1} \qquad \text{(B.11)}$$

$$\mathbf{b}_s' = (V_{1,s+1},V_{2,s+1})\boldsymbol{\beta} \qquad \text{(B.12)}$$

$$\qquad \text{(B.13)}$$

with

$$x_{h,s+1}=x_h(z,\mathbf{b}_{s+1})=\mathbf{b}_{s+1}+z\lambda+\frac{z^2}{2},\quad h=1,2$$

The functions V and W are defined as follows:

$$\begin{aligned} V_{h,s+1} &= V_h(x_{h,s+1}, \mu_h) = \frac{x_{h,s+1}\mu_h}{1 - x_{h,s+1}\mu_h}, \quad h = 1, 2 \\ W_{1,s+1} &= W_1(x_{1,s+1}, \mu_1) = \log(1 - x_{1,s+1}\mu_1) \end{aligned}$$

B.2 No-arbitrage condition

The SDF assumed is

$$M_{t,t+1} = \frac{\exp(-\delta_2 y_{t+1} - \delta_{11} f_{1,t+1} - \delta_{12} f_{2,t+1})}{\mathbb{E}^{\mathbb{P}}[\exp(-\delta_2 y_{t+1} - \delta_{11} f_{1,t+1} - \delta_{12} f_{2,t+1})|\mathcal{F}_t]} \tag{B.14}$$

The no-arbitrage conditions are

$$\mathbb{E}^{\mathbb{P}}[M_{s,s+1}|\mathcal{F}_s] = 1 \qquad \text{for} \quad s \in \mathbb{N} \tag{B.15}$$

$$\mathbb{E}^{\mathbb{P}}[M_{s,s+1}e^{y_{s+1}}|\mathcal{F}_s] = e^{r_{s+1}} \qquad \text{for} \quad s \in \mathbb{N} \tag{B.16}$$

Let $\boldsymbol{\delta}_1 = (\delta_{11}, \delta_{12})'$. To enforce no arbitrage, I use Proposition 2 in Majewski et al. (2015), which shows that absence of arbitrage is equivalent to

$$\begin{aligned} \mathcal{A}(1 - \delta_2, -\boldsymbol{\delta}_1) &= r + A(-\delta_2, -\boldsymbol{\delta}_1) \\ \mathcal{B}(1 - \delta_2, -\boldsymbol{\delta}_1) &= B(-\delta_2, -\boldsymbol{\delta}_1) \end{aligned}$$

These equalities are implied by

$$\begin{aligned} x_1(1 - \delta_2, -\boldsymbol{\delta}_1) &= x_1(-\delta_2, -\boldsymbol{\delta}_1) \\ x_2(1 - \delta_2, -\boldsymbol{\delta}_1) &= x_2(-\delta_2, -\boldsymbol{\delta}_1). \end{aligned}$$

For this to hold, it is easy to check that it is sufficient to impose

$$\delta_2 = \lambda + \frac{1}{2} \tag{B.17}$$

So, the no-arbitrage condition fix the level of the equity risk premium, while both the continuous and discontinuous variance risk premia are free parameters to be calibrated on option sample.

B.3 MGF for VARG-J model under $\mathbb{Q}$

Let $y_{t,T} = \log(S_T/S_t)$

$$\varphi^{\mathbb{Q}}_{\delta_2,\boldsymbol{\delta_1}}(t,T,z) = \mathbb{E}^{\mathbb{Q}}\left(e^{zy_{t,T}}|\mathcal{F}_t\right) = \exp\left(a^* + \mathbf{b}^{*\prime}\mathbf{f}_t\right) \tag{B.18}$$

where:

$$\begin{aligned} a^*_s &= a^*_{s+1} + \mathcal{A}_s(z-\delta_2, \mathbf{b}^*_{s+1} - \boldsymbol{\delta_1}) - \mathcal{A}_s(-\delta_2, -\boldsymbol{\delta_1}) && \text{(B.19)}\\ \mathbf{b}^*_s &= \mathcal{B}(z-\delta_2, \mathbf{b}^*_{s+1} - \boldsymbol{\delta_1}) - \mathcal{B}(-\delta_2, -\boldsymbol{\delta_1}) && \text{(B.20)}\\ & && \text{(B.21)} \end{aligned}$$

subject to the initial conditions:

$$a^*_T = 0, \quad \boldsymbol{b}^*_T = \mathbf{0}$$

We now specialize these expression for the setup outlined in section 2.2.1. Consider equation (B.19). Using (B.8), I get

$$a^*_s = a^*_{s+1} + zr_s - \nu(W^*_{1,s+1} - W^y_1) + d_1(V^*_{1,s+1} - V^y_1) + d_2(V^*_{2,s+1} - V^y_2) \tag{B.22}$$

where:

$$\begin{aligned} x^*_{h,s+1} &= x_h(z-\delta_2, \boldsymbol{b}^*_{s+1} - \boldsymbol{\delta}_1), \quad h=1,2\\ y^*_h &= x_h(-\delta_2, -\boldsymbol{\delta}_1) = -\delta_{1h} - \delta_2\lambda + \frac{\delta_2^2}{2}, \quad h=1,2\\ V^*_{h,s+1} &= V_h(x^*_{h,s+1}, \mu_h), \quad h=1,2\\ V^y_h &= V_h(y^*_h, \mu_h), \quad h=1,2\\ W^*_{1,s+1} &= W_1(x^*_{1,s+1}, \mu_1)\\ W^y_1 &= W_1(y^*_1, \mu_1) \end{aligned}$$

Using (B.9) and (B.10), equation (B.20) becomes:

$$\mathbf{b}^{*\prime}_s = (V^*_{1,s+1} - V^y_1\ ,\ V^*_{2,s+1} - V^y_2)\boldsymbol{\beta} \tag{B.23}$$

Note that o compute the MGF of $y_{t,T}$ under $\mathbb{P}$ it is simply needed to plug $\delta_2 = 0$ and $\boldsymbol{\delta_1} = \mathbf{0}$ in the expression of $\varphi^{\mathbb{Q}}_{\delta_2,\boldsymbol{\delta_1}}(t,T,z)$.

The MGFs $\varphi^{\mathbb{Q}}_{\delta_2,\delta_1}(t,T,z)$ and $\varphi^{\mathbb{P}}_{0,\mathbf{0}}(t,T,z)$ derived above depend on the parameters under $\mathbb{P}$, $\boldsymbol{\psi}$, defined in (2.14), and on the risk premium parameters $\delta = (\delta_2, \delta_1')'$ introduced in the SDF (B.14). I now show that the MGF under $\mathbb{Q}$ can be rewritten as the MGF under $\mathbb{P}$ using a new set of parameters, $\boldsymbol{\psi}^{\mathbb{Q}}$, the risk-neutral ones

$$\boldsymbol{\psi}^{\mathbb{Q}} = [\lambda^{\mathbb{Q}}, \nu^{\mathbb{Q}}, \mu_1^{\mathbb{Q}}, \mu_2^{\mathbb{Q}}, \mathbf{d}^{\mathbb{Q}}, \boldsymbol{\beta}^{\mathbb{Q}}] \tag{B.24}$$

To derive the expression of $\boldsymbol{\psi}^{\mathbb{Q}}$ as a function of $\boldsymbol{\psi}^{\mathbb{Q}}$ and δ, I match the parameters using the identity:

$$\varphi^{\mathbb{Q}}_{\delta_2,\delta_1}(t,T,z;\boldsymbol{\psi},\delta) = \varphi_{0,\mathbf{0}}(t,T,z;\boldsymbol{\psi}^{\mathbb{Q}}) \tag{B.25}$$

It is useful to denote

$$\begin{aligned}
x^{\mathbb{Q}}_{h,s+1} &= x_h(z, \boldsymbol{b}^*_{s+1}; \boldsymbol{\psi}^{\mathbb{Q}}), \quad h = 1,2 \\
V^{\mathbb{Q}}_{h,s+1} &= V_h(x^{\mathbb{Q}}_{h,s+1}, \mu^{\mathbb{Q}}_h), \quad h = 1,2 \\
W^{\mathbb{Q}}_{1,s+1} &= W_1(x^{\mathbb{Q}}_{1,s+1}, \mu^{\mathbb{Q}}_1)
\end{aligned}$$

For this to hold, (B.22) needs to be matched with (B.11) and (B.23) with (B.12) , where (B.11) and (B.12) are evaluated at $V^{\mathbb{Q}}_{1,s+1}$, $V^{\mathbb{Q}}_{2,s+1}$ and $W^{\mathbb{Q}}_{1,s+1}$. Note that since I start from the same initial conditions (B.25) requires

$$\nu(W^*_{1,s+1} - W^y_1) = \nu^{\mathbb{Q}} W^{\mathbb{Q}}_{1,s+1} \tag{B.26}$$

$$d_1(V^*_{1,s+1} - V^y_1) = d^{\mathbb{Q}}_1 V^{\mathbb{Q}}_{1,s+1} \tag{B.27}$$

$$d_2(V^*_{2,s+1} - V^y_2) = d^{\mathbb{Q}}_2 V^{\mathbb{Q}}_{2,s+1} \tag{B.28}$$

$$(V^*_{1,s+1} - V^y_1 \ , \ V^*_{2,s+1} - V^y_2)\boldsymbol{\beta} = (V^{\mathbb{Q}}_{1,s+1} \ , \ V^{\mathbb{Q}}_{2,s+1})\boldsymbol{\beta}^{\mathbb{Q}} \tag{B.29}$$

for all s.

Consider (B.26). This requires

$$\nu[\log(1 - x^*_{1,s+1}\mu_1) - \log(1 - y^*_1\mu_1)] = \nu^{\mathbb{Q}} \log(1 - x^{\mathbb{Q}}_{1,s+1}\mu^{\mathbb{Q}}_1)$$

Sufficient conditions for this equality to hold are

$$\nu^{\mathbb{Q}} = \nu, \quad \mu^{\mathbb{Q}}_1 = \frac{\mu_1}{1 - y^*_1\mu_1} \quad \text{and} \quad x^{\mathbb{Q}}_{1,s+1} = x^*_{1,s+1} - y^*_1.$$

In turn, it can be checked that the latter equality is valid if I pose

$$\lambda^{\mathbb{Q}} = -\frac{1}{2}.$$

Note that under these conditions I also have $x^{\mathbb{Q}}_{2,s+1} = x^*_{2,s+1} - y^*_2$.

I now consider (B.27):

$$d_1 \left(\frac{x^*_{1,s+1}\mu_1}{1 - x^*_{1,s+1}\mu_1} - \frac{y^*_1\mu_1}{1 - y^*_1\mu_1} \right) = d^{\mathbb{Q}}_1 \frac{x^{\mathbb{Q}}_{1,s+1}\mu^{\mathbb{Q}}_1}{1 - x^{\mathbb{Q}}_{1,s+1}\mu^{\mathbb{Q}}_1}. \tag{B.30}$$

Substituting the expressions for $x^{\mathbb{Q}}_{1,s+1}$ and $\mu^{\mathbb{Q}}_1$, I get

$$d^{\mathbb{Q}}_1 = \frac{d_1}{1 - y^*_1\mu_1}.$$

Note that (B.27) also implies the equality $V^{\mathbb{Q}}_{1,s+1} = (1 - y^*_1\mu_1)(V^*_{1,s+1} - V^y_1)$.

Now turn to (B.28):

$$d_2 \left(\frac{x^*_{2,s+1}\mu_2}{1 - x^*_{2,s+1}\mu_2} - \frac{y^*_2\mu_2}{1 - y^*_2\mu_2} \right) = d^{\mathbb{Q}}_2 \frac{x^{\mathbb{Q}}_{2,s+1}\mu^{\mathbb{Q}}_2}{1 - x^{\mathbb{Q}}_{2,s+1}\mu^{\mathbb{Q}}_2}.$$

If I substitute for $x^{\mathbb{Q}}_{2,s+1}$ and $\mu^{\mathbb{Q}}_2$ the expressions obtained above, I get an identity with the same structure as (B.30). Its validity thus requires expressions for $d^{\mathbb{Q}}_2$ and $\mu^{\mathbb{Q}}_2$ which are specular to those for $d^{\mathbb{Q}}_1$ and $\mu^{\mathbb{Q}}_1$:

$$\mu^{\mathbb{Q}}_2 = \frac{\mu_2}{1 - y^*_2\mu_2} \quad \text{and} \quad d^{\mathbb{Q}}_2 = \frac{d_2}{1 - y^*_2\mu_2}$$

Note that these solutions also imply that $V^{\mathbb{Q}}_{2,s+1} = (1 - y^*_2\mu_1)(V^*_{2,s+1} - V^y_2)$

Finally, I turn to (B.29) which implies:

$$\beta^{\mathbb{Q}}_h = \frac{\beta_h}{1 - y^*_h\mu_h}, \quad h = 1, 2$$

Chapter 3

Explaining Contagion during European Debt Crisis with Mutually Exciting Jump Processes

The European sovereign debt crisis has raised many questions about financial contagion. The sovereign debt crisis has influenced the general economic conditions across European countries, affecting, among others, the banking market. This paper is concerned with two main questions: *a*) whether the crisis in Greece spreads in other European countries; *b*) whether contagion, if present, is mainly due to being a member of a single currency union or to regional proximity. The answer to both questions is addressed by estimating an econometric model for the dynamics of asset returns by Aït-Sahalia et al. (2015) in which contagion is defined as the within and between country transmission of shocks. According to the empirical evidence, the European banking markets are exposed to contagion which is transmitted through inter-bank connections. Countries belonging to the monetary union are less exposed to contagion with respect to Eurozone countries. Moreover, news about the crisis in Greece affects European markets, especially Eurozone countries, generating panic about domestic financial stability as well as domestic news.

JEL-Classification: C58, G01, C32

Keywords: Contagion, Sovereign debt crisis, Jumps, Hawkes process

3.1 Introduction

Contagion represents the mechanism by which financial instability becomes so widespread that a crisis in one country reaches systemic dimensions, as proved by the recent Greek crisis. The Greek debt crisis officially started on late 2009, when the prime minister revealed a large gap in Greece's accounts. In early 2010, the major agencies downgraded Greece's credit rating. Very soon, the situation became tense because of the fear that the Greek's crisis could cause contagion and, in turn, cause the default of those countries with a similarly economy, i.e GIPS countries[1]. These tensions immediately reflected on all major financial markets. The major rating agencies also lowered the rating of different European countries and, consequently, banks based in those countries or with substantial exposure to government bonds of downgraded countries were in difficulty. In many cases, the market turmoil was amplified. Hence, the sovereign debt crisis contaminated the general economic conditions across European countries, affecting, among others, the banking market.

There is consensus in the literature that sovereign crises can be transmitted to banks through three main channels. First, the bank's funding capacity depends on the value of government bonds held in its portfolio. This means that a loss in the value of government bonds implies a reduction in credit supply (see Angeloni and Wolff, 2012). Second, banks typically use government bonds as collateral both in interbank transactions and in operation with central banks: a reduction in the value of sovereign bonds has adverse effect on credit supply (see Correa et al., 2013). Third, the yield on sovereign debt as well as the sovereign rating may impact on the cost of credit to the economy. Indeed, a rise in the government bond yields increases the banks' funding cost and can worsen the access of banks to money market and deposit markets (see Arezki et al., 2011).

There is a wide literature on contagion (see Caporin et al., 2013, Ang and Longstaff, 2013, Longstaff et al., 2011, Mauro et al., 2002 and Pan and Singleton, 2008) as well as a vast literature on contagion between sovereign and banking market (see Acharya et al., 2014, Alter and Beyer, 2014, Alter and Schüler, 2012, De Bruyckere et al., 2013, Angeloni and Wolff, 2012 and Gross and Kok, 2013). Both streams of literature show significant evidence of contagion across European countries, during the recent sovereign crisis. To the best of my knowledge, there is

[1]GIPS is an acronym coined during the recent sovereign crisis which refers to Greece, Italy, Portugal and Spain. These countries were identified as the economies where sovereign debt had increased sharply due to banks bailouts

no empirical work that looks at whether mutual contagion exists at the bank level[2], i.e. once the sovereign crisis affects the banking system in one country, whether banks based in other European countries are affected or not by the crisis.

I contribute to the empirical literature on financial contagion by analyzing whether there is empirical evidence of contagion at the bank level, measuring the direction and the size of contagion transmission between European banking markets. Specifically, I am interested in studying mutual contagion in the European banking marketand in understanding whether the Greek sovereign crisis spread across countries. In order to understand and quantify the contagion transmission on the banking market, I estimate the econometric model by Aït-Sahalia et al. (2015) in which contagion is defined as the within and between country transmission of shocks. One of the main driver of the Greek crisis was the downgrade of rating agencies. Credit rating changes are extreme events which are reflected into downfalls of stock prices. In order to capture the magnitude and the direction of financial contagion, asset returns are directly modeled as a Hawkes jump diffusion process. This process is able to reproduce both time and space propagation during a crisis period: a jump in one market raises the probability of future jumps both in the same market and elsewhere generating clusters both in time and in space. Once there is an extreme news event in financial markets, jumps occur and the occurrence of jumps raises the probability of having future jumps both in time and across countries. The cross-excitation could be asymmetric between countries. Indeed, contagion materializes under the assumption of mutually exciting jump processes. This specification allows to have a data generating process for returns that is able to generate the salient features of contagion: the cross-sectional and serial dependence observed across markets during the European sovereign debt crisis.

This paper is concerned with two main questions: *a*) whether the crisis in Greece spread in other European countries; *b*) whether contagion, if present, is mainly due to being a member of a single currency union or to regional proximity. The latter question is particularly relevant since banks tend to hold a large amount of government debt securities because of the absence of risk weight and limit on large exposures to sovereign bonds issued in domestic currency, according to the Basel Accords [3]. Manasse and Zavalloni (2013) look at the empirical evidence

[2]There exists a vast literature on systemic risk (see Allen et al., 2009 for a comprehensive overview), which is strictly related to contagion since the latter can be the motivation for the failure of multiple financial institutions.

[3]The Basel Accords allows for a 0% risk weight to be assigned to government bonds issued in domestic currency. Moreover, the Accord exempts sovereign debt issued in domestic currency from the 25% limit on large exposures, at odds to all other assets holdings.

of contagion during the sovereign debt crisis for several European countries both inside and outside the monetary union. The authors find that before the crisis being part of the monetary union worked as protection against contagion and systemic risk. However, during the Greek crisis, countries belonging to the Euro area were more exposed to contagion with respect to countries non belonging to the monetary union. Nevertheless, Manasse and Zavalloni (2013) find that Sweden, Denmark, and UK were largely affected by sovereign debt crisis.
The answer to the former question is addressed considering bank market indexes for Greece, Italy, and Spain, grouped as *Eurozone* while the latter question is addressed considering Denmark as *no-Eurozone* country. The estimation procedure is based on the generalized method of moments (GMM) in which the moment restrictions are derived in closed-form. The empirical analysis indicates that there is clear evidence of contagion from Greece to European countries as well as self-contagion in all countries. Moreover, changes over time of contagion coefficients can be linked to unexpected news about Greek economic conditions and to domestic news.
The rest of the paper is organized as follows. Section 3.2 describes the Hawkes jump diffusion process used to analyze contagion. Section 3.3 discusses the empirical analysis and Section 3.4 concludes.

3.2 The Hawkes jump diffusion process

3.2.1 Asset return dynamics

I consider the bivariate version of the Hawkes jump diffusion process of Aït-Sahalia et al. (2015), since it is more tractable. In particular, country 1 is always assumed to be Greece while country 2 changes for a total of three models, i.e. one for each pair (Greece-Italy, Greece-Spain and Greece-Denmark).
In the Hawkes jump diffusion process, asset log-returns for each country follow a semimartingale dynamics and consist in a drift term, a volatility dynamics and mutually exciting jumps:

$$\begin{aligned} dX_{1t} &= \mu_1 dt + \sigma_1 dW_{1t} + Z_{1t} dN_{1t} \\ dX_{2t} &= \mu_2 dt + \sigma_2 dW_{2t} + Z_{2t} dN_{2t} \end{aligned} \tag{3.1}$$

where W_{1t} and W_{2t} are standard Brownian motions with a constant correlation coefficient and N_{1t} and N_{2t} are Hawkes processes defined by their intensity processes λ_{1t} and λ_{2t}. These

intensities describe the conditional mean jump rate per unit of time:

$$\begin{cases} \mathbb{P}\left(N_{it+\Delta} - N_{it} = 0 | \mathcal{F}_t\right) = 1 - \lambda_{it}\Delta + o(\Delta) \\ \mathbb{P}\left(N_{it+\Delta} - N_{it} = 1 | \mathcal{F}_t\right) = \lambda_{it}\Delta + o(\Delta) \\ \mathbb{P}\left(N_{it+\Delta} - N_{it} > 1 | \mathcal{F}_t\right) = o(\Delta) \end{cases} \tag{3.2}$$

where $\mathcal{F}_t$ is the information set until time t, and $i = 1, 2$.
Contagion is defined as the transmission of extreme shocks to economies. I assume that, λ_{1t} and λ_{2t} have a time-varying dynamics:

$$\begin{aligned} d\lambda_{1t} &= \alpha_1(\lambda_\infty - \lambda_{1t})dt + \beta_{11}dN_{1t} + \beta_{12}dN_{2t} \\ d\lambda_{2t} &= \alpha_2(\lambda_\infty - \lambda_{2t})dt + \beta_{21}dN_{1t} + \beta_{22}dN_{2t} \end{aligned} \tag{3.3}$$

Each jump intensity is stochastic, mean-revearting and depends on the path of the past jumps both in its own country and in the other countries. Jumps are mutually dependent both in space (across countries) and in time (across various business days in a single country). These features are defined by Aït-Sahalia and co-authors as *self-* (or *time*) and *cross-* (or *space*) *excitation*. A crisis is made up of jumps and it is amplified in time and across markets. Hence, the model produces clusters of jumps over time in each country and generates successive jumps across European markets through the time-varying and predictable jump intensity. For example, let's consider two countries; once a jump occurs in country 1, the arrival rate of jumps in this country moves up in response to the most recent jump by β_{11} and then decays back towards the level $\lambda_{1\infty}$ at speed α_1. At the same time, the arrival rate of jumps of country 2 jumps up by β_{21} and then decays exponentially back. The cross-excitation between markets could be asymmetric and it is signaled by the magnitude of β_{ij} for $i \neq j$. This mutually-excitation makes the model appealing to measure the financial contagion on the banking market, focusing on *time* and *cross* dimensions.
Finally, Z_{1t} and Z_{2t} are jump sizes, cross-sectionally and serially independently distributed. The distribution of the jumps size is not specified and only the moments are parameterized assuming the following cumulative probability distribution:

$$F_{Z_i}(x) = \begin{cases} p_i e^{-\gamma_{i1}(-x)} & -\infty < x \leq 0, \quad i = 1, 2 \\ p_i + (1 - p_i)(1 - e^{-\gamma_{i2}x}) & 0 < x < \infty, \quad i = 1, 2 \end{cases} \tag{3.4}$$

where γ_{i1} and γ_{i2} are strictly positive and p_i are probabilities. Since the moments of the process are provided as functions of the generic moments of the jump size Z_t, the law F_{Z_i} allows to express the generic moment of order k as:

$$E\left(Z_i^k\right) = (-1)^k \frac{k! p_i}{\gamma_{1i}^k} + \frac{k!(1-p_i)}{\gamma_{i2}^k} \tag{3.5}$$

3.2.2 Estimation procedure

The model presented in the Section 3.2.1 is estimated using the Generalized Method of Moments (GMM). Since the Hawkes jump diffusion model does not have a closed-form likelihood, Aït-Sahalia et al. (2015) derive closed-form expressions for the key moments of the returns. In particular, the authors derive explicit expressions for the moments as functions only of the observable state variable, i.e. returns, integrating out the latent state variables, i.e. point processes N_{it} and the intensity processes λ_{it}, for $i = 1, 2$.
The moment conditions specified for the estimation are:

$$\begin{cases} \mathbb{E}\left[\Delta X_{it}\right], & i = 1,2 \\ \mathbb{E}\left[\left(\Delta X_{it} - \mathbb{E}\left(\Delta X_{it}\right)\right)^r\right], & r = 2,\dots,4 \\ \mathbb{E}\left[\Delta X_{it}\Delta X_{jt} - \mathbb{E}\left[\Delta X_{it}\right]\mathbb{E}\left[\Delta X_{jt}\right]\right], & i \neq j, \quad i,j = 1,2 \\ \mathbb{E}\left[\left(\Delta X_{it+\tau}\right)^r \left(\Delta X_{jt}\right)^r - \mathbb{E}\left[\left(\Delta X_{it}\right)^r\right]\mathbb{E}\left[\left(\Delta X_{jt}\right)^r\right]\right], & r = 1,2, \quad \tau = 1,2,3,4 \end{cases} \tag{3.6}$$

For ease of exposition I do not report the final expression of these moments and I refer to Aït-Sahalia et al. (2015), in particular Section 3.3 and Appendix B.
The moment condiitons are specified as difference between the corresponding sample moment of the log-returns and its closed-form expression derived under the model. This ensures that the GMM estimator, presented below, is consistent.
Let θ be the vector of parameters to be estimated. The GMM estimator $\hat{\theta}_T$ is the value of $\theta \in \Theta$ that minimizes the quadratic form:

$$\hat{\theta}_T := \arg\min_{\theta} \quad g_T(y,\Delta,\theta)' W_T g_T(y,\Delta,\theta) \tag{3.7}$$

where $g_T(y,\Delta,\theta)$ represents the sample moment restrictions. Furthermore, W_T is an $M \times M$ positive definite weighting matrix where M is the number of moments used in the estimation procedure, chosen greater than the number of parameters.

During the estimation I set $W_T \neq I$ and I put extra weights in the third and fourth moments to make them relatively comparable in magnitude to the other moments. Specifically, the weights are chosen to be inversely related to the variance of the corresponding sample moments. Under standard regularity conditions, as $T \to \infty$, $\hat{\theta}_T$ is asymptotically normal with asymptotic variance Ω equal to:

$$\Omega^{-1} := \Delta^{-1} D'WD \left(D'WSWD\right)^{-1} D'WD \tag{3.8}$$

where D is the gradient of the moment conditions and S is the asymptotic variance of the sample moment function. If W_T is chosen optimally, $W = S^{-1}$, then $\Omega^{-1} = \Delta^{-1} D' S^{-1} D$.
A nonnegative estimator for S is:

$$\hat{S}_T = \hat{\Gamma}_{0,T} + \sum_{v=1}^{q} \left(1 - \frac{v}{q+1}\right) \left(\hat{\Gamma}_{v,T} + \hat{\Gamma}'_{v,T}\right) \tag{3.9}$$

where

$$\hat{\Gamma}_{v,T} = \frac{1}{N} \sum_{n=v+1}^{N} h(y_{n\Delta}, \Delta, \tilde{\theta}) h(y_{(n-v)\Delta}, \Delta, \tilde{\theta})' \tag{3.10}$$

Here $\tilde{\theta}$ is an initial consistent estimate of θ_0 which is obtained by minimizing the quadratic form in the equation (3.7) with $W_T \neq I$.

3.3 Empirical analysis

3.3.1 Data

In order to asses the direction and the size of financial contagion transmission among European banking markets, I use the Datastream Bank Sector Index for both *Eurozone* and *no-Eurozone* countries. For all countries, the sample starts on January 4, 1988 and ends on January 31, 2017. From Table 3.1, Greece and Italy exhibit negative mean returns, indicating downward movements for the price Index. The kurtosis is much larger than that for the standard Gaussian distribution and the skewness for Greece and Italy is negative. Jumps in the Bank Sector Index can generate large kurtosis and if the mean jumps in price are negative they can cause negative skewness.

Figure 3.1 plots Bank Sector Index returns for Greece, Italy, Spain and Denmark from January 1, 2009 to December 31, 2010. The period displayed represents the period of highest turmoil on all European markets. In fact, the beginning of the Greek debt crisis can be identified in late 2009 while 2010 is the year in which the most important monetary policy interventions

Table 3.1: Summary Statistics Datastream Bank Sector Index. This table reports summary statistics for the log-returns of the Datastream Bank Sector Index for Greece, Italy and Spain.The sample period considered is from January 4, 1988 to January 30, 2017 for all countries.

Country	Mean	Std.Dev.	Skew.	Kurt.	N. Obs.
Greece (GR)	-0.00064	0.0308	-0.829	20.03	7585
Italy (IT)	-3.2e-06	0.0177	-0.395	12.97	7585
Spain (SP)	0.00011	0.0168	0.013	14.40	7585
no-Eurozone					
Denmark (DEN)	0.00028	0.0146	0.003	11.60	7585

were done. Generally papers that study the Greek debt crisis and its spreading around Europe, consider this limited time period. In this paper, I look at a longer time period since contagion is captured through mutually exciting jumps processes

3.3.2 Parameter estimates

Exacts daily means, autocovariances and cross-covariances of squared returns, third and fourth moments and autocovariances and cross-covariances over 4 daily-lags [4] represent the moment conditions used to identify and to estimate, the parameters of model (3.1)-(3.3). Following the Monte Carlo simulation results in Aït-Sahalia et al. (2015), parameters are restricted as follows: $\alpha_1 = \alpha_2$, $\lambda_{1\infty} = \lambda_{2\infty} = \lambda_\infty$ and $\gamma_{1i} = \gamma_{2i} = \gamma_i$. Moreover, γ_i and p_i are pre-identified: the former is identified computing the average of absolute returns that exceed 2% (price movements of this magnitude are considered as jumps) while the latter is derived from the ratio of the third and fourth moments. These two parameters do not enter θ and are treated as known and fixed during the estimation. Table 3.2 records the estimates and the standard errors of the parameters in θ for all countries.

Contagion in this context is defined as the within and between country transmission of shocks. In other words, if contagion is present, the beta coefficients should be large and significantly different from zero. Specifically, β_{11} and β_{22} measure the time dimension of contagion, i.e. the degree of self-excitation in each country, while β_{12} and β_{21} represent the level of cross-excitation from country 1 (Greece) to country 2 and from country 2 to country 1, respectively.

[4]The autocovariances and the cross-covariances are chosen looking at the autocorrelograms and cross-correlograms for each pair of countries. Visual inspection gives insights about the degree of excitation and the fourth represents the highest significant correlation among all pairs.

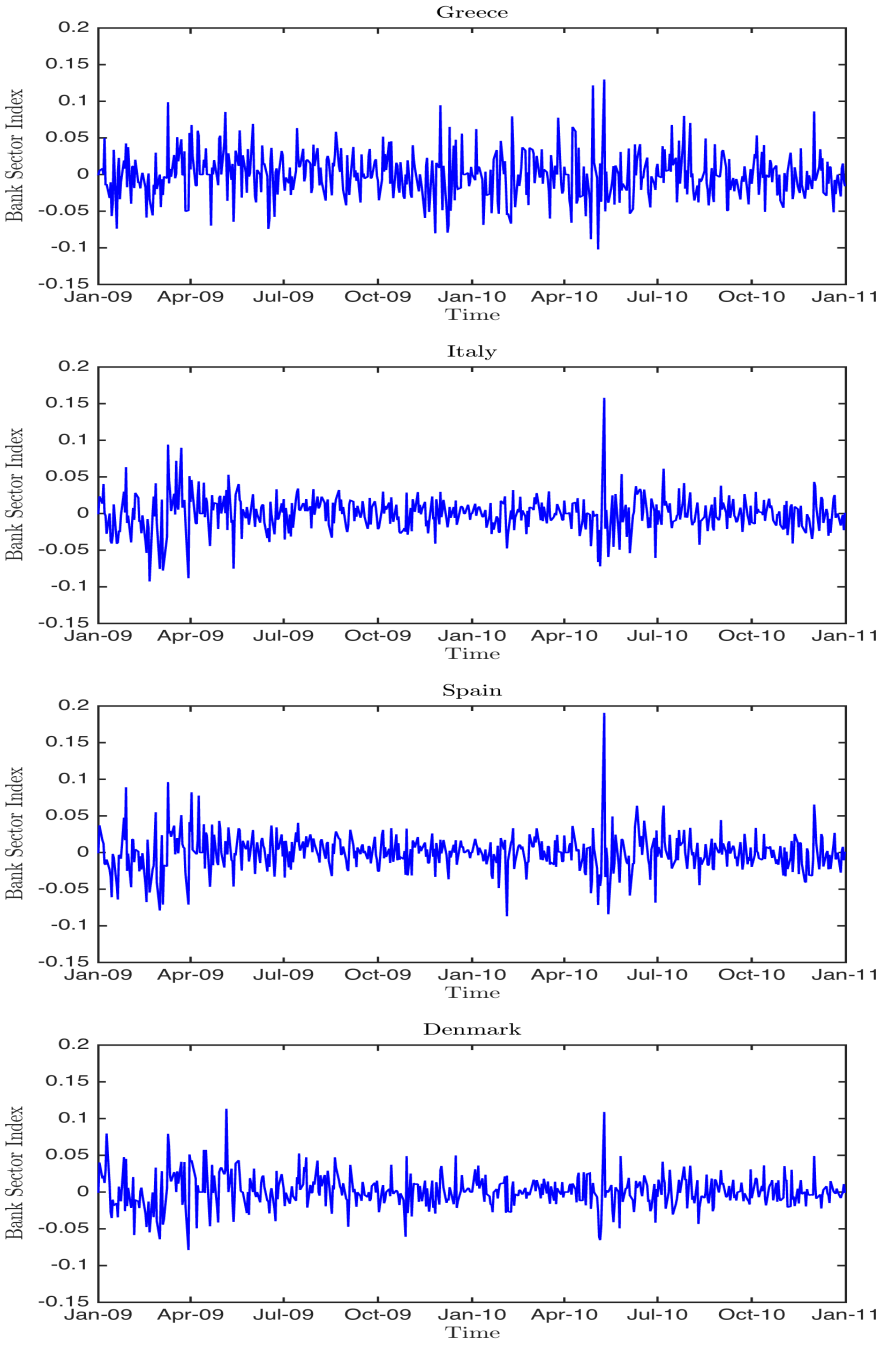

Figure 3.1: This figure shows the the Bank Sector Index returns for *Eurozone* (Greece, Italy and Spain) and for *no-Eurozone* (Denmark). The sample period is from January 1, 2009 to December 31, 2010, which is the period of the explosion of the Greek sovereign debt crisis.

Looking at the first dimension of contagion, jumps in Greece are transmitted to the domestic banking market generating self-excitation, i.e the probability of future jumps is higher. The same evidence is reported for other countries: β_{22} is significantly different from zero. Through β_{12} and β_{21}, is possible to understand the direction of contagion transmission between Greece

Table 3.2: Parameter estimates for the bivariate Hawkes jump diffusion process. This table reports the GMM estimates for the parameters of the bivariate Hawkes jump diffusion process. Standard errors are in parenthesis. The sample period considered is from January 4, 1988 to January 30, 2017 for all countries.

	Eurozone		*no-Eurozone*
Country 1 Country 2	Greece Italy	Greece Spain	Greece Denmark
α	4,1261 (0,0021)	12,045 (0,1423)	1,9950 (0,0013)
β_{11}	4,0958 (0,0088)	11,821 (0,1004)	1,9888 (0,0012)
β_{12}	0,0267 (0,0363)	0,0000 (0,3739)	0,0000 (0,0210)
β_{21}	0,0813 (0,0357)	0,2006 (0,1016)	0,0071 (0,0038)
β_{22}	3,6032 (0,0079)	10,039 (0,3407)	1,8705 (0,0022)
λ_1	46,641 (0,0004)	46,641 (0,0043)	46,641 (0,0001)
λ_2	9,4709 (0,0019)	9,8656 (0,0205)	4,9654 (0,0005)
λ_∞	0,2811 (0,0644)	0,8668 (0,2335)	0,1440 (0,0168)
σ_1	0,2092 (0,0831)	0,2095 (0,1024)	0,2107 (0,0915)
σ_2	0,2485 (0,0247)	0,2308 (0,0091)	0,2222 (0,0132)
ρ	0,8071 (0,0033)	0,8322 (0,4419)	0,6500 (0,0028)
μ_1	0,3431 (0,7406)	0,3431 (0,7404)	0,3431 (0,7391)
μ_2	0,0766 (0,3456)	0,0246 (0,4750)	0,0704 (0,3266)
$1/\gamma_1$	0,0448	0,0448	0,0448
$1/\gamma_2$	0,0345	0,0333	0,0326
p_1	0,6207	0,6207	0,6207
p_2	0,6186	0,4964	0,4990

and each other country: a difference in terms of estimated coefficients signals an asymmetric form of contagion. According to the evidence in Table 3.2, Greece affects the other countries but it is not affected by the other European countries. Indeed it is possible to answer to the first question of this paper saying that there is a clear evidence of contagion from Greece to

European countries. On the one hand, jumps in Greece cause a significant increase in other countries' jump intensity. On the other hand, jumps in all other countries do not increase the probability of having jumps in Greece. Furthermore, as expected, there is evidence of self-contagion in all countries.

In this paper, the model for Greece and Denmark is estimated in order to undestand if the contagion is mainly due to being a member of a currency union or for regional proximity. Referring to the Basel Accords we could expect that systemic risk should be less relevant for these European countries with respect to those that belongs to the *Eurozone*. The evidence from Table 3.2 does not confirm this conjecture: the contagion spread also in the *no-Eurozone* country and the self- and cross-excitation is comparable to those observed for *Eurozone*, in terms of statistical significance while the estimated value of β_{21} for Denmark is smaller. This means that *no-Eurozone* countries were affected by Greek crisis but were less exposed to contagion with respect to *Eurozone* countries.

3.3.3 Analyzing jump intensities

An interesting feature of the model presented in Aït-Sahalia et al. (2015) is represented by the jump intensity of the Hawkes process. In particular, λ_{it} provides information on the probability of having jumps in country i at each point in time. Since the jump intensity for country i depends on the jumps that take place both in country i and in country j, λ_{it} (as well as λ_{jt}) gives insights into the interaction over time between the two countries. In other words, if λ_{it} is high then the probability of having jumps in country i increases, and if contagion spreads, then also λ_{jt} increases. I use the parameter estimates displayed in Table 3.2 to compute each jump intensity and I assume, coherently with the identification of γ_i, that a jump takes place in a country if, at a given point in time, the observed absolute return for that country is greater than 2%.

Figure 3.2 plots the jump intensity for the *Eurozone* pairs of countries analyzed in this paper, from January 1, 2007 to January 30, 2017. The left panel shows both jump intensities for Greece and Italy while the right panel plots the jump intensities for Greece and Spain. It is evident that crisis in Greece affects the Italian and the Spanish banking markets. Moreover, referring to the estimates recorded in Table 3.2, the contagion from Greece to Spain is stronger than the contagion from Greece to Italy. In all cases, the two intensities follow the same evolution during the sovereign Greek crisis (2010-2014) while, before and after the crisis co-movements, still present, seem to be less evident. At odds, the evolution of the jump intensity of Denmark

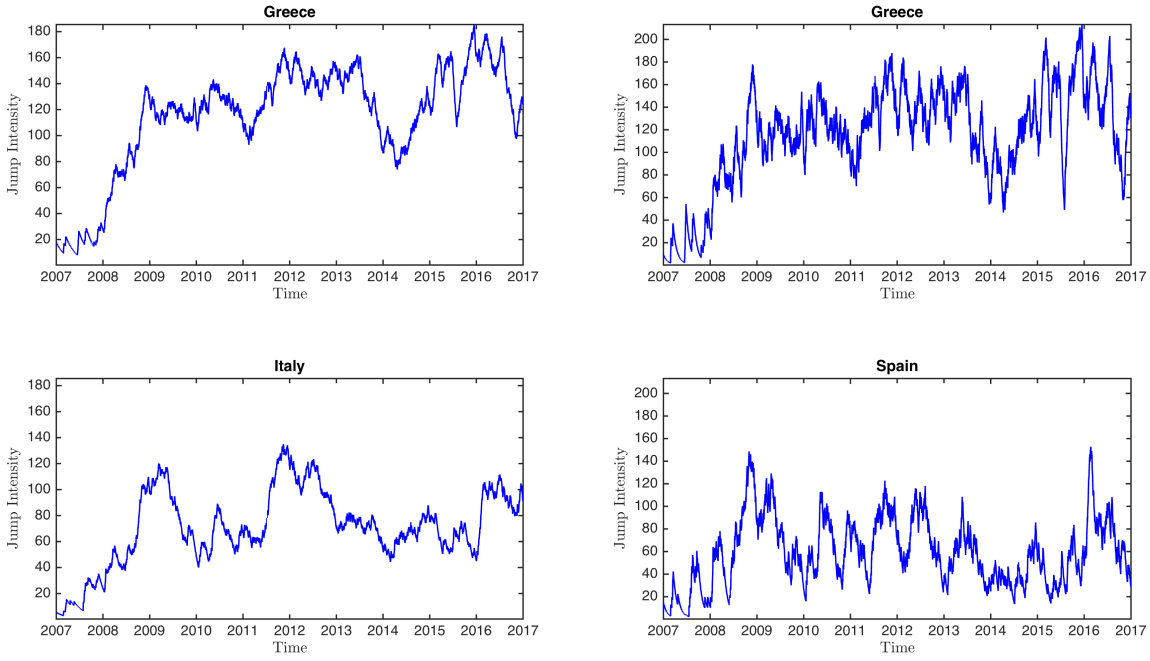

Figure 3.2: This figure shows the jump intensity for the *Eurozone* pairs of countries (Greece vs Italy and Greece vs Spain). The sample period is from January 1, 2007 to January 30, 2017. The parameters used to obtain all jump intensities are those displayed in Table 3.2.

does not perfectly mimic that of Greece, as displayed in Figure 3.3. In fact, looking at the estimates in Table 3.2, the jump intensity of Denmark is mainly due to its own jumps.

3.3.4 Contagion: rolling estimation

The beta parameters trace the contagion effect across banking markets on two dimensions: cross- and time- excitation. First I focus on the former dimension in order to understand how the crisis in Greece affects other European countries. In particular, I report the time evolution of β_{21} which indicates the contagion from the Greek bank market (country 1) to the other European countries (country to 2). In order to understand how β_{21} parameter behaves over time, I perform a monthly rolling window estimation. Namely, the in-sample period starts on January 4, 1988 and ends on May 20, 2007 while the out-of-sample period starts on May 21, 2007 and ends on January 30, 2017. Indeed, the one-month coefficients are computed using the in-sample period. Then the in-sample period is moved forward by one period (one month or 22 days) and the estimation is run again. This procedure is repeated step by step until the end of the sample. Notice that in model (3.1)- (3.3) mutual excitation is also present during "normal" market conditions because of the interconnection among banking markets. But if the connec-

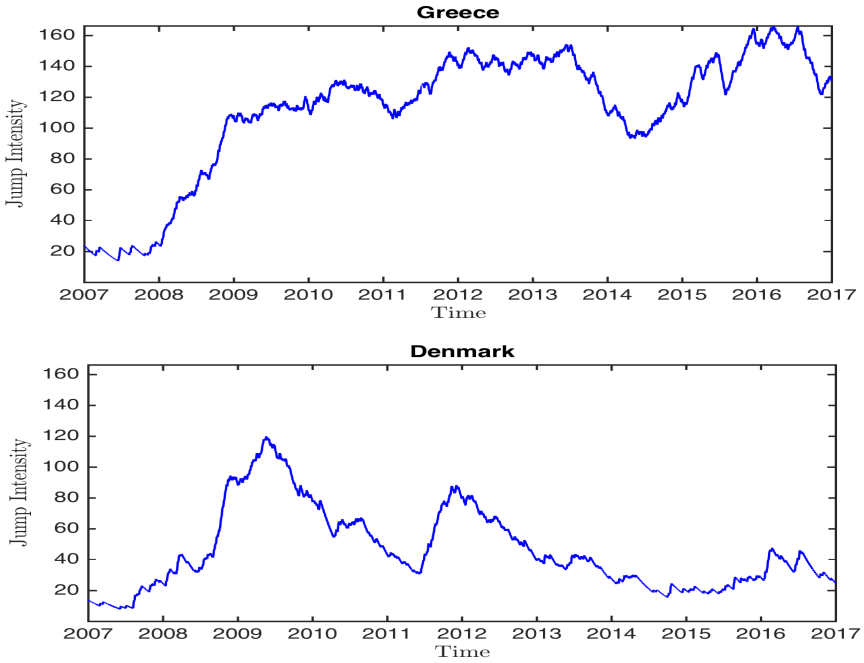

Figure 3.3: This figure shows the jump intensity for the *no-Eurozone* pair of countries (Greece vs Denamrk). The sample period is from January 1, 2007 to January 30, 2017. The parameters used to obtain both jump intensities are those displayed in Table 3.2.

tion becomes stronger due to the contagion effect, then β_{21} should increase, causing a rise in the probability of future jumps, at least for the period in which unexpected news propagate in Europe. The left panel of Figure 3.4 plots the rolling estimates of β_{21} for Italy and Spain. The cross-excitation parameter changes over time capturing the domestic market reactions to Greek news. Specifically, β_{21} records some changes from 2009 to 2010 for both countries, indicating that news about a possible crisis in Greece affected European markets generating panic about domestic financial stability. Moreover, β_{21} reaches a peak at the end of 2011 for both countries. The peak can be due to important decisions of the European Central Bank (ECB) about measures to enhance the response to the crisis in Greece and around Europe. For example, on October 27, 2011, European leaders agreed on a comprehensive package of measures focused on Greece and Europe and on measures strengthening the governance of the euro area and the budgetary discipline. Finally, in the right panel of Figure 3.4, I plot the behaviour of the time-excitation parameter for Italy and Spain. β_{22} measures the effect of domestic news on the probability of having future jumps in the domestic country/market. The effect of the US subprime crisis is evident from the peak across 2008-2009 for both countries. Italy records a rise in the time-contagion at the end of 2011 (e.g. December 4, 2001 Mario Monti reveals a comprehensive package of measures to ensure Italy's stability and to enhance growth). Furthermore,

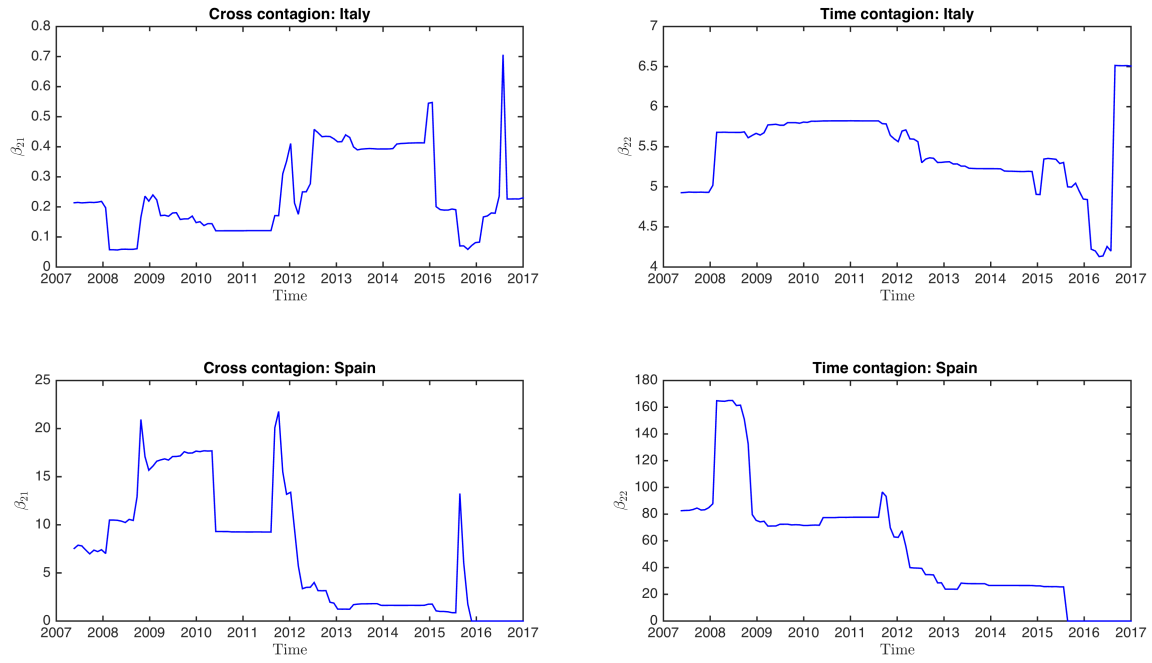

Figure 3.4: This figure shows the behaviour of the cross- and time- excitation coefficients for Italy and Spain.The out-of-sample period is from May 20, 2007 to January 30, 2017. The parameters are estimated using the rolling procedure described in Section 3.3.4

β_{22} registers an increase at the beginning of 2012 for both Italy and Spain. For example, on January 13, 2012 Standard and Poor's carries out a sweeping downgrade Italian and Spanish sovereigns which are stripped of A- and A status; at February 3, 2012 the Spanish government adopts a new set of measures to reform and strengthen the domestic banking market. Figure 3.5 plots the rolling estimates of β_{21} and β_{22} for Denmark. The former records some increases (see 2012) but they are very small: the effect of unexpected news in Greece affect Denmark but the impact is very low. Instead, β_{22} registers larger changes, due to unexpected news about the Danish economy.

A quantitative assessment of the effect of different news (good or bad, about interventions by ECB, IMF or politicians speech) on contagion is outside the scope of this paper and is left to future research.

3.4 Conclusion and direction of future works

This paper sheds some light on financial contagion during the European debt crisis at the bank level. Specifically, this paper aims at understanding whether the crisis in Greece spread in

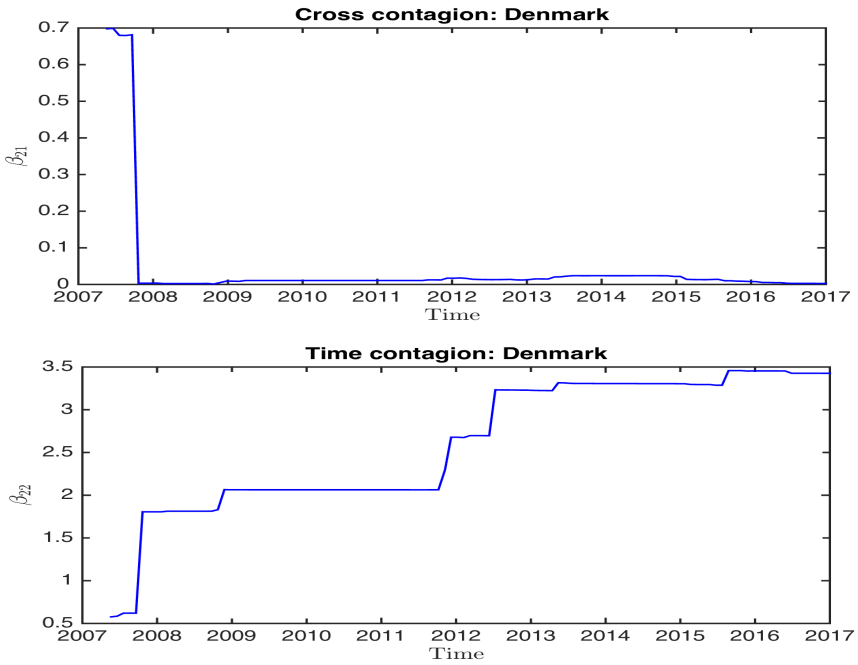

Figure 3.5: This figure shows the behaviour of the cross- and time- excitation coefficients for Denmark.The out-of-sample period is from May 20, 2007 to January 30, 2017. The parameters are estimated using the rolling procedure described in Section 3.3.4

other European countries and whether contagion, if present, is mainly due to being a member of the European monetary union or to regional proximity, looking at the propagation of shocks at the bank level. In order to measure the direction and the size of contagion transmission between banking markets, I consider the Datastream Bank Sector Index for Greece, Italy, and Spain and I estimate the discrete time multivariate Hawkes processes proposed by Aït-Sahalia et al. (2015) for several European countries. The empirical evidence suggests that unexpected news in Greece cause a significant increase in the jump intensity of the other countries . This means that, during the Greek crisis, the European banking markets are exposed to contagion which is transmitted through inter-bank connections, especially for those countries belonging to the monetary union. Moreover, I perform a monthly rolling estimation in order to understand how the contagion parameters change over time. Through a visual inspection of the beta parameters behaviour, I can first conclude that news about the crisis in Greece affect European markets generating panic about domestic financial stability. In particular for *no-Eurozone* countries contagion is present but it is less evident; I can conclude that it is mainly due to regional proximity. Second, the main changes are linked to important announcements of interventions to control the Greek crisis but also to domestic (Italian, Spanish or Danish) news.

I plan to enlarge the analysis presented in this paper, considering several European countries

for both groups, i.e. *Eurozone* and *no-Eurozone*. Nevertheless, I think that a quantitative assessment of the effect of different news on contagion can have important monetary policy implications. In this sense, an event study can be done in order to understand if specific interventions to limit the crisis propagation around Europe generating a decrease in the contagion or not. I leave these questions to future research.

Bibliography

Acharya, V., I. Drechsler, and P. Schnabl (2014). A pyrrhic victory? bank bailouts and sovereign credit risk. *The Journal of Finance 69*(6), 2689–2739.

Aït-Sahalia, Y., J. Cacho-Diaz, and R. J. Laeven (2015). Modeling financial contagion using mutually exciting jump processes. *Journal of Financial Economics 117*(3), 585–606.

Aït-Sahalia, Y. and J. Jacod (2010). Analyzing the spectrum of asset returns: Jump and volatility components in high frequency data. Technical report, National Bureau of Economic Research.

Aït-Sahalia, Y. and L. Mancini (2008). Out of sample forecasts of quadratic variation. *Journal of Econometrics 147*(1), 17–33.

Aït-Sahalia, Y., P. A. Mykland, and L. Zhang (2005). How often to sample a continuous-time process in the presence of market microstructure noise. *Review of Financial studies 18*(2), 351–416.

Aït-Sahalia, Y., P. A. Mykland, and L. Zhang (2011). Ultra high frequency volatility estimation with dependent microstructure noise. *Journal of Econometrics 160*(1), 160–175.

Albertazzi, U., T. Ropele, G. Sene, and F. M. Signoretti (2014). The impact of the sovereign debt crisis on the activity of italian banks. *Journal of Banking & Finance 46*, 387–402.

Alitab, D., G. Bormetti, F. Corsi, and A. A. Majewski (2015). A jump and smile ride: Continuous and jump variance risk premia in option pricing. *Available at SSRN 2631155*.

Alizadeh, S., M. W. Brandt, and F. X. Diebold (2002). Range-based estimation of stochastic volatility models. *The Journal of Finance 57*(3), 1047–1091.

Allen, F., A. Babus, and E. Carletti (2009). Financial crises: theory and evidence. *Annu. Rev. Financ. Econ. 1*(1), 97–116.

Alter, A. and A. Beyer (2014). The dynamics of spillover effects during the european sovereign debt turmoil. *Journal of Banking & Finance 42*, 134–153.

Alter, A. and Y. S. Schüler (2012). Credit spread interdependencies of european states and banks during the financial crisis. *Journal of Banking & Finance 36*(12), 3444–3468.

Andersen, T. G. and T. Bollerslev (1998). Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. *International economic review*, 885–905.

Andersen, T. G., T. Bollerslev, P. F. Christoffersen, and F. X. Diebold (2012). Financial risk measurement for financial risk management. Technical report, National Bureau of Economic Research.

Andersen, T. G., T. Bollerslev, and F. X. Diebold (2007). Roughing it up: Including jump components in the measurement, modeling, and forecasting of return volatility. *The Review of Economics and Statistics 89*(4), 701–720.

Andersen, T. G., T. Bollerslev, F. X. Diebold, and H. Ebens (2001a). The distribution of realized stock return volatility. *Journal of financial economics 61*(1), 43–76.

Andersen, T. G., T. Bollerslev, F. X. Diebold, and P. Labys (2003). Modeling and forecasting realized volatility. *Econometrica 71*(2), 579–625.

Andersen, T. G., T. Bollerslev, F. X. Diebold, and C. Vega (2002). Micro effects of macro announcements: Real-time price discovery in foreign exchange. Technical report, National bureau of economic research.

Andersen, T. G., T. Bollerslev, P. Frederiksen, and M. Ørregaard Nielsen (2010). Continuous-time models, realized volatilities, and testable distributional implications for daily stock returns. *Journal of Applied Econometrics 25*(2), 233–261.

Andersen, T. G., O. Bondarenko, and M. T. Gonzalez-Perez (2011). A corridor fix for vix: developing a coherent model-free option-implied volatility measure. In *2nd Osaka Conference on „ÄúHigh-Frequency Data Analysis in Financial Markets„Äù Osaka*.

Ang, A. and F. A. Longstaff (2013). Systemic sovereign credit risk: Lessons from the us and europe. *Journal of Monetary Economics 60*(5), 493–510.

Angeloni, C. and G. B. Wolff (2012). Are banks affected by their holdings of government debt? Technical report, Bruegel working paper.

Arezki, R., B. Candelon, and A. N. R. Sy (2011). Sovereign rating news and financial markets spillovers: Evidence from the european debt crisis.

Audrino, F. and Y. Hu (2016). Volatility forecasting: Downside risk, jumps and leverage effect. *Econometrics 4*(1), 8.

Bai, J., C. Julliard, and K. Yuan (2012). Eurozone sovereign bond crisis: Liquidity or fundamental contagion. *Federal Reserve Bank of New York Working Paper*.

Bakshi, G., C. Cao, and Z. Chen (1997). Empirical performance of alternative option pricing models. *The Journal of finance 52*(5), 2003–2049.

Bandi, F. M. and R. Reno (2016). Price and volatility co-jumps. *Journal of Financial Economics 119*(1), 107–146.

Barndorff-Nielsen, O. E. (2002). Econometric analysis of realized volatility and its use in estimating stochastic volatility models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology) 64*(2), 253–280.

Barndorff-Nielsen, O. E. and N. Shephard (2004). Power and bipower variation with stochastic volatility and jumps. *Journal of financial econometrics 2*(1), 1–37.

Barone-Adesi, G., R. F. Engle, and L. Mancini (2008). A garch option pricing model with filtered historical simulation. *Review of Financial Studies 21*(3), 1223–1258.

Bates, D. S. (1996). Jumps and stochastic volatility: Exchange rate processes implicit in deutsche mark options. *Review of financial studies 9*(1), 69–107.

Bates, D. S. (2000). Post-'87 crash fears in the s&p 500 futures option market. *Journal of Econometrics 94*(1), 181–238.

Bates, D. S. (2012). Us stock market crash risk, 1926–2010. *Journal of Financial Economics 105*(2), 229–259.

Bedendo, M. and P. Colla (2015). Sovereign and corporate credit risk: Evidence from the eurozone. *Journal of Corporate Finance 33*, 34–52.

Beirne, J. and M. Fratzscher (2013). The pricing of sovereign risk and contagion during the european sovereign debt crisis. *Journal of International Money and Finance 34*, 60–82.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of econometrics 31*(3), 307–327.

Bollerslev, T. and V. Todorov (2011). Tails, fears, and risk premia. *The Journal of Finance 66*(6), 2165–2211.

Boswijk, H., R. LAEVEN, and X. Yang (2014). Testing for self-excitation in jumps. Technical report, Working Paper, University of Amsterdam.

Brownlees, C. T. and G. M. Gallo (2010). Comparison of volatility measures: a risk management perspective. *Journal of Financial Econometrics 8*(1), 29–56.

Caporin, M., L. Pelizzon, F. Ravazzolo, and R. Rigobon (2013). Measuring sovereign contagion in europe. Technical report, National Bureau of Economic Research.

Caporin, M., E. Rossi, P. S. de Magistris, et al. (2015). Volatility jumps and their economic determinants. *Journal of Financial Econometrics 14*(1), 29–80.

Chan, W. H. and J. M. Maheu (2002). Conditional jump dynamics in stock market returns. *Journal of Business & Economic Statistics 20*(3), 377–389.

Chiu, C.-L., M.-C. Lee, and J.-C. Hung (2005). Estimation of value-at-risk under jump dynamics and asymmetric information. *Applied Financial Economics 15*(15), 1095–1106.

Christensen, K., R. C. Oomen, and M. Podolskij (2014). Fact or friction: Jumps at ultra high frequency. *Journal of Financial Economics 114*(3), 576–599.

Christoffersen, P., R. Elkamhi, B. Feunou, and K. Jacobs (2009). Option valuation with conditional heteroskedasticity and nonnormality. *Review of Financial studies*, hhp078.

Christoffersen, P., B. Feunou, and Y. Jeon (2015). Option valuation with observable volatility and jump dynamics. *Journal of Banking & Finance 61*, S101–S120.

Christoffersen, P., S. Heston, and K. Jacobs (2013). Capturing option anomalies with a variance-dependent pricing kernel. *Review of Financial Studies 26*(8), 1963–2006.

Christoffersen, P., K. Jacobs, C. Ornthanalai, and Y. Wang (2008). Option valuation with long-run and short-run volatility components. *Journal of Financial Economics 90*(3), 272–297.

Christoffersen, P. F. (1998). Evaluating interval forecasts. *International economic review*, 841–862.

Christoffersen, P. F. and F. X. Diebold (2000). How relevant is volatility forecasting for financial risk management? *Review of Economics and Statistics 82*(1), 12–22.

Clements, M. P., A. B. Galvão, and J. H. Kim (2008). Quantile forecasts of daily exchange rate returns from forecasts of realized volatility. *Journal of Empirical Finance 15*(4), 729–750.

Correa, R., H. Sapriza, and A. Zlate (2013). Liquidity shocks, dollar funding costs, and the bank lending channel during the european sovereign crisis.

Corsi, F. (2009). A simple approximate long-memory model of realized volatility. *Journal of Financial Econometrics*, nbp001.

Corsi, F., N. Fusari, and D. La Vecchia (2013). Realizing smiles: Options pricing with realized volatility. *Journal of Financial Economics 107*(2), 284–304.

Corsi, F., D. Pirino, and R. Reno (2010). Threshold bipower variation and the impact of jumps on volatility forecasting. *Journal of Econometrics 159*(2), 276–288.

Corsi, F. and R. Renó (2009). Har volatility modelling with heterogeneous leverage and jumps. *Available at SSRN 1316953*.

Corsi, F. and R. Renò (2012). Discrete-time volatility forecasting with persistent leverage effect and the link with continuous-time volatility modeling. *Journal of Business & Economic Statistics 30*(3), 368–380.

Creal, D., S. J. Koopman, and A. Lucas (2013). Generalized autoregressive score models with applications. *Journal of Applied Econometrics 28*(5), 777–795.

De Bruyckere, V., M. Gerhardt, G. Schepens, and R. Vander Vennet (2013). Bank/sovereign risk spillovers in the european debt crisis. *Journal of Banking & Finance 37*(12), 4793–4809.

Diebold, F. X. and R. S. Mariano (2012). Comparing predictive accuracy. *Journal of Business & economic statistics*.

Duan, J.-C. et al. (1995). The garch option pricing model. *Mathematical finance 5*(1), 13–32.

Duan, J.-C., P. Ritchken, and Z. Sun (2006). Approximating garch-jump models, jump-diffusion processes, and option pricing. *Mathematical Finance 16*(1), 21–52.

Duffie, D., J. Pan, and K. Singleton (2000). Transform analysis and asset pricing for affine jump-diffusions. *Econometrica 68*(6), 1343–1376.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica: Journal of the Econometric Society*, 987–1007.

Engle, R. F. and S. Manganelli (2004). Caviar: Conditional autoregressive value at risk by regression quantiles. *Journal of Business & Economic Statistics 22*(4), 367–381.

Eraker, B. (2004). Do stock prices and volatility jump? reconciling evidence from spot and option prices. *The Journal of Finance 59*(3), 1367–1403.

Eraker, B., M. Johannes, and N. Polson (2003). The impact of jumps in volatility and returns. *The Journal of Finance 58*(3), 1269–1300.

Fabozzi, F. J., R. Giacometti, and N. Tsuchida (2016). Factor decomposition of the eurozone sovereign cds spreads. *Journal of International Money and Finance 65*, 1–23.

Fang, F. and C. W. Oosterlee (2008). A novel pricing method for european options based on fourier-cosine series expansions. *SIAM Journal on Scientific Computing 31*(2), 826–848.

Fulop, A., J. Li, and J. Yu (2014). Self-exciting jumps, learning, and asset pricing implications. *Review of Financial Studies*, hhu078.

Gagliardini, P., C. Gourieroux, and E. Renault (2011). Efficient derivative pricing by the extended method of moments. *Econometrica 79*(4), 1181–1232.

Gençay, R., F. Selçuk, and A. Ulugülyağci (2003). High volatility, thick tails and extreme value theory in value-at-risk estimation. *Insurance: Mathematics and Economics 33*(2), 337–356.

Ghysels, E., P. Santa-Clara, and R. Valkanov (2004). The midas touch: Mixed data sampling regression models. *Finance*.

Giot, P. and S. Laurent (2004). Modelling daily value-at-risk using realized volatility and arch type models. *journal of Empirical Finance 11*(3), 379–398.

Giot, P., S. Laurent, and M. Petitjean (2010). Trading activity, realized volatility and jumps. *Journal of Empirical Finance 17*(1), 168–175.

Gouriéroux, C. and J. Jasiak (2006). Autoregressive gamma processes. *Journal of Forecasting 25*(2), 129–152.

Gourieroux, C. and A. Monfort (2007). Econometric specification of stochastic discount factor models. *Journal of Econometrics 136*(2), 509–530.

Gross, M. and C. Kok (2013). Measuring contagion potential among sovereigns and banks using a mixed-cross-section gvar.

Hampel, F. R., E. M. Ronchetti, P. J. Rousseeuw, and W. A. Stahel (2011). *Robust statistics: the approach based on influence functions*, Volume 114. John Wiley & Sons.

Hansen, L. P. (1982). Large sample properties of generalized method of moments estimators. *Econometrica: Journal of the Econometric Society*, 1029–1054.

Hansen, P. R. and A. Lunde (2006). Realized variance and market microstructure noise. *Journal of Business & Economic Statistics 24*(2), 127–161.

Harvey, A. and A. Luati (2014). Filtering with heavy tails. *Journal of the American Statistical Association 109*(507), 1112–1122.

Heston, S. L. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of financial studies 6*(2), 327–343.

Heston, S. L. and S. Nandi (2000). A closed-form garch option valuation model. *Review of Financial Studies 13*(3), 585–625.

Heyde, C., S. Kou, and X. Peng (2006). What is a good risk measure: Bridging the gaps between data, coherent risk measures and insurance risk measures. *Preprint, Columbia University*.

Huang, X. and G. Tauchen (2005). The relative contribution of jumps to total price variance. *Journal of financial econometrics 3*(4), 456–499.

Hung, J.-C., M.-C. Lee, and H.-C. Liu (2008). Estimation of value-at-risk for energy commodities via fat-tailed garch models. *Energy Economics 30*(3), 1173–1191.

Jacod, J. and V. Todorov (2009). Testing for common arrivals of jumps for discretely observed multidimensional processes. *The Annals of Statistics*, 1792–1838.

Kou, S. G. (2002). A jump-diffusion model for option pricing. *Management science 48*(8), 1086–1101.

Kupiec, P. H. (1995). Techniques for verifying the accuracy of risk measurement models. *THE J. OF DERIVATIVES 3*(2).

Liu, L. Y., A. J. Patton, and K. Sheppard (2015). Does anything beat 5-minute rv? a comparison of realized measures across multiple asset classes. *Journal of Econometrics 187*(1), 293–311.

Longstaff, F. A., J. Pan, L. H. Pedersen, and K. J. Singleton (2011). How sovereign is sovereign credit risk? *American Economic Journal: Macroeconomics 3*(2), 75–103.

Maheu, J. M. and T. H. McCurdy (2004). News arrival, jump dynamics, and volatility components for individual stock returns. *The Journal of Finance 59*(2), 755–793.

Majewski, A. A., G. Bormetti, and F. Corsi (2015). Smile from the past: A general option pricing framework with multiple volatility and leverage components. *Journal of Econometrics*.

Manasse, P. and L. Zavalloni (2013). Sovereign contagion in europe: Evidence from the cds market.

Maneesoonthorn, W., C. S. Forbes, and G. M. Martin (2014). Inference on self-exciting jumps in prices and volatility using high frequency measures. *arXiv preprint arXiv:1401.3911*.

Mauro, P., N. Sussman, and Y. Yafeh (2002). Emerging market spreads: then versus now. *The Quarterly Journal of Economics 117*(2), 695–733.

Merton, R. C. (1976). Option pricing when underlying stock returns are discontinuous. *Journal of financial economics 3*(1), 125–144.

Monfort, A., F. Pegoraro, J.-P. Renne, and G. Roussellet (2014a). Recursive discrete-time affine processes and asset pricing. *Technical report, mimeo.*.

Monfort, A., F. Pegoraro, J.-P. Renne, and G. Roussellet (2014b). Staying at zero with affine processes: A new dynamic term structure model. In *Paris December 2014 Finance Meeting EUROFIDAI-AFFI Paper*.

Müller, U. A., M. M. Dacorogna, R. D. Davé, R. B. Olsen, O. V. Pictet, and J. E. von Weizsäcker (1997). Volatilities of different time resolutions—analyzing the dynamics of market components. *Journal of Empirical Finance 4*(2), 213–239.

Pan, J. (2002). The jump-risk premia implicit in options: Evidence from an integrated time-series study. *Journal of financial economics 63*(1), 3–50.

Pan, J. and K. J. Singleton (2008). Default and recovery implicit in the term structure of sovereign cds spreads. *The Journal of Finance 63*(5), 2345–2384.

Parkinson, M. (1980). The extreme value method for estimating the variance of the rate of return. *Journal of Business*, 61–65.

Press, S. J. (1967). A compound events model for security prices. *Journal of business*, 317–335.

Santa-Clara, P. and S. Yan (2010). Crashes, volatility, and the equity premium: Lessons from s&p 500 options. *The Review of Economics and Statistics 92*(2), 435–451.

Todorov, V. and G. Tauchen (2011). Volatility jumps. *Journal of Business & Economic Statistics 29*(3), 356–371.

Wu, L. (2003). Jumps and dynamic asset allocation. *Review of Quantitative Finance and Accounting 20*(3), 207–243.