Università
Ca'Foscari
Venezia

Scuola Dottorale di Ateneo
Graduate School

Dottorato di ricerca
in Economia
Ciclo XXVII
Anno di discussione 2016

Price and Volatility Jumps in the Stock Market

Settore scientifico disciplinare di afferenza: SECS-P/05

Tesi di Dottorato di Patrick, Zoi, Matricola: 955703

Coordinatore del Dottorato
Prof. Giacomo Pasini

Tutore del Dottorando
Prof. Monica Billio

Co-tutore del Dottorando
Prof. Loriana Pelizzon

Price and Volatility Jumps in the Stock Market

by

Patrick, Zoi
Matricola 955703

A thesis submitted in partial fulfillment
for the degree of

Doctor of Philosophy in Economics

Department of Economics
Ca' Foscari University of Venice

Supervisors :

Prof. Monica Billio
Prof. Loriana Pelizzon

September, 2016

Declaration

I, Patrick Zoi, hereby declare that this Ph.D thesis titled "Price and Volatility Jumps in the Stock Market" is a presentation of my original research work for the degree of Doctor of Philosophy in Economics under the guidance and supervision of Prof. Monica Billio and Prof. Loriana Pelizzon. Wherever contributions of others are involved, every effort is made to indicate this clearly, with due reference to the literature, and acknowledgement of collaborative research and discussions.

I affirm that this work has not been submitted previously in the same or similar form to another examination committee at another department, university or country.

Signed:

_____

Date:

_____

Copyright © 2016 by Patrick, Zoi
All rights reserved

Studente: Patrick Zoi Matricola: 955703
Dottorato di ricerca in Economia
Ciclo XXVII
Titolo della tesi: Price and Volatility Jumps in the Stock Market

Abstract

The contributions of this thesis are discussed in three self-contained chapters.

Chapter 1 We study the bivariate jump process involving the S&P 500 and the Euro Stoxx 50 index with jumps extracted from high frequency data using non-parametric methods. Our analysis, based on a generalized Hawkes process, reveals the presence of self-excitation in the jump activity which is responsible for jump clustering but has a very small persistence in time. Concerning cross-market effects, we find statistically significant co-jumps occurring when both markets are simultaneously operating but no evidence of contagion in the jump activity, suggesting that the role of jumps in volatility transmission is negligible. Moreover, we find a negative relationship between the jump activity and the continuous volatility indicating that jumps are mostly detected during tranquil market conditions rather than in periods of stress. Importantly, our empirical results are robust under different jump detection methods.

Chapter 2 We construct a new class of non-parametric robust to jumps estimators for the realized volatility combining multiple measures applied to high frequency data. Collecting information from several estimators, this method provides a higher asymptotic efficiency and allows to improve finite sample properties. We use such combinations to construct non-parametric tests for the detection of jumps in asset prices: our Monte Carlo study shows that the new tests can achieve substantially more power compared to other common methods.

Chapter 3 We introduce a new stochastic process generalizing the Autoregressive Gamma (ARG) of Gourieroux and Jasiak (2006). This process is based on a new and more flexible specification of the conditional distribution which extends the non-central gamma preserving the same analytical tractability. We propose an empirical a application to model the realized volatility measured from high frequency data. The period of our analysis includes the sub-prime and the Euro Sovereign crisis. Our results highlight the superior performances of the new process compared the standard ARG and confirm the need of at least two stochastic factors for a satisfactory description of the volatility dynamics: one factor is responsible for small volatility fluctuations while the second factor generates large upward shocks featuring volatility jumps.

Contents

List of Tables

List of Figures

Acknowledgements

Firstly, I would like to express my sincere gratitude to my advisors Prof. Monica Billio and Prof. Loriana Pelizzon from for their fundamental guidance and constant support to my research activity and to the development of this thesis.

Besides my advisors, I would like to thank the rest of my thesis committee: Prof. Fulvio Corsi, Prof. Alain Monfort and Prof. Roberto Renò for their insightful comments and their stimulating suggestions which encouraged me to deepen my research from various perspectives.

My sincere thanks also goes to Prof. Giacomo Bormetti for stimulating fruitful discussions and providing extremely valuable comments, to Dr. Fabrizio Ferriani for his active collaboration in the development of the first chapter and for his support in various phases underlying the evolution of this work.

I thank my PhD fellows for the interesting discussions which provided the basis for the development of my research ideas.

Last but no least I would like to thank my family: my partner Nirvana, my daughter Matilde and my son Giacomo who always patiently supported me during the whole PhD.

Chapter 1

The Dynamics of Price Jumps in the Stock Market: an Empirical Study on Europe and U.S.

1.1 Introduction

Prices of traded assets are sometime subject to sudden movements which are hardly described by a continuous process. Such events are commonly referred as "*Jumps*" in order to emphasize their instantaneous impact on the asset prices. Jumps may differ in terms of sign (positive vs. negative), magnitude (large vs. small) and frequency (recurring vs rare). They are commonly associated with a sudden flow of new information but there is no general consensus on which kind of market events can more likely generate discontinuous price reactions. Calcagnile et al. (2015) find that such discontinuities are only partially related to scheduled news announcements and their occurrence is largely unpredictable, Bajgrowicz et al. (2016) instead show that common jump detection methods entail a significant risk of misclassification claiming that jumps are rare and mostly related to news announcements. Aït-Sahalia, Cacho-Diaz, and Laeven (2015) (ADL henceforth) suggest that jump are also a vehicle of contagion across worldwide markets. This happens if asset price jumps spread from an originating region to a different one.

Nowadays, financial literature considers the presence of jumps as a cardinal component of the asset price dynamics. Starting from the seminal paper of Merton (1976) who first introduced them in a diffusion model, a plenty of scholars has extensively contributed to the scientific research in this field. As highlighted by Ait-Sahalia (2004) the study of jumps is extremely relevant for investors in terms of asset allocation and portfolio optimization as large price movements may generate significant losses and encourage the demand for higher risk premia, see Liu et al. (2003), Wright

and Zhou (2009), Bates (2008), Bollerslev and Todorov (2011), Ait-Sahalia and Hurd (2015) among others. For risk management purposes jumps are important because they can generate fat tails with a significant impact on the Value at Risk (see for instance Duffie and Pan 1997 and Pan and Duffie 2001). For asset pricing, jumps are also extremely relevant since they are responsible for market incompleteness with the implication that the jump risk cannot be perfectly hedged (see Duffie et al. 2000, Eraker et al. 2003 among many others). The literature on jumps has largely benefited from the increasing availability of high frequency data which foster a copious scientific production in the field of jump detection. Some seminal contributions include Barndorff-Nielsen and Shephard (2004, 2006), Huang and Tauchen (2005), Andersen et al. (2007b) (ABD henceforth), Andersen et al. (2010a), Bollerslev et al. (2013). In these papers jumps are identified thanks to non-parametric techniques which rely on the comparison between two realized measures of volatility, one determined by continuous price changes and another one including also jumps. Several alternative volatility measures that are also robust to jumps have been proposed, for a more detailed discussion we refer to Chapter 2.

This chapter partially follows ADL who introduce a reduced-form model for asset returns that is able to capture the time clustering of jump events within the same market (self-excitation) as well as the transmission across markets (cross-excitation). These features of jumps are of particular concern for investors, regulators and policymakers: given that jumps are an important source of risk, the study of their dynamics at the multivariate level can shed some light on their simultaneous occurrence (co-jumps) as well as on their possible transmission across different markets. However, we differ from their contribution and we add to the existing literature in several ways. First of all our study is based on high frequency data and modern non-parametric jump detection methods, instead of using daily returns. Moreover, we also investigate jumps using multivariate Hawkes processes (introduced in Hawkes 1971b and Hawkes 1971a) but we rely on a generalized version of the method proposed by Bowsher (2007) which allows to introduce some additional explanatory variables. We are able to estimate the model via maximum likelihood, knowing the exact intraday jump times as well as their size and we are able to apply the specification tests described in Bowsher (2007) in addition to the standard ones. Interestingly, our conclusions are substantially different: we find no cross-excitation while self-excitation are significant but exhibit a very low persistence compared to

ADL. In this regard, our results are also complementary to the novel contribution of Bajgrowicz et al. (2016) who find no significant jump clustering effect at the daily time scale: according to our estimates such effects are characterized by a short persistence and can be measured only at short time scales once that the exact intraday jump times are identified. We stress the importance of taking into account the continuous volatility that, according to our evidence, exhibits an inverse relation with the jumps intensity. This indicates that the number of jumps detected during low-volatility periods is higher than in period of stress. Such a result, can be simply determined by a reduction of the power of the tests when the volatility is high. The decrease in the jump activity during those periods can be a mere statistical artifact rather then a true empirical evidence and we are not allowed to draw robust conclusions in this regard. Nevertheless, we have a clear and robust evidence on a decreasing relative contribution of jumps to the total price variation during the sub-prime and the Euro Sovereign crisis. This is in stark contrast with the dynamics assumed by ADL in their parametric model where price jumps characterize periods of market turmoil. We find that the continuous intraday volatility for the S&P 500 measured from high frequency data reached an annualized level above 120% just a few days before the Leheman bankruptcy. A possible explanation for this discrepancy is that the ADL parametric approach does not take into account the potential effects generated by fast volatility changes and cross-market volatility spillovers which may substantially affect their results.

Our findings are also relevant to provide an accurate mathematical description of stock index returns, which represents a fundamental task in finance. The jump clustering as well as the absence of cross-excitation effects in the jump activity have serious consequences for portfolio optimization, risk management and option pricing. For example, if jumps propagate across markets as in the ADL model, a risk manager has to take into account the transmission mechanism to mitigate the jump risk and rational investors would require a risk premium to face the risk of contagion. Jump clustering is also remarkably important since it implies a higher probability to observe multiple jumps within the same trading day compared to a standard Poisson processes, so affecting the shape of the return distribution.

The rest of this chapter is organized as follows. Section 1.2 describes the data and Section 1.3 reviews the most common approaches for jump identification and test. Section 1.4 presents the

multivariate Hawkes framework which is adopted to model the asset price dynamics and examines the presence of market co-jumps, while Section 1.6 concludes. Technical aspects are relegated to the Appendix.

1.2 Data description

Our data set comes from Olsendata and contains information on the S&P 500 and the Euro Stoxx 50 indexes in between 2007-09-13 and 2014-04-30; the two indexes are traded at the NYSE and at the Frankfurt Stock Exchange, respectively. The 7-year period covered by our analysis includes the sub-prime crisis leading to the bankruptcy of Lehman Brothers on September 15^{th} 2008 and the subsequent European sovereign crisis in 2011. For both markets, we compute the total return from prices reported every 5 minutes. This frequency is widely recognized to offer a reasonable balance between a fine sampling frequency on the one hand and robustness to market microstructure noise on the other (see for instance Andersen et al., 2010a). The NYSE and the Frankfurt Stock Exchange normally operate respectively form 9:30 to 16:00 and from 9:00 to 17:30 in local times, with the first price observed 5 minutes after market opening time. Each ordinary trading day has respectively 77 intraday returns for the S&P 500 and 100 returns for the Euro Stoxx 50. For the Euro Stoxx 50 we decide to ignore the first 10 minutes of activity due to a remarkably higher price variability compared to the rest of the day. This choice is consistent with most of the empirical literature where the first observations are usually excluded due to the potentially erratic price behavior produced by market opening procedures. We also exclude from the data set an extremely small number of days containing an anomalous number of price observations. At the end of the data cleaning process, our sample consists of 1674 trading days for the S&P 500 and 1691 for the Euro Stoxx 50. The cumulated log-return and the intraday annualized volatility measured from high frequency returns are reported on Table 1.1 where we can observe the highest volatility peaks during the Sub-prime and the Euro Sovereign crisis

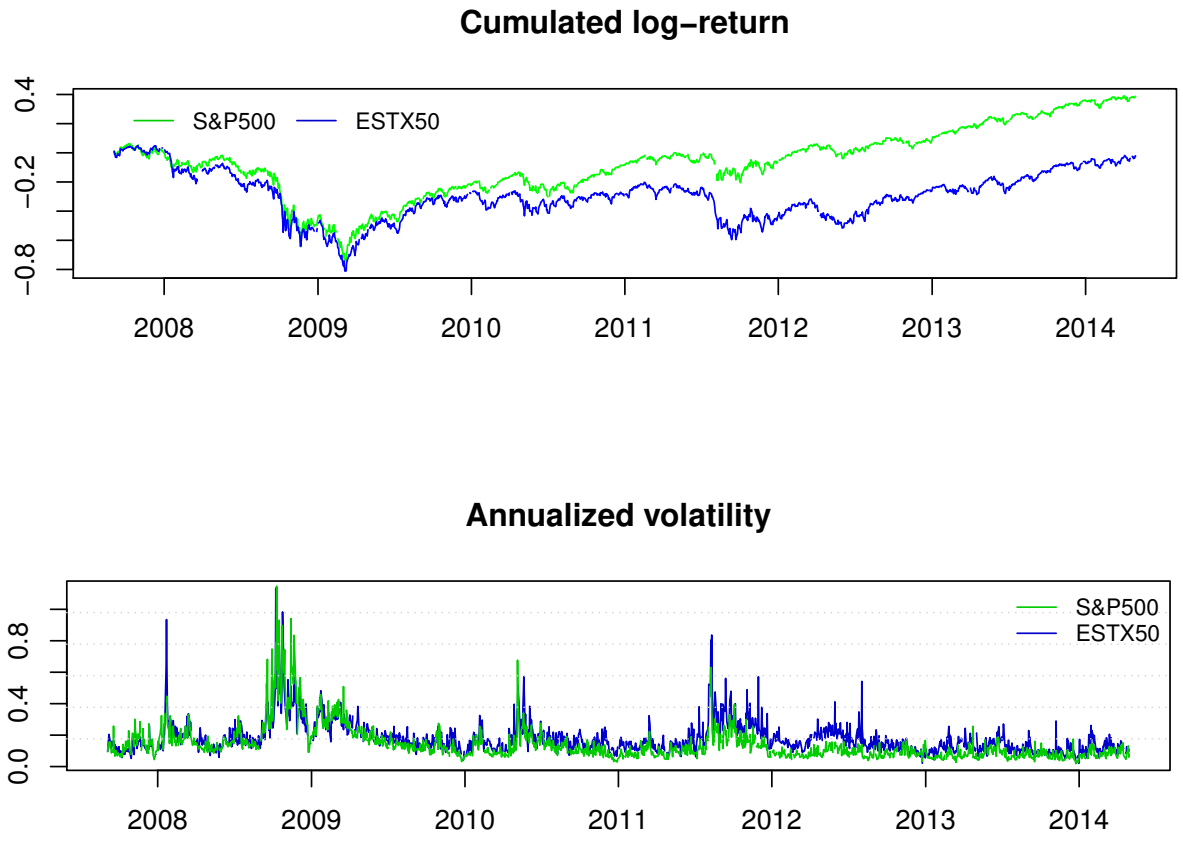


FIGURE1.1: The top panel exhibits the time series of cumulated log-returns (including the overnight). The bottom panel reports the annualized volatility (the square root of the realized quadratic variation multiplied by 252 which is the number of trading days in a year)

1.3 Jumps Identification

A detailed study of the statistical features of jumps requires the identification of their occurrence in a framework that is free from any specific parametric assumption on the price evolution. This task is crucial to our analysis and can be accomplished using non-parametric tests based on high frequency data. The field of jump testing developed significantly in the last decade, starting from the seminal work of BNS who developed a non-parametric method which relies on the comparison between two realized measures of volatility: the bipower variation and the quadratic variation. The former is driven exclusively by continuous price changes while the latter also includes jumps. Afterwards, several alternative tests have been proposed. In order to summarize the most relevant contributions,

we conveniently distinguish two main families. The first one is the BNS family which includes all the tests that are constructed by using the bipower variation or an alternative measure robust to jumps volatility. The second important family has been introduced by Lee and Mykland (2008) (LM henceforth) and is based on the idea that jumps can be identified when a return exceeds a certain threshold determined adaptively according to the instantaneous volatility. The BNS family of tests includes the contributions of Corsi, Pirino, and Renò (2010), Andersen, Dobrev, and Schaumburg (2012) and Podolskij and Ziggel (2010). Conversely, the proposals of Andersen, Bollerslev, and Diebold (2007b), Bollerslev et al. (2013) and Bormetti et al. (2015) belong to the LM family and they differentiate on the basis of the methodology employed to determine the volatility and the threshold level. Beyond these two families, it is worthwhile to mention also the "Swap Variance" tests of Jiang and Oomen (2008), looking at the difference between relative returns and log-returns, and the test of Aït-Sahalia and Jacod (2009) which is based on absolute return moments calculated at different sampling frequencies.

In a recent study, Schwert (2010) empirically shows that the applications of alternative identification methods can generally lead to substantially different conclusions on the presence of jumps. Therefore, the choice of a specific identification test may potentially drive our results on the statistical properties of the detected jumps. Unfortunately, none of the identification methods is generally preferable to the others, and the recent simulation studies of Dumitru and Urga (2011) and Gilder et al. (2014) among others show that the performances of the various tests in finite samples are related to the features of the data generating process as well as to the time frequency of prices observations. In view of this, we recognize that the choice of the jump identification technique deserves an accurate discussion and we follow an approach that is common in the literature: we compare the results obtained under different jump identification method. For brevity, we consider three alternative sets of jumps obtained by selecting one test from the BNS family and a second test from LM family; the third test is derived as the intersection of the previous two. According to Dumitru and Urga (2011), the intersection of two jump tests generally leads to a substantial reduction of the effective size compared to the nominal one and can be regarded to some extent as a reduction of the significance level. The technical motivations driving our choices will be largely discussed in the next subsection.

1.3.1 Jump Tests

Assume, as usual, that prices follow a continuous-time semi-martingale and let the log-price p_t be described by the stochastic differential equation

$$p_t = \int_0^t \mu_s \, ds + \int_0^t \sigma_s dW_s + \int_0^t J_s dN_s \tag{1.1}$$

where the drift μ_t has finite variation, the volatility σ_t is càdlàg, W_t is a standard Brownian motion, N_t is a finite activity counting process with possibly stochastic intensity λ_t and J_t is the random jump size. The stochastic processes encompassed by equation 1.1 exclude infinite activity jumps. However, the class of models covered is widely recognized to be flexible enough to capture the main features of financial time series at high frequency (see for instance Andersen et al. 2007a, Andersen et al. 2010a).

Assume also that each trading day t has duration 1 and $M + 1$ log-prices $p_{t,0}, \ldots, p_{t,M+1}$ are observed at equally spaced times. The intraday log-returns are indicated as $r_{t,i} = p_{t,i+1} - p_{t,i}$ for $i = 1, \ldots, M$ or alternatively with a single index r_i to denote the i-th log-return in the entire time series: $i = 1, \ldots, M \cdot T$ where T is the total number of trading days.

The BNS family of tests

While the LM type of tests requires substantial restriction to feature the volatility process, those belonging to the BNS family are more flexible and remain consistent even in presence of volatility jumps, although their power and their size in finite samples can be negatively affected by violent volatility shocks. The following volatility metrics is essential to the computation of the test statistics:

$$RV_t = \sum_{i=1}^{M} r_{t,i}^2 \tag{1.2}$$

the realized variance RV_t converges in probability to the quadratic variation as $M \to \infty$ (Andersen, Bollerslev, and Diebold, 2010b):

$$p \lim_{M \to \infty} RV_t = QV_t = \int_{t-1}^{t} \sigma_s^2 ds + \int_{t-1}^{t} {J_s}^2 dN_s \tag{1.3}$$

and in absence of jumps the quadratic variation corresponds to the integrated variance

$$QV_t = IV_t = \int_{t-1}^{t} \sigma_s^2 ds \tag{1.4}$$

To separate the contribution to the realized variance due to continuous price variation from the contribution of jumps, BNS introduce the bipower variation:

$$BPV_t \equiv \mu_1^{-2} \left(\frac{M}{M-1} \right) \sum_{i=2}^{M} |r_{t,i-1}| \, |r_{t,i}| \tag{1.5}$$

where $\mu_\gamma = \mathbb{E}\left(|u|^\gamma\right)$ and $u \sim N(0,1)$. In the asymptotic limit

$$p \lim_{M \to \infty} BPV_t = \int_{t-1}^{t} \sigma_s^2 ds$$

moreover, in absence of jumps, under other regularity conditions, the joint asymptotic distribution of RV_t and IV_t is normally distributed as

$$\sqrt{M} \begin{pmatrix} RV_{t,M} - IV \\ BV_{t,M} - IV \end{pmatrix} \xrightarrow{D} N \left(0, \begin{bmatrix} 2 & 2 \\ 2 & 2.62 \end{bmatrix} IQ_t \right) \tag{1.6}$$

where $IQ_t = \int_{t-1}^{t} \sigma_s^4 ds$ is the integrated quarticity. BNS propose some alternative statistics to compute the test, the most common is based on the relative jump measure:

$$RJ_t \equiv \frac{RV_t - \hat{IV}_t}{RV_t} \tag{1.7}$$

where $\hat{IV}_t$ denotes some jump robust measure of the integrated variance. We generally define the test statistics for BNS family as

$$Z_t \equiv \frac{RJ_t}{\sqrt{\frac{1}{M} \left(v_{I\hat{V}_t} - v_{RV_t} \right) \frac{I\hat{Q}_t}{I\hat{V}_t^2}}} \tag{1.8}$$

where $\hat{IQ}_t$ is a consistent estimate of the integrated quarticity, $v_{I\hat{V}}$ and v_{RV} are constant such that in absence of jumps

$$Var\left[\hat{IV}_t\right] = \frac{v_{I\hat{V}_t}}{M} IQ_t + O\left(M^{-2}\right) \quad \text{and} \quad Var\left[RV_t\right] = \frac{v_{RV_t}}{M} IQ_t$$

therefore $v_{RV} = 2$ while $v_{I\hat{V}}$ depends on $\hat{IV}$. Z_t is a pivotal quantity that converges asymptotically to a standard normal random variable. A jump is detected with the confidence level $1-\alpha$ when $Z_t > \Phi^{-1}_{1-\alpha}$ being $\Phi^{-1}_{1-\alpha}$ the inverse standard normal distribution evaluated at $1-\alpha$. In the original proposal of BNS, $\hat{IV}_t$ coincides with BPV_t and

$$\hat{IQ}_t = \max\left(\hat{IV}_t^2, QP_t\right)$$

being QP_t the quad-power quarticity

$$QP_t = \mu_1^{-4} M \left(\frac{M}{M-3}\right) \sum_{i=4}^{M} |r_{t,i-3}|\,|r_{t,i-2}|\,|r_{t,i-1}|\,|r_{t,i}| \tag{1.9}$$

The large diffusion of this test statistics is due to its suitable finite sample properties highlighted by Huang and Tauchen (2005). Equations 1.7 and 1.8 define the BNS family of tests throughout this and the following chapters. An interesting alternative volatility measure is the corrected threshold bipower variation introduced by Corsi, Pirino, and Renò (2010):

$$C-TBPV_t \equiv \mu_1^{-2} \sum_{i=2}^{M} Z_1\left(r_{t,i-1}, c_\theta^2 \hat{v}_{t,i-1}; c_\theta\right) Z_1\left(r_{t,i}, c_\theta^2 \hat{v}_{t,i}; c_\theta\right) \tag{1.10}$$

where

$$Z_\gamma\left(x, y; c_\theta\right) \equiv \begin{cases} |x|^\gamma & \text{if } x^2 \leqslant y \\ \frac{1}{2\Phi(-c_\theta)\sqrt{\pi}} \left(\frac{2}{c_\theta^2}\, y\right)^{\gamma/2} \Gamma\left(\frac{\gamma+1}{2}, \frac{c_\theta^2}{2}\right) & \text{if } x^2 > y \end{cases} \tag{1.11}$$

Φ is the cumulative standard normal distribution and $\Gamma(\alpha, x)$ is the upper incomplete gamma function. The $C-TBPV_t$ replaces the absolute returns exceeding the threshold by their conditional expected value under the normality assumption:

$$\mathbb{E}\left[\,|r_{t,i}|^\gamma \,|\, r_{t,i}^2 > c_\theta^2\right] = Z_\gamma\left(r_{t,i}, c_\theta^2 \hat{v}_{t,i}; c_\theta\right)$$

The corrected threshold tripower quarticity is analogously defined as:

$$C-TTPV_t \equiv \mu_1^{-2} \sum_{i=3}^{M} Z_{4/3}\left(r_{t,i-2}, c_\theta^2 \hat{v}_{t,i-2}; c_\theta\right) Z_{4/3}\left(r_{t,i-1}, c_\theta^2 \hat{v}_{t,i-1}; c_\theta\right) Z_{4/3}\left(r_{t,i}, c_\theta^2 \hat{v}_{t,i}; c_\theta\right) \quad (1.12)$$

Asymptotically, the $C-TBPV_t$ and the $C-TTPV_t$ behave analogously to the bipower variation and the tripower quarticity in absence of jumps. In presence of jumps these measures offer the advantage of reducing the upward bias which characterizes the multipower variation measures of Barndorff-Nielsen et al. (2006), with positive effects on the power of the test. The simulation study of Corsi, Pirino, and Renò (2010) shows that the gain is particularly relevant in presence of consecutive jumps: in this case the bias affecting the multipower variation can become extremely large with detrimental effects on jump detection. Anyway, the main disadvantage of the threshold multipower measures is the necessity to estimate the local volatility and to set a judgmental threshold level.

Andersen et al. (2012) introduce two jump robust volatility measures based on nearest neighbor truncation that can be regarded as special cases of the quantile-based realized volatility measures of Christensen et al. (2010). The bias generated in finite samples by the presence of jumps and stochastic volatility is generally small for these estimators but they have a lower asymptotic efficiency compared to the bipower and the threshold bipower variation. However, the bias may become very large in presence of consecutive jumps: this is of particular concern in our study given that our analysis reveals quite often the presence of consecutive jumps due to jump clustering effects[1].

The test proposed by Podolskij and Ziggel (2010) belongs also to the BNS family but it is constructed on the threshold estimator of Mancini (2009) that removes returns larger then a certain size. It has the advantage of being efficient efficiency but it is strongly dependent on the threshold level. Such a dependence implies a serious risk of retaining jumps when the threshold is too high or removing continuous returns when the threshold is too low. In addition, the test requires an external perturbation to allow for the development of the required asymptotic theory and the the form of this perturbation entails additional arbitrarity.

Given the considerations above, the testing methodology of Corsi et al. (2010) appears as the most appropriate within the BNS family for the purposes of our empirical analysis. As a proxy for the

[1] Jump clustering is documented for instance by ADL, Chen and Poon (2013), Bormetti et al. (2015) among others.

instantaneous volatility we adopt the estimator defined by equation 1.19 in the next paragraph. A potential problem for this kind of test concerns the presence of stale quotes which may generate a downward bias in the multipower variation, hence increasing the probability of spurious jumps detection. To avoid this issue we also remove the stale quotes from our sample before computing the volatility measures.

Identification of intraday jump times All the tests belonging to the BNS family are designed to reveal the presence of at least one jump over a certain time period, typically a single trading day. To analyse the statistical properties of jumps we need to classify each single intraday return as a jump or alternatively as a continuous price fluctuation. In this chapter we follow the iterative procedure described in Andersen et al. (2010a) which replicates the BNS test but removes at each step the contribution of the largest absolute return from realized variance. However, we adopt this method with some important modifications:

1. The BNS test is replaced by the methodology of Corsi et al. (2010) since it produces a smaller bias in presence of jumps.

2. The test is calculated after rescaling high frequency returns to remove the intraday periodicity of volatility as recommended by Rognlie (2010). This procedure reduces the intraday variability of volatility which induces the downward bias in the bipower variation and increases the spurious detection rate.

3. To reduce the impact of stale, we remove zero intraday returns from the sample before computing the realized measures. The presence of an isolated null return causes two consecutive blocks in the bipower to be zero, while the impact is smaller on the realized variance. The negative bias in the bipower is therefore removed with the preliminary exclusion of stale quotes from the calculation.

4. Following Gilder et al. (2014) we classify a jump as the largest absolute log return after adjusting for the intraday volatility pattern.

5. Differently from Andersen et al. (2010a), we recalculate at each step the threshold bipower variation and the tripower quarticity to guarantee the complete removal of the upward bias

in case of jumps.

6. The iterative procedure is based on multiple hypotheses testing and the size of the test can be seriously affected. To deal with this issue we apply the conservative Holm-Bonferroni correction. Note that without such a correction the maximum number of jumps detected on a single trading day increases remarkably[2]: for the S&P 500 index it passes from 5 to 8, for the Euro Stoxx 50 from 7 to 11.

The LM Family of Tests

The LM test is based on a measurement of the instantaneous volatility of the continuous component in Equation 1.1. Such a measurement is feasible with asymptotically infinite precision only if the drift μ_t and the volatility σ_t changes "slowly" in time (see Lee and Mykland, 2008 for details). This restriction is the major limitation for these tests since consistency is not guaranteed in the presence of volatility jumps. Note that the presence of volatility bursts and simultaneous occurrence of price and volatility jumps are increasingly documented in the literature via parametric and non-parametric methods. Recent contributions are Todorov and Tauchen (2011), Jacod and Todorov (2010), Corsi and Renò (2012), Wei (2012), Christensen et al. (2014), Bandi and Renò (2016). The tests belonging to the LM family are based on the following statistic:

$$z_{t,i} = \frac{|r_{t,i}|}{\sqrt{\hat{V}_{t,i}}} \tag{1.13}$$

where $\hat{V}_{t,i}$ is an estimate of the instantaneous volatility and $z_{t,i}$ is the normalized absolute return. As the sampling frequency increases, $\hat{V}_{t,i}$ converges to the unobserved instantaneous volatility and $z_{t,i}$ distributes as the absolute value of a standard normal random variable. A jump is detected whenever $z_{t,i}$ exceeds a predetermined threshold θ. The various LM tests proposed in the literature differ for the methodology used to determine the threshold level and for the estimation of the instantaneous volatility. Concerning the threshold, as the test is applied for every intraday return, the issue of false discovery rate arising in the context of multiple hypotheses testing must be properly taken

[2] It is worthwhile to mention that Andersen et al. find just a few occurrences of multiple jumps in a single day with at most 3 jumps even though they do not apply any correction to the confidence level. According to our evidence, this is motivated by the fact that they do not recalculate the bipower variation at each step, thus the positive bias probably "compensates" for the decreasing confidence level.

into account. The simplest solution (proposed by Andersen et al., 2007b) consists in the application of the Šidák approach: given a certain daily size α, the corresponding size for each intraday test is $\beta = 1 - (1-\alpha)^{1/M}$ and the associated threshold level is $\theta = \Phi^{-1}_{1-\beta/2}$. However, finite sample volatility is always measured with an error and the Šidák approach often leads to over-rejection of the null. LM propose to calculate critical values from the limiting distribution of the maximum of the test statistics: as $M \to \infty$ the quantity

$$\xi_M = \frac{\max_i (z_{t,i}) - C_M}{S_M}$$

with $C_M = (2\log M)^{1/2} - [\log\pi + \log(\log M)]/2\sqrt{2\log M}$ and $S_M = (2\log M)^{-1/2}$, distributes as a Gumbel random variable. This method is more conservative and reduces the probability of detecting spurious jumps.

With regard to the estimators used for the instantaneous volatility, LM propose the bipower variation calculated over a time window of size K depending on the sampling frequency (they recommend $K = \sqrt{252\,M}$ where 252 is the number of trading days in a year)

$$\hat{V}_i^{LM} = \frac{1}{\mu_1^2 (K-2)} \sum_{j=i-K+1}^{i-1} |r_j|\,|r_{j-1}| \tag{1.14}$$

Andersen et al. (2007b) measure the local volatility using the bipower variation calculated over the entire trading day:

$$\hat{V}_{t,i}^{ABD} = \frac{BPV_t}{M} = \frac{1}{\mu_1^2 (M-1)} \sum_{i=2}^{M} |r_{t,i-1}|\,|r_{t,i}| \tag{1.15}$$

The estimators 1.14 and 1.15 are very similar, the main difference is that ABD use the returns belonging to a single trading day while the LM estimator is constructed on past returns collected from different trading days. It is important to remark that both estimators are upward biased in case of jumps and may substantially lose accuracy when the instantaneous volatility moves rapidly. Both issues are extremely relevant for our purposes: the former reduces the detection power of the test, especially when multiple jumps occur closely in time, and it also influences the clustering pattern; the latter increases the the error affecting our volatility estimates and therefore the probability of

spurious jump detection.

To remove the bias, Bormetti et al. (2015) construct an estimator similar to LM which is based on the threshold bipower variation: the past information is weighted through an exponential moving average. It is well known that volatility is generally higher at the beginning and at the end of the trading day, following a U-shaped intraday pattern that is largely documented in the literature (see for instance Bollerslev et al. 2013 and Gilder et al. 2014). This pattern must be taken into account to improve our volatility estimates. In details, let $\tilde{r}_{t,i}$ denote the log-return scaled by a proper factor to remove the intraday periodicity:

$$\tilde{r}_{t,i} = r_{i,t}/\zeta_i \qquad \text{where} \qquad \zeta_i = \frac{1}{T}\sum_{t=1}^{T}\frac{|r_{t,i}|}{sd_t} \tag{1.16}$$

being sd_t the standard deviation of the absolute returns on day t. The estimator is defined by the following equations

$$\tilde{V}_i^{BEW} = \frac{\alpha}{\mu_1^2}\left|\tilde{r}_{j'}\right|\left|\tilde{r}_j\right| + (1-\alpha)\,\tilde{V}_{i-1}^{BEW} \quad i = 1,\dots,M\cdot T \tag{1.17}$$

$$\hat{V}_{t,k}^{BEW} = \zeta_k\,\tilde{V}_{t,k}^{BEW} \qquad t = 1,\dots,T \quad k = 1,\dots,M \tag{1.18}$$

where $\tilde{V}$ indicates the volatility estimates purified by the intraday pattern that is instead included in $\hat{V}$, $j < j' \leqslant i-1$, $|\tilde{r}_j|/\sqrt{\hat{V}_j^{BEW}} \leqslant \theta$ and $|\tilde{r}_l|/\sqrt{\hat{V}_l^{BEW}} > \theta \quad \forall\, j < l < j'$. This estimator represents a moving average weighted bipower variation which excludes all the observation exceeding the threshold θ. However, the inaccuracy in presence of fast volatility changes still remains a critical issue. We introduce three modifications: i) we correct the intraday volatility patterns following the method proposed by Boudt et al. (2011) which ensures efficiency and consistency in presence of jumps (technical details are reported in Appendix A.1); ii) the moving average parameter α is determined in order to reduce the autocorrelation of the standardized squared returns[3] ; iii) to improve the accuracy in presence of sharp volatility changes, we also consider the estimator $\hat{V}^{FEW}$ based on forward information and defined exactly as $\hat{V}^{BEW}$ but on the time reversed series (i.e.

[3] Our estimated optimal half life time is around 150 minutes.

the series obtained substituting the index i with $M \cdot T - i + 1$). Our new estimator is

$$\hat{V}_i^{SEW} = \frac{1}{2}\left(\hat{V}_i^{BEW} + \hat{V}_i^{FEW}\right) \tag{1.19}$$

which is symmetric in time, i.e. it equally weighs past and future information. An extensive simulation study that will be presented in Section 2.6 shows that this test has remarkably more power compared to the original specification proposed by LM because our local volatility estimator is not only able to reduce the bias due to price jumps but is also more responsive to quick changes in volatility. In presence of an upward volatility jump, for instance, the backward estimator tends to underestimate volatility and it is likely to signal spurious price jumps; the same can happen to the forward estimator in presence of negative volatility jumps. This finds also some confirmation in our empirical analysis. We apply the test at the 99% confidence level calculated according to the conservative approach of LM (the threshold level is 4.07 and 4.11 standard deviations for U.S. and EU respectively). For the S&P 500, the symmetric $\hat{V}^{SEW}$ estimator reveals 236 days containing at least one jump and a total number of 315 intraday jumps. Measuring the volatility with the backward $\hat{V}^{BEW}$ estimator leads to the detection of 506 jumps while the forward estimator $\hat{V}^{BEW}$ recognizes 319 jumps. The results obtained for the Euro Stoxx 50 index are also similar: 631, 957 and 700 jumps are detected under $\hat{V}^{BEW}$, $\hat{V}^{FEW}$,$\hat{V}^{SEW}$ respectively. It is quite evident that measuring volatility accurately is crucial for a correct identification of identification jumps. This point has been also recently analyzed by Christensen et al. (2014) and Bajgrowicz et al. (2016) who find that many price fluctuations classified as jumps at 5 minutes by standard tests are spurious. An increasing empirical evidence is pointing toward volatility jumps (Todorov and Tauchen 2011, Jacod and Todorov 2010, Corsi and Renò 2012, Wei 2012, Christensen et al. 2014, Bandi and Renò 2016) which can jeopardize the outcomes of our jump identification methods. To have some indication about the possible occurrence of volatility jumps in correspondence of price jumps, we construct the relative difference between the backward and the forward volatility estimator as follows

$$\delta_i = 2\,\frac{\hat{V}_i^{FEW} - \hat{V}_i^{BEW}}{\hat{V}_i^{FEW} + \hat{V}_i^{BEW}} \tag{1.20}$$

the mean of δ_i calculated unconditionally as well as in case of price jumps is reported in Table

1.1. Results are very similar for both markets: when the price jumps, then the forward-estimated volatility is substantially larger indicating the presence of a sharp upward volatility movement. Table 1.1 also shows that the volatility increase is larger when price jumps are negative, however both jump directions seem to be accompanied by positive volatility shocks. Though this analysis provides only some indications on the possibility to have jumps in volatility, it can represent a starting point to design formal tests whose development is left for future research.

	S&P 500	ESTX 50
unconditional	1.4%	1.5%
conditional on a jump	24.1%	19.5%
conditional on a negative jump	31.7%	26.4%
conditional on a positive jump	15.1%	11.9%

Table 1.1: Conditional and unconditional mean of δ_i for S&P 500 and Euro Stoxx 50. (equation 1.20)

1.3.2 Results

Table 1.2 reports the summary statistics for the three alternative sets of jumps: the first set obtained from our modified version of the LM test (m-LM henceforth), the second set from sequential version of the CPR test (s-CPR), while the third set comes as the intersection of the previous two. All jumps are reported on figures 1.2 and 1.3 for the Euro Stoxx and for the S&P respectively. First of all, we note that there are significant differences between the outcomes of m-LM and the s-CPR methods: the m-LM test always detects more jumps (almost twice of those detected under the s-CPR test). Table 1.3 reports result from a simulation study conducted using the same setting proposed in Chapter 2 where jump clustering effects are introduced through a stochastic jump intensity which evolves according to a univariate self-exciting Hawkes process (see Section 1.4 for details). We notice that the LM-type of tests have generally more power than the BNS test confirming the results of Dumitru and Urga (2011) and Gilder et al. (2014). According to our evidence the m-LM test proposed in this paper has a remarkably smaller size compared to the standard version of LM. We observe that all the jump detection methods benefit significantly from the intraday volatility pattern correction which largely reduces the size in finite samples while the market microstructure noise has indeed only minor effects at 5 minutes. Importantly the outcomes of the m-LM and the s-CPR tests are not independent: the size of the intersection is larger than the the product of their

individual size. The intersection ensures anyway a large decrease of the significance level making the detection mechanism more severe. Bajgrowicz et al. (2016) claim that the role of jumps in the literature is probably overstated due to a high rate of spurious detections. To avoid drawing wrong conclusions about the dynamics of jumps they suggest to control the Family Wise Error Rate (FWER) or the False Detection Rate (FDR). The first approach aims at controlling the probability of having one or more type I errors in the whole sample and leads to extremely severe significance levels which may result in a serious loss of power. The letter instead is designed to control the ratio between false detection and the total number of detected events and allows to preserve more power. The expected false discovery rates on Table 1.2 are calculated using to the size estimated on our numerical simulation.

The relative contribution of jumps to the total price variance is calculated as the average over all trading days of the ratio between the sum of squared detected jumps and the realized quadratic variation. Taking the sample mean of RJ_t ad defined in equation 1.7 with $\hat{IV}_t = BPV_t$, Huang and Tauchen (2005) find that about 7.3% of the quadratic variation on the S&P 500 is due to jumps. A similar calculation performed on our data with $\hat{IV} = C - TBPV_t$ gives an average ratio of 8.0% for the S&P 500 and 9.3% for the Euro Stoxx 50. Note that these estimates differ significantly from the results of Table 1.2. Interestingly, the mean of RJ_t on days where no jumps are detected according to the s-CPR test at the 99% confidence level is respectively 4.5% and 4.6% for the two indexes. This result can be explained in two different ways that are not mutually exclusive: 1) our choice of the confidence level is too severe to effectively remove the majority of jumps; 2) even after the corrections we adopted the threshold bipower variation is still seriously downward biased. Christensen et al. (2014) find that the contribution of jumps to total price variance extracted from 5 minutes data is usually overestimated and intraday volatility bursts[4] are often misclassified as jumps. Using data sampled at a higher frequency and applying specific corrections for the microstructure noise, they find that the contribution of jumps is much smaller (around 1% for the equity market). Such results are also confirmed by Bajgrowicz et al. (2016). We therefore retain our confidence level at 99% to avoid an excessive spurious detection rate.

[4] The recent paper of Christensen et al. (2016) shows that flash crashes are indeed characterized by drift bursts with a continuous path. Such events cannot be clearly distinguished from jumps at 5 minutes but require a higher frequency to be properly investigated.

Note that under all identification procedures, the number of jumps as well as their size and their relative contribution to total price variance are smaller for the U.S. index compared to the Euro Stoxx 50, which can be plausibly related to the lower diversification of the European index.

Figures 1.5 and 1.4 show the intraday distribution of jump times. Even if we do not dispose of a detailed data set reporting all the relevant news announcements, the pattern of intraday jumps clearly suggests that at least a significant part of detected jumps can be related to macroeconomic releases and other scheduled announcements. For the U.S. market we notice a peak at about 30 minutes after the market opening (less pronounced under the s-CPR method) which corresponds to the macroeconomic announcements scheduled around 10 o'clock (see Gilder et al. 2014). This effect is also visible in the intraday volatility pattern and we defer to Appendix A.1 for further details. A second and more evident peak on the U.S. market is located around 14:00 which is the time the Federal Fund Target rate is publicly communicated after the FOMC meeting. For the European market we observe a large number of jumps located within the 5 minute interval starting at 14:30 in local time corresponding to the start of the pre-negotiation time at the NYSE. A second and smaller peak is visible 1 hour and a half later, in correspondence of the U.S. macroeconomic announcements previously mentioned. This evidence suggests some dependence of the jump activity on the European market from the news concerning the U.S. economy. This type of information may generate simultaneous reactions in both markets which are usually referred as co-jumps. A detailed analysis of this topic is deferred to Section 1.4.

Figures 1.6 and 1.7 report the intraday annualized volatility measured by the square root of the quadratic variation (thus also including also the contribution of jumps). The average number of jumps and the average contribution of jumps to the quadratic variation are also reported. All figures show that jumps occur more frequently during low volatility periods. Remarkably, also the relative contribution of jumps is larger when the volatility is lower, regardless of the identification method. Note that the inverse relation between the volatility level and the relative contribution of jumps is even more pronounced when the continuous volatility is measured by the threshold bipower variation or using the $MinRV_t$ and the $MedRV_t$ measures proposed by Andersen et al. (2012). While the detection of a decreasing number of jumps when volatility rises can be determined by a deterioration of the power of the tests, the evidence concerning a diminished relative contribution

of the jump component to the total price variance calculated on non-parametric measures of the integrated volatility is much more striking. In presence of a quickly changing volatility for instance, the bipower variation is downward biased and the relative contribution of jumps is overestimated. Importantly, we can conclude according to the empirical evidence that both the sub-prime and the Euro Sovereign crisis have been characterized by a large upward volatility shocks while jumps played a only minor role.

Empirical Results

	S&P 500			ESTX 50		
	m-LM	s-CPR	m-LM ∩ s-CPR	m-LM	s-CPR	m-LM ∩ s-CPR
days with jumps	236	167	106	465	286	222
total jumps	315	196	119	637	351	255
max. jumps per day	6	5	3	5	7	3
contrib. price var.	3.7%	2.3%	1.8%	6.5%	4.0%	3.5%
average jump size	0.44%	0.40%	0.48%	5.0%	5.2%	5.9%
max. jump size	3.6%	2.6%	2.6%	3.4%	3.4%	3.4%
min. jump size	0.09%	0.08%	0.1%	0.1%	0.04%	0.1%
FDR	7.6%	14.5%	3.2%	3.9%	9.0%	1.5%

Table 1.2: Summary statistics for jumps detected under different methods. The m-LM and the s-CPR tests are applied with a nominal confidence level equal to 99%. The contribution of jumps to the total price variance is calculated here as the sample average of the ratio between the sum of squared detected jumps and the realized quadratic variation on each trading day(overnight returns are excluded from the denominator). The FDR is obtained multiplying the size of the test reported in the first panel of Table 1.3 by the number of trading days and the dividing by the number of detected jump days.

Simulation Results

Correction for intraday volatility pattern applied

	No microstructure noise				Strong microstructure noise			
	LM	m-LM	CPR	m-LM ∩ CPR	LM	m-LM	CPR	m-LM ∩ CPR
Power	71.2%	71.2%	63.1%	62.4%	68.7%	69.3%	60.6%	60.0%
Size	3.91%	1.07%	1.52%	0.20%	3.61%	1.01%	1.50%	0.21%

Correction for intraday volatility pattern not applied

	No microstructure noise				Strong microstructure noise			
	LM	m-LM	CPR	m-LM ∩ CPR	LM	m-LM	CPR	m-LM ∩ CPR
Power	72.0%	71.7%	63.0%	62.40%	70.0%	70.0%	61.0%	60.0%
Size	8.52%	3.33%	2.48%	0.280%	6.88%	1.01%	2.40%	0.62%

Table 1.3: Simulations results for the different tests (details of the simulation setting are provided in Section 2.6 of Chapter 2). Each simulated process is sampled every 5 minutes.

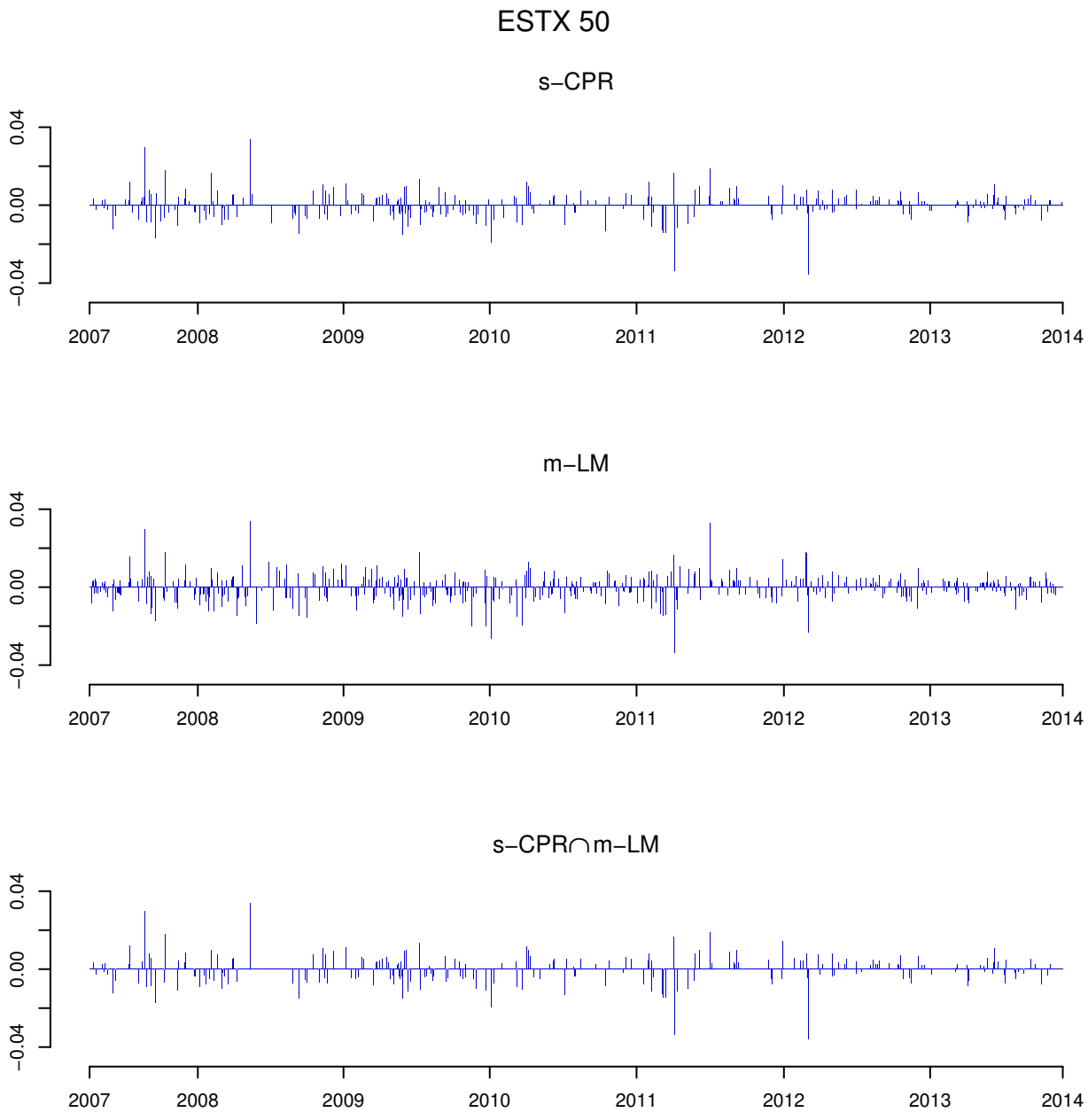


FIGURE1.2: Jumps in the Euro Stoxx 50 identified under different detection methods

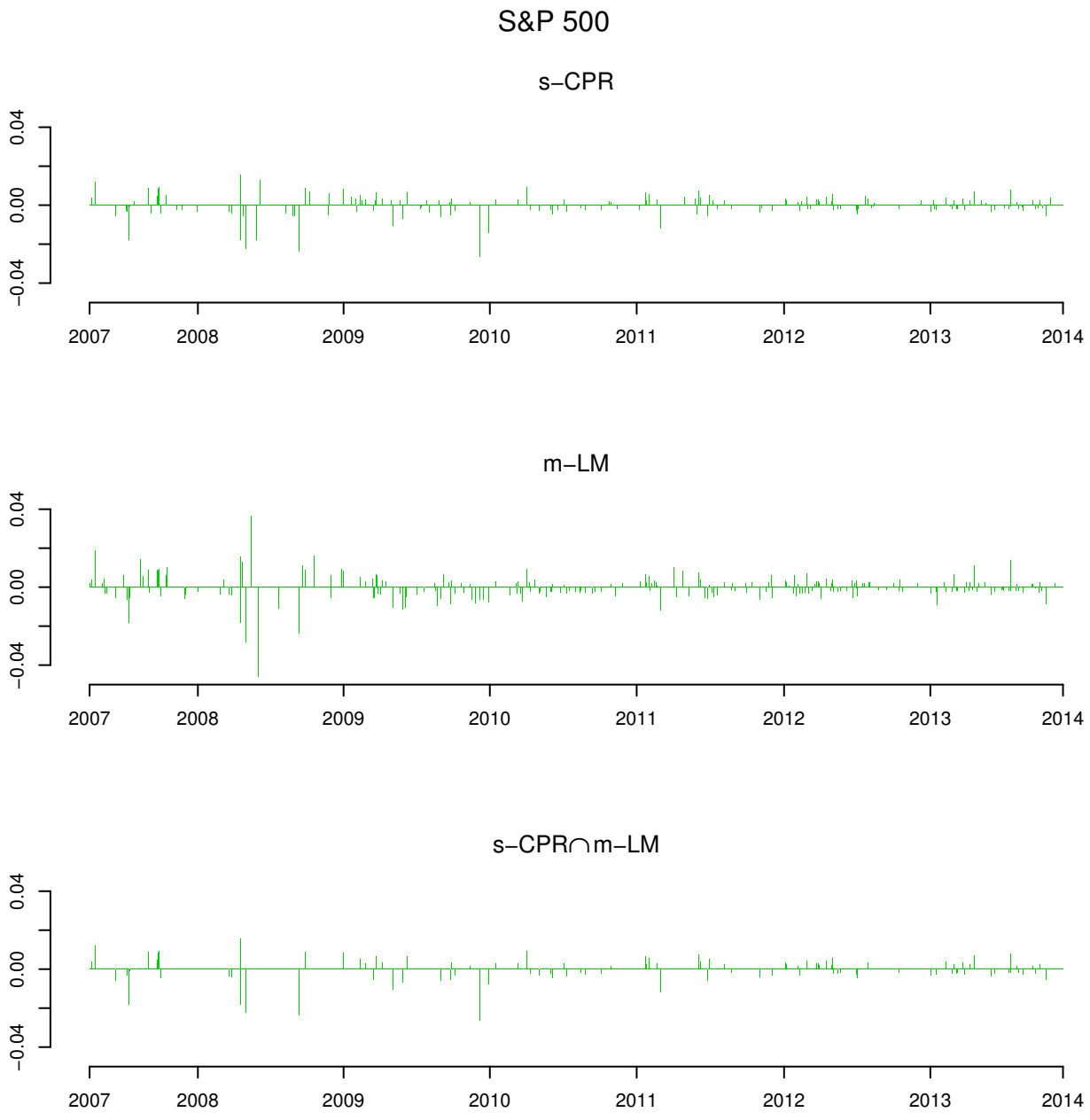


FIGURE1.3: Jumps in the S&P 500 identified under different detection methods

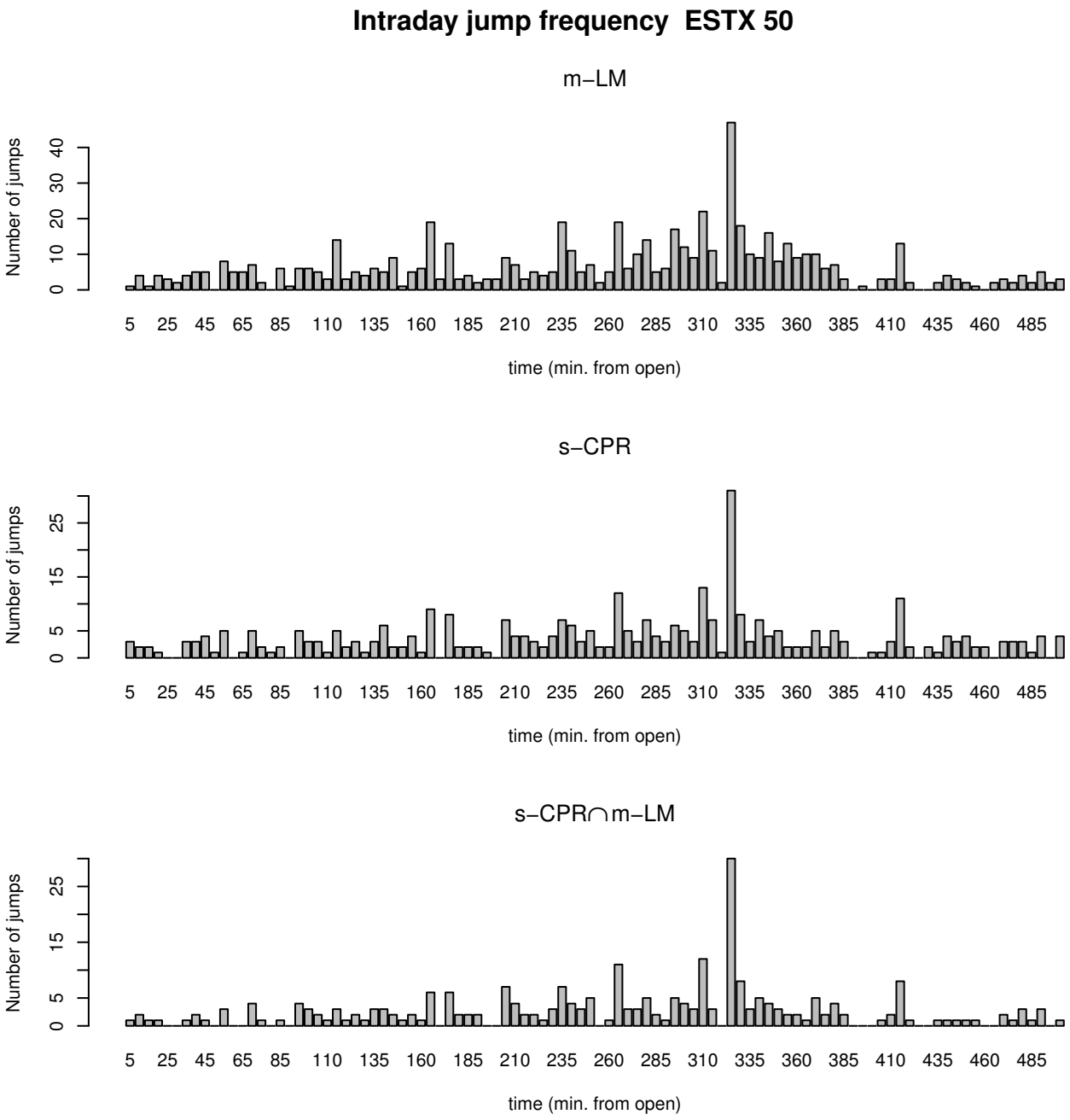


FIGURE1.4: Distribution of intraday jump times for the Euro Stoxx 50.

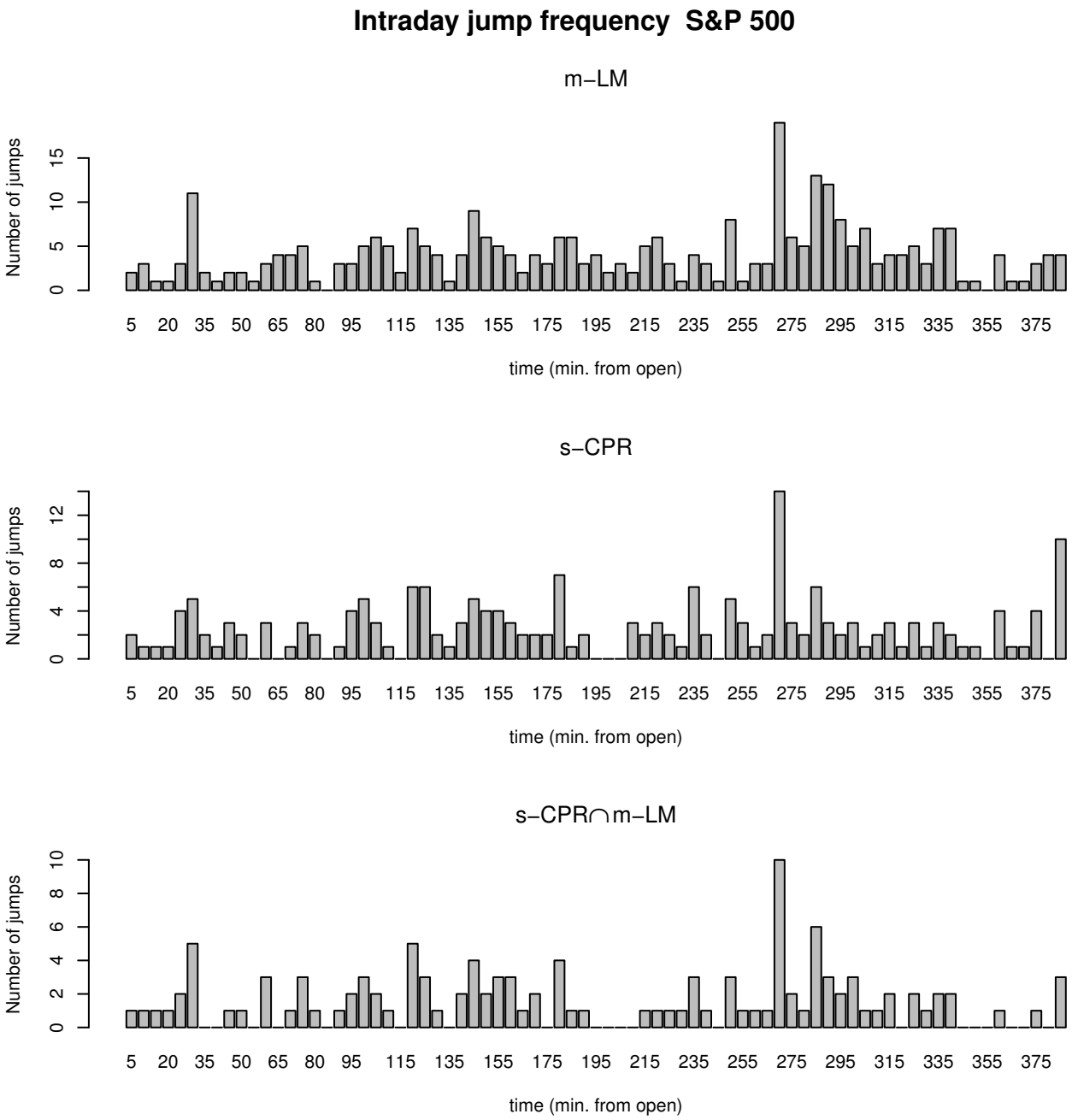


FIGURE1.5: Distribution of intraday jump times for the S&P 500.

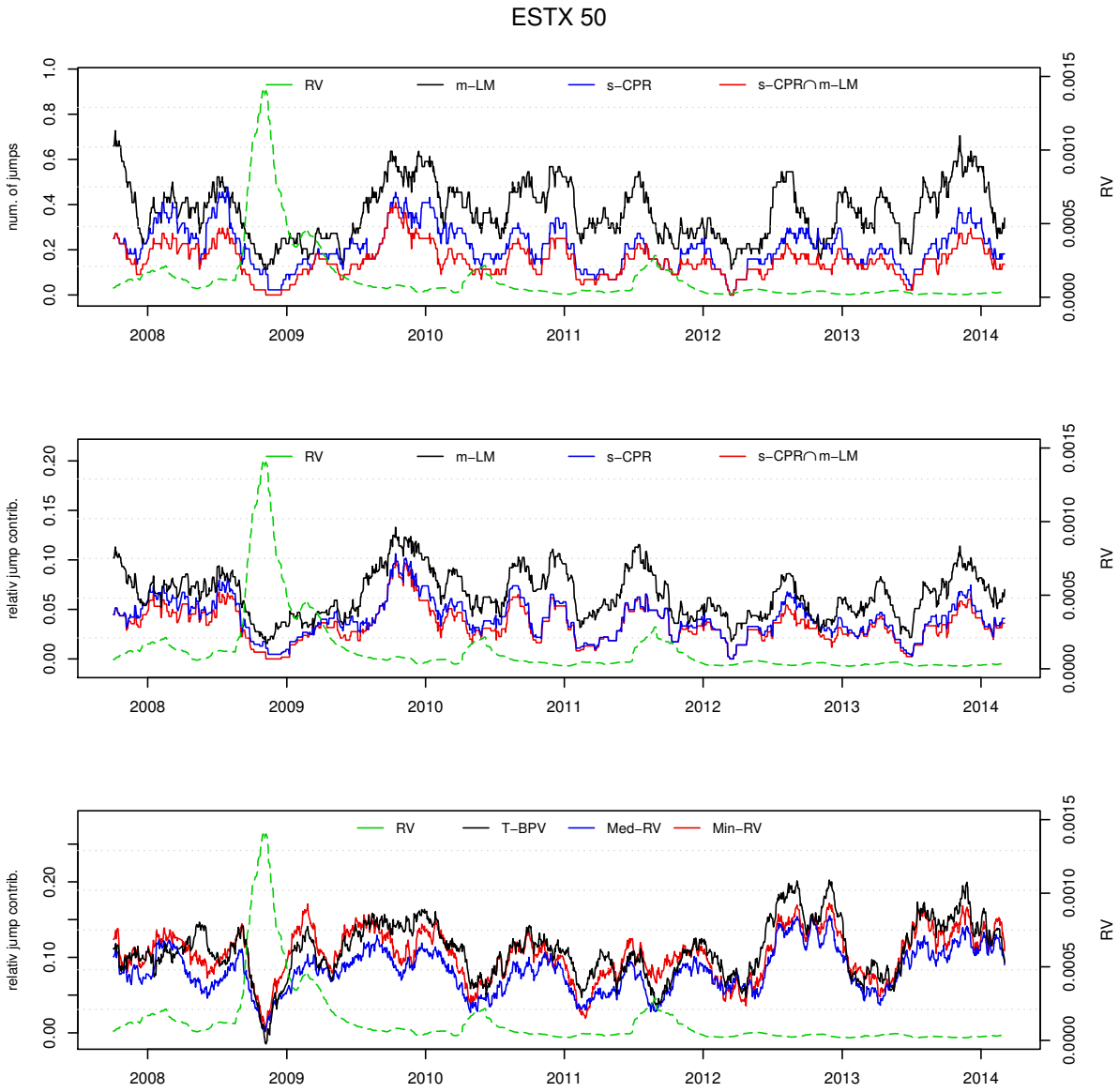


FIGURE1.6: The top panel reports the average number of jumps, the central and the bottom panels show the average contribution of jumps to total price variance calculated respectively from detected jumps and from jump robust realized measures. The dashed line is the intraday realized volatility. All these quantities are averaged over a rolling window of two months centered on the reference date.

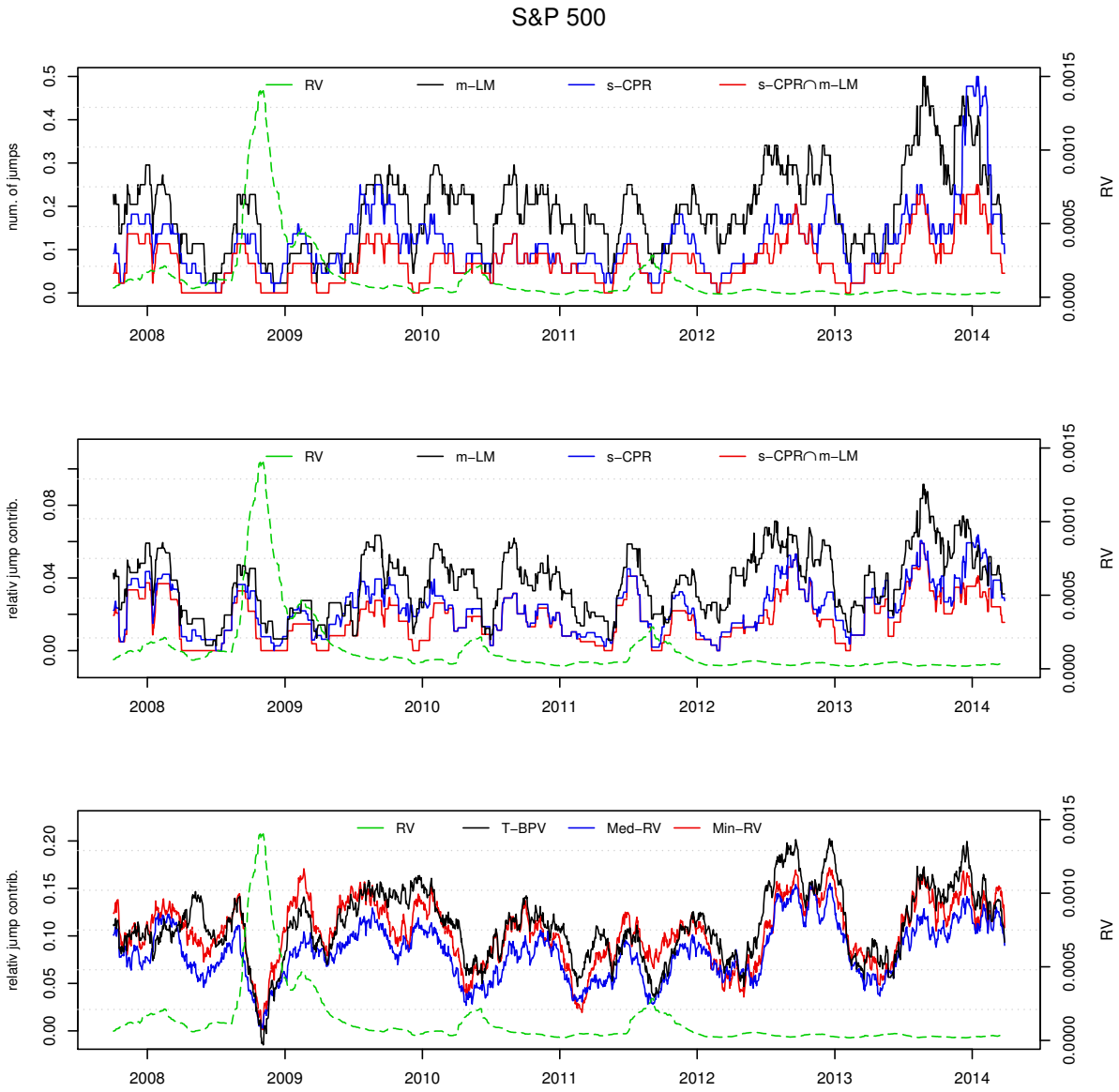


FIGURE1.7: The top panel reports the average number of jumps, the central and the bottom panels show the average contribution of jumps to total price variance calculated respectively from detected jumps and from jump robust realized measures. The dashed line is the intraday realized volatility. All these quantities are averaged over a rolling window of two months centered on the reference date.

1.4 Modelling Jumps with Multivariate Hawkes Processes

Hawkes processes belong to the class of multivariate point processes, they have been originally introduced by Hawkes (1971b) and Hawkes (1971a) and widely employed to model earthquakes. The use of Hawkes processes in finance has been proposed by Bowsher (2007) to model market transactions, ADL to model contagion jump cascades involving multiple markets, Bormetti et al. (2015) to analyze the multivariate jumps dynamics in the Italian stock market and Granelli and Veraart (2016) to study the variance risk premium on an index whose constituents are subject to contagion. In this chapter, we use the Hawkes processes to model the jump intensities $\lambda_{l,\tau}$ where l assumes the value 1 for the Euro Stoxx 50 index and 2 for the S&P 500. We will use the notation $\tau(t,i)$ to denote the time corresponding to the i^{th} interval on day t. The standard specification for the jumps intensities is

$$\lambda_{l,\tau} = \theta_l + \sum_{l'=1}^{2} \int_{-\infty}^{\tau} g_{l,l'}(\tau - s)\, dN_{l',s} \qquad l = 1,2 \tag{1.21}$$

where $N_{l,\tau}$ is the counting process for market l and the function $g_{l,l'}$, which is usually exponential, measures the effect that an event on market l' generates on the intensity of market l. This model is able to produce jump clustering, because past jumps increase the current intensity whenever $g_{l,l} > 0$, as well as cross-excitation effects from l' to l when $g_{l,l'} > 0$ for $l \neq l'$. ADL use this type of process to describe jump cascades involving financial markets located in different geographical regions. Their GMM estimates performed on daily data show statistically significant spillovers in the jump activity related to the propagation of large shocks during periods of crisis. Our purpose is to use this simple model and some of its variants to statistically describe the dynamics of jumps detected from high frequency data. This application requires some modification of the original model in order to take into account that markets operate simultaneously only for a small portion of the trading day. Moreover trading days can differ across countries due to specific national holidays. When a market is closed its jump intensity is obviously zero, nevertheless we want to allow the information coming from the operating markets to possibly influence the jump intensity on the next trading day. Equation 1.21 can describe market l during its operating time. When the market l' is closed, only self-excitation can take place for l since no jumps occur on l'. The missing part of

the dynamics is the overnight evolution of λ_l which obviously requires some specific assumptions. Let $o_{l,t}$ and $c_{l,t}$ denote the opening and closing time of market l measured according to some time convention (for instance UTC); for non-trading days we simply assume $o_{l,t} = c_{l,t}$. We assume the following recursive evolution to take into account the events occurring in market l' when market l is closed :

$$\lambda_{l,\tau} = \begin{cases} \lambda_{l,o_{l,t}} + \theta_l + \sum_{l'=1}^{2} K_{l,l'} \int_{o_{l,t}}^{\tau} e^{-\gamma_{l,l'}(\tau-s)} dN_{l',s} & \tau \in [o_{l,t}, c_{l,t}) \\ 0 & \tau \in [c_{l,t}, o_{l,t+1}) \end{cases} \tag{1.22}$$

$$\lambda_{l,o_{t+1}} = \lambda_{l,c_{l,t}} + \theta_l + \sum_{l' \neq 1}^{2} \tilde{K}_{l,l'} \int_{c_{l,t}}^{t} e^{-\tilde{\gamma}_{l,l'}(\tau-s)} dN_{l',s} \tag{1.23}$$

Where $K_{l,l'}$, $\tilde{K}_{l,l'}$ generate the self and mutual excitation of a possibly different magnitude in distinct time periods: $K_{l,l'}$ refers to periods of simultaneous market activity while $\tilde{K}_{l,l'}$ refers to periods when only one market is open. According to Equation 1.23, the overnight evolution of the jump intensity is similar to the intraday evolution but has a possibly different mean reversion rate as well as a different impact of past jumps. These features are designed to capture the slowdown of the information flow which characterizes the overnight period. Table 1.4 reports the maximum likelihood estimates of our model when single elements of the dynamics are progressively included. Model 1 is a simple Poisson process with constant intensity obtained imposing $K_{l,l'} = \tilde{K}_{l,l'} = 0$, $\lambda_{l,0} = \theta_l$ for $l, l' = 1, 2$. Model 2 is a univariate Hawkes process that allows for self-excitation: the restrictions are $K_{l,l'} = \tilde{K}_{l,l'} = 0$. The likelihood function exhibits a low dependence on $\tilde{\gamma}$ which is never found to be statistically significant, thus we estimate the model under two alternative assumptions: no overnight mean reversion ($\tilde{\gamma} = 0$) and fast overnight mean reversion ($\tilde{\gamma} = 1$). Under the first assumption, the starting value of the intensity on a trading day equals to the last value of the previous trading day; conversely, under the second assumption, the starting value for the jump intensity on each trading day corresponds to the minimum level θ. For brevity, in Table 1.4 we only report the estimates performed under the fast mean reversion assumption since they provide the highest likelihood for all identification methods. The effect of jumps on future intensity exhibits a very small persistence: for the S&P 500 the half-life time ranges from 21 min to 1 hour, for the ESTX from 21 min to 1 hour and half. For both markets the s-CPR method has

the highest persistence, the intersection the lowest. When allowing for self-excitation, we obtain a remarkable increse of the likelihood under all jump identification methods suggesting that it represents a relevant feature to describe the statistical properties of jumps. When a jump occurs, its impact on the intensity is remarkably large, generally one order of magnitude larger then the baseline intensity level θ. These results are very similar to those obtained by Bormetti et al. (2015) analyzing the Italian stock market.

Model 3 also allows for spillovers in the jump activity (cross-excitation). To limit the number of free parameters we impose $\tilde{K}_{l,l'} = K_{l,l'}$ and we perform our estimates under the assumption of an equal mean reversion rate for the overnight and the intraday periods $\tilde{\gamma}_{l,l'} = \gamma_{l,l'}$ (note that leaving more freedom to the parameters does not lead to significant improvements). From the results reported in Table 1.4 we do not see any significant cross-excitation effect. However, under the s-CPR method, the standard tests reject the null $K_{l,l'} = 0$ at the 5% confidence level for both markets suggesting the possible presence of some cross market effect. Anyway, a further analysis anyway reveals that these weak cross dependencies immediately disappear once that the we allow the jump intensity to depend on the continuous volatility. According to our empirical evidence, jumps detected from high frequency do not seem to play a relevant role in volatility transmission across markets.

To extend our analysis, we also explore some additional features of the jump process: the dependence from continuous volatility and the role of the jump size. To this purpose we move from the standard Hawkes processes to the class of generalized Hawkes processes whose properties are discussed in details by Bowsher (2007). The generalized model allows the deterministic component θ to be time dependent and we use a power of normalized jump size as an explanatory in place of the increments to the counting process N_t. The full model is specified as follows:

$$\lambda_{l,\tau} = \begin{cases} \lambda_{l,o_{l,t}} + \theta_{l,\tau} + \sum_{l'=1}^{2} K_{l,l'} \int_{o_{l,t}}^{\tau} e^{-\gamma_{l,l'}(\tau-s)} \left| \frac{J_{l',s}}{\sqrt{v_{l',s}}} \right|^{\alpha} dN_{l',s} & \tau \in [o_{l,t}, c_{l,t}) \\ 0 & \tau \in [c_{l,t}, o_{l,t+1}) \end{cases} \tag{1.24}$$

$$\lambda_{l,o_{t+1}} = \lambda_{l,c_{l,t}} + \theta_{l,\tau} + \sum_{l' \neq 1}^{2} \tilde{K}_{l,l'} \int_{c_{l,t}}^{\tau} e^{-\tilde{\gamma}_{l,l'}(\tau-s)} \left| \frac{J_{l',s}}{\sqrt{v_{l',s}}} \right|^{\alpha} dN_{l',s} \tag{1.25}$$

where $\alpha \geqslant 0$. We consider the following parametrization for the deterministic time dependent

component $\theta_{l,\tau}$ used to accommodate an explicit dependence on the continuous volatility level:

$$\theta_{l,\tau} = \exp\left(a_l + b_l \log v_{l,\tau}\right) \quad a_l,\ b_l \in \mathbb{R} \tag{1.26}$$

where our volatility proxy is

$$v_{l,\tau(t,i)} = IV_t\, \zeta_i^2 / M \tag{1.27}$$

IV_t is the integrated volatility on day t , ζ_i is the intraday volatility corrector (see Appendix A.1) and M is the number of intraday returns. Equation 1.27 is a proxy for the instantaneous volatility on a specific time interval i on day t. The integrated volatility can be measured in different ways, the simplest and theoretically more efficient being the sum of all the squared returns that are not identified as jumps. Unfortunately, this choice is subject to an endogeneity bias as the error in its measurement unavoidably depends on the number of jumps detected and consequently by jump intensity itself. To circumvent this issue, we use a forecast of the integrated volatility built on the information available up to day $t-1$ and based on a bivariate extension of the HAR-type regressions of Corsi and Renò (2012). The results are discussed in Section 1.5 and show an extremely high in-sample forecasting power, meaning that the forecast values represent a good proxy for the realized volatility. At the same time, we observe significant volatility spillovers from U.S. to Europe with a time lag of one day. Interestingly, in the opposite direction we find significant lagged cross-leverage effects that are unprecedented in the literature to the best of our knowledge: negative returns in the Euro Stoxx 50 affect the volatility of the S&P 500 (probably such effects are mostly generated during the Euro Sovereign crisis).

The dependence on the jump size is introduced in our generalized Hawkes model when $\alpha > 0$ and it is determined by the absolute size of the jump normalized by the instantaneous volatility: the idea is that the impact of a jump is proportional to its size compared to the typical size of continuous returns on the same period. The Results for the alternative specifications are reported in Table 1.5 where the distinctive features of the generalized process are gradually introduced. Model 4 extends the univariate Model 2 and also introduces the volatility dependence under the constraints $\alpha = 0$, $K_{l,l'}$, $\tilde{K}_{l,l'} = 0$. Importantly, Table 1.5 shows that all estimates confirm a significant inverse dependence on the volatility level. This result is consistent with the analysis of Wei (2012) who finds that the volatility is on average lower on trading days with jumps. This result confirms what we

qualitatively observe in Figures 1.6 and 1.7: jumps at high frequency characterize mostly tranquil market conditions rather than periods of turmoil. From a statistical perspective, we must take into account that we are not analyzing the true (unobserved) jump process but the jumps detected by some non-parametric test. The inverse dependence that we measure could simply reflect the difficulty of such tests to detect jumps when the volatility is high. This may be the case if jump are i.i.d.: in presence of high volatility levels, the magnitude of continuous price fluctuations observed at a fixed sampling frequency may become close to the magnitude of jumps. The detection of discontinuities would then require a finer time resolution which is usually not achievable in practice due to the presence of the microstructure noise.

With regard to cross-market effects, according to further analysis not reported here for brevity, none of the cross-excitation coefficients is statistically significant when taking into account the continuous volatility and this holds regardless of the method used to detect jumps. Interestingly, when volatility is omitted, standard tests reject the null $K_{l,l'} = 0$ in some cases suggesting that cross-market dependence on the jump activity can be induced by the well documented cross-market volatility spillovers.

Concerning the role played by the jump size, we see that there is no agreement across the different detection methods: under the m-LM procedure, large jumps seem to have a larger impact on the intensity with a convex response ($\alpha > 1$); this effect disappears under the s-CPR method and for the intersection set. A possible motivation for this difference would be the presence of volatility jumps: contrary to the s-CPR method that is asymptotically robust to these events, the m-LM method is subject to an increase of the false detection rate. A rapid increase in volatility may be erroneously identified as a jump under the m-LM method which may lead to the kind of result that we have observed.

		ESTX			S&P 500		
		m-LM	s-CPR	Intersection	m-LM	s-CPR	Intersection
model 1	θ	$7.44 \cdot 10^{-4***}$	$4.16 \cdot 10^{-4***}$	$3.02 \cdot 10^{-4***}$	$4.90 \cdot 10^{-4***}$	$3.05 \cdot 10^{-4***}$	$1.85 \cdot 10^{-4***}$
		$(2.20 \cdot 10^{-5})$	$(1.38 \cdot 10^{-5})$	$(1.21 \cdot 10^{-5})$	$(1.81 \cdot 10^{-5})$	$(1.58 \cdot 10^{-5})$	$(9.68 \cdot 10^{-6})$
	$\log L$	-4147.614	-2518.953	-1911.414	-2208.438	-1467.042	-950.0483
model 2	θ	$6.32 \cdot 10^{-4***}$	$3.65 \cdot 10^{-4***}$	$2.81 \cdot 10^{-4***}$	$3.96 \cdot 10^{-4***}$	$2.37 \cdot 10^{-4***}$	$1.69 \cdot 10^{-4***}$
		$(2.86 \cdot 10^{-5})$	$(1.63 \cdot 10^{-5})$	$(1.86 \cdot 10^{-5})$	$(2.41 \cdot 10^{-5})$	$(1.85 \cdot 10^{-5})$	$(1.07 \cdot 10^{-5})$
	γ	$3.05 \cdot 10^{-2***}$	$7.47 \cdot 10^{-3***}$	$3.25 \cdot 10^{-2*}$	$2.90 \cdot 10^{-2***}$	$1.15 \cdot 10^{-2**}$	$3.30 \cdot 10^{-2*}$
		$(7.62 \cdot 10^{-3})$	$(1.68 \cdot 10^{-3})$	$(1.30 \cdot 10^{-2})$	$(6.67 \cdot 10^{-3})$	$(4.21 \cdot 10^{-3})$	$(1.38 \cdot 10^{-2})$
	$K_{l,l}$	$4.68 \cdot 10^{-3***}$	$1.21 \cdot 10^{-3***}$	$2.27 \cdot 10^{-3*}$	$5.93 \cdot 10^{-3***}$	$1.57 \cdot 10^{-3**}$	$3.06 \cdot 10^{-3*}$
		$(1.10 \cdot 10^{-3})$	$(2.63 \cdot 10^{-4})$	$(9.43 \cdot 10^{-4})$	$(1.37 \cdot 10^{-3})$	$(5.21 \cdot 10^{-4})$	$(1.43 \cdot 10^{-3})$
	$\log L$	-4062.034	-2489.863	-1895.516	-2122.551	-1448.232	-934.451
model 3	θ	$5.82 \cdot 10^{-4***}$	$3.48 \cdot 10^{-4***}$	$2.69 \cdot 10^{-4***}$	$3.76 \cdot 10^{-4***}$	$2.06 \cdot 10^{-4***}$	$1.64 \cdot 10^{-4***}$
		$(4.27 \cdot 10^{-5})$	$(1.90 \cdot 10^{-5})$	$(1.96 \cdot 10^{-5})$	$(2.91 \cdot 10^{-5})$	$(3.87 \cdot 10^{-5})$	$(1.55 \cdot 10^{-5})$
	$\gamma_{l,l}$	$3.06 \cdot 10^{-2***}$	$7.95 \cdot 10^{-3***}$	$3.24 \cdot 10^{-2*}$	$2.87 \cdot 10^{-2***}$	$1.19 \cdot 10^{-2**}$	$3.27 \cdot 10^{-2*}$
		$(7.70 \cdot 10^{-3})$	$(1.80 \cdot 10^{-3})$	$(1.33 \cdot 10^{-2})$	$(5.57 \cdot 10^{-3})$	$(4.34 \cdot 10^{-3})$	$(1.36 \cdot 10^{-2})$
	$K_{l,l}$	$4.68 \cdot 10^{-3***}$	$1.24 \cdot 10^{-3***}$	$2.53 \cdot 10^{-3*}$	$5.87 \cdot 10^{-3***}$	$1.56 \cdot 10^{-3**}$	$3.03 \cdot 10^{-3*}$
		$(1.11 \cdot 10^{-3})$	$(2.75 \cdot 10^{-4})$	$(9.47 \cdot 10^{-4})$	$(1.35 \cdot 10^{-3})$	$(5.24 \cdot 10^{-4})$	$(1.41 \cdot 10^{-3})$
	$\gamma_{l,l}$	$3.74 \cdot 10^{-4}$	$3.26 \cdot 10^{-4}$	$5.62 \cdot 10^{-4}$	$5.65 \cdot 10^{-3}$	$1.39 \cdot 10^{-4}$	$1.27 \cdot 10^{-3}$
		$(4.57 \cdot 10^{-4})$	$(2.65 \cdot 10^{-4})$	$(5.45 \cdot 10^{-4})$	$(4.47 \cdot 10^{-3})$	$(1.50 \cdot 10^{-3})$	$(1.31 \cdot 10^{-2})$
	$K_{l,l'}$	$2.39 \cdot 10^{-4}$	$1.16 \cdot 10^{-4}$	$2.26 \cdot 10^{-4}$	$1.6 \cdot 10^{-4}$	$7.70 \cdot 10^{-5}$	$3.00 \cdot 10^{-5}$
		$(2.16 \cdot 10^{-5})$	$(9.6 \cdot 10^{-5})$	$(2.15 \cdot 10^{-4})$	$(7.94 \cdot 10^{-4})$	$(4.12 \cdot 10^{-5})$	$(4.33 \cdot 10^{-4})$
	$\log L$	-4057.068	-2488.735	-1893.671	-2121.027	-1444.632	-934.156

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

Table 1.4: Estimates for the standard Hawkes process: time is measured in minutes and standard errors are reported in parenthesis.

		ESTX			S&P 500		
		m-LM	s-CPR	Intersection	m-LM	s-CPR	Intersection
model 4	a	$-1.22 \cdot 10^{1***}$	$-1.11 \cdot 10^{1***}$	$-1.20 \cdot 10^{1***}$	$-12.1 \cdot 10^{1***}$	$-1.08 \cdot 10^{1***}$	$-1.30 \cdot 10^{1***}$
		$(8.37 \cdot 10^{-1})$	$(1.00 \cdot 10^{-1})$	$(1.08 \cdot 10^{-1})$	$(9.30 \cdot 10^{-1})$	$(9.00 \cdot 10^{-1})$	(1.12)
	b	$-3.41 \cdot 10^{-1***}$	$-2.23 \cdot 10^{-1**}$	$-2.71 \cdot 10^{-1***}$	$-2.94 \cdot 10^{-1***}$	$-1.77 \cdot 10^{-1**}$	$-3.01 \cdot 10^{-1**}$
		$(5.86 \cdot 10^{-2})$	$(7.09 \cdot 10^{-2})$	$(7.63 \cdot 10^{-2})$	$(6.44 \cdot 10^{-2})$	$(6.13 \cdot 10^{-2})$	$(7.80 \cdot 10^{-2})$
	γ	$3.18 \cdot 10^{-2***}$	$5.74 \cdot 10^{-2***}$	$3.29 \cdot 10^{-2*}$	$3.07 \cdot 10^{-2***}$	$1.22 \cdot 10^{-2**}$	$3.56 \cdot 10^{-2*}$
		$(8.47 \cdot 10^{-3})$	$(1.70 \cdot 10^{-3})$	$(1.70 \cdot 10^{-2})$	$(7.78 \cdot 10^{-3})$	$(4.32 \cdot 10^{-3})$	$(1.62 \cdot 10^{-2})$
	$K_{l,l}$	$4.63 \cdot 10^{-3***}$	$1.05 \cdot 10^{-3***}$	$2.20 \cdot 10^{-3*}$	$6.07 \cdot 10^{-3***}$	$1.57 \cdot 10^{-3**}$	$3.18 \cdot 10^{-3*}$
		$(1.12 \cdot 10^{-3})$	$(2.42 \cdot 10^{-3})$	$(9.49 \cdot 10^{-4})$	$(1.50 \cdot 10^{-3})$	$(5.22 \cdot 10^{-4})$	$(1.56 \cdot 10^{-3})$
	$\log L$	-4039.128	-2483.786	-1888.9	-2111.314	-1445.279	-929.475
model 5	a	$-1.21 \cdot 10^{1***}$	$-1.11 \cdot 10^{1***}$	$-1.19 \cdot 10^{1***}$	$-1.20 \cdot 10^{1***}$	$-1.08 \cdot 10^{1***}$	$-1.31 \cdot 10^{1***}$
		$(8.29 \cdot 10^{-1})$	$(1.00 \cdot 10^{-1})$	$(9.35 \cdot 10^{-1})$	$(9.30 \cdot 10^{-1})$	(1.09)	(1.28)
	b	$-3.35 \cdot 10^{-1***}$	$-2.23 \cdot 10^{-1**}$	$-2.71 \cdot 10^{-1***}$	$-2.92 \cdot 10^{-1***}$	$-1.77 \cdot 10^{-1*}$	$-3.04 \cdot 10^{-1***}$
		$(5.82 \cdot 10^{-2})$	$(7.09 \cdot 10^{-2})$	$(6.65 \cdot 10^{-2})$	$(6.43 \cdot 10^{-2})$	$(7.49 \cdot 10^{-2})$	$(8.81 \cdot 10^{-2})$
	γ	$3.22 \cdot 10^{-2***}$	$5.74 \cdot 10^{-2***}$	$346 \cdot 10^{-2*}$	$3.00 \cdot 10^{-2***}$	$1.27 \cdot 10^{-2}$	$4.00 \cdot 10^{-2*}$
		$(7.87 \cdot 10^{-3})$	$(1.70 \cdot 10^{-3})$	$(1.73 \cdot 10^{-3})$	$(7.68 \cdot 10^{-3})$	$(7.66 \cdot 10^{-3})$	$(1.96 \cdot 10^{-2})$
	$K_{l,l}$	$1.90 \cdot 10^{-4}$	$1.05 \cdot 10^{-3***}$	$6.41 \cdot 10^{-4*}$	$2.91 \cdot 10^{-4}$	$9.40 \cdot 10^{-4}$	$5.40 \cdot 10^{-4}$
		$(1.27 \cdot 10^{-4})$	$(2.42 \cdot 10^{-3})$	$(8.35 \cdot 10^{-4})$	$(2.19 \cdot 10^{-4})$	$(9.20 \cdot 10^{-4})$	$(8.55 \cdot 10^{-4})$
	α	1.78^{***}	–	$6.89 \cdot 10^{-1}$	1.68^{***}	$3.39 \cdot 10^{-1}$	1.00
		$(3.06 \cdot 10^{-1})$	–	$(6.27 \cdot 10^{-1})$	$(3.49 \cdot 10^{-1})$	$(5.60 \cdot 10^{-1})$	$(6.82 \cdot 10^{-1})$
	$\log L$	-4023.449	-2483.786	-1888.446	-2096.817	-1445.112	-928.376

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

Table 1.5: Estimates for the extended univariate Hawkes process: time is measured in minutes and standard errors are reported in parenthesis. Note that for model 5 for under the s-CPR method the constraint $\alpha \geqslant 0$ is binding.

Co-jumps

Hawkes processes are designed to capture jump clustering effects across markets but they fail to capture the simultaneous occurrence of jump events that are commonly referred to as co-jumps. The occurrence of co-jumps for stocks traded on the same market has been largely documented by Gilder et al. (2014), Bormetti et al. (2015), Calcagnile et al. (2015) among others and specific statistical tests have been recently designed for their detection (see Jacod and Todorov 2010 and Caporin et al. 2014). To the best of our knowledge, co-jumps involving market indexes located on different geographical regions have not yet been studied by the literature. One reason is probably that different markets are often operating asynchronously and the overlap of their activity represents just a small portion of the trading day. This is exactly our case: the overlap ranges between one

and three hours (due to differences in the adoption of light saving time) and it normally amounts to a couple of hours. A systematic study of co-jumps is definitely beyond the purposes of this chapter, nevertheless the simultaneous occurrence of jumps is sometimes observed in our sample. To keep our approach simple, we base our investigation on the co-exceedance criterion whereby a co-jump is detected whenever jumps are detected in the price of both indexes within the same time interval. According to this rule the number of co-jumps under the m-LM, the s-CPR and their intersection is respectively equal to 18, 13 and 7. Given the small number of events it is natural to question whether they can be generated by random statistical fluctuations. In this regard, we first notice that all the identified co-jumps have the same sign in both markets suggesting that they are effectively originated by the same type of information. However, to establish their significance from a statistical perspective we need to perform a formal test. In practice the co-exceedance criterion is applied on a discrete timeline whose granularity is 5 minutes and the probability to observe jumps falling within the same time interval is clearly larger than zero even when jumps occur independently conditionally on the intensities. Assume Δt is a small positive time interval, under the null hypothesis that the two processes are independent with intensities $\lambda_{1,t}$ and $\lambda_{2,t}$, the probability of observing a co-jump in the interval $[t,\, t+\Delta t]$ is approximately $1-\exp\left(-\lambda_{1,t}\cdot\lambda_{2,t}\cdot\Delta t^2\right)$. The test statistics is represented by the total number of co-jumps n that under the null is a Poisson random variable with mean $\Lambda=\sum_{i=1}^{T}\lambda_{1,i}\cdot\lambda_{2,i}\cdot\Delta t^2$ where we assume $\lambda_{l,t}=0$ when market l is closed. To calculate the test some reliable estimates of the jump intensity on both markets are needed and, to this purpose, we use estimates from Model 5 for the m-LM method (since in this case the coefficient α is significant) and Model 4 in all other cases. To check the reliability of these estimates we first apply the specification tests proposed by Bowsher (2007)[5].

According to our results, the null hypothesis of correct model specification is generally not rejected at the 5% significance level for all the jump detection methods adopted. Conversely, the hypothesis that all the co-jumps observed are generated by random statistical fluctuations is always rejected

[5] The main idea behind these tests consists in the application of a random time change defined as $\tilde{t}_l\left(t\right)=\int_0^t\lambda_{l,s}ds$. In the new time coordinates, under the null of a correct model specification, each jump processes behave as a Poisson processes with constant unitary intensity. The residuals, defined as $e_{l,j}=\int_{\tau_{j-1}}^{\tau_j}$ with $\tau_{l,0}=0$ and $\tau_{l,1}\ldots,\tau_{l,N_l}$ representing the jump times on market l, are i.i.d. exponentially distributed under the null. In practice, independence is verified applying the Ljung-Box test to residuals and squared residuals up to 15 lags; exponentiality can be checked applying the Engle and Russel (1998) test for excess dispersion.

at the 0.1% significance level.

1.5 Forecasting the Integrated Volatility

In this Section we discuss some details on the volatility proxy used in Section 1.4. The objective is to provide accurate estimates of the integrated volatility on each trading day t using the previously available information. The aim of this approach is to circumvent some endogeneity issues which may arise if the integrated volatility is calculated subtracting the contribution of jumps from the quadratic variation introducing therefore some dependence between the volatility estimates and the price jumps occurred in the same period.

The dependent variables that we want to model are the integrated volatility of the Euro Stoxx 50 ($IV_{EU,t}$) and of the S&P 500 ($IV_{EU,t}$). Let $r_{l,t}$ denote the close to close log-return on day t while $J_{l,t}$ is the absolute contribution of jumps to the quadratic variation. More specifically the integrated volatility is calculated as the average variance of the intraday log-returns that are not identified as jumps:

$$IV_{l,t} = \frac{M}{M - \sum_{i=1}^{M} Jump_{t,i}} \sum_{i=1}^{M} r_{t,i}^2 \left(1 - Jump_{t,i}\right)$$

where $Jump_{t,i}$ is the jump indicator taking the value 1 each time that a jump is identified and zero otherwise. Clearly, this method produces distinct volatility measures for each different jump identification method. Our approach is based on a bivariate extension of the LHAR-C-CJ regression of Corsi and Renò (2012), i.e. a parsimonious regression of the HAR type (introduced by Corsi, 2009) which also includes lagged leverage effects and jumps. We propose a straightforward bivariate extension that is able to capture the cross market effects. As an example, for the Euro Stoxx 50, the regression reads as follows:

$$\begin{aligned}
\log IV_{EU,t+1} &= c + \beta_1 \log IV_{EU,t} + \beta_2 \log IV^{(5)}_{EU,t} + \beta_3 \log IV^{(22)}_{EU,t} + \beta_4 \log\left(1 + J_{EU,t}\right) + \beta_5 \log\left(1 + J^{(5)}_{EU,t}\right) \\
&+ \beta_6 \log\left(1 + J^{(22)}_{EU,t}\right) + \beta_7 r^{-}_{EU,t} + \beta_8 r^{(5)-}_{EU,t} + \beta_9 r^{(22)-}_{EU,t} + \beta_{10} \log IV_{US,t} \\
&+ \beta_{11} \log IV^{(5)}_{US,t} + \beta_{12} \log IV^{(22)}_{US,t} + \beta_{13} \log\left(1 + J_{US,t}\right) + \beta_{14} \log\left(1 + J^{(5)}_{US,t}\right) \\
&+ \beta_{15} \log\left(1 + J^{(22)}_{US,t}\right) + \beta_{16} r_{US,t} + \beta_{17} r^{(5)-}_{US,t} + \beta_{18} r^{(22)-}_{US,t}
\end{aligned}$$

where for a generic observable X we have:

$$\begin{aligned}
X^{-} &= \min\left(X, 0\right) \\
X^{(h)}_t &= \frac{1}{h} \sum_{j=1}^{h} X_{t-h+1} \\
X^{(h)-} &= \min\left(X^{(h)}, 0\right) \\
X^{-(h)} &= \frac{1}{h} \sum_{j=1}^{h} \min\left(X_{t-h+1}, 0\right)
\end{aligned}$$

with h representing the order of the LHAR-C-CJ component.

The results shown in Table 1.6 are very similar under the alternative jump identification schemes adopted. For both indexes the strong volatility persistence is confirmed: the coefficients relative to the daily, weekly and monthly components are positive and significant. The persistence of the leverage effects is also relevant, especially for the Euro Stoxx 50 index. The main differences with respect to Corsi and Renò (2012) are found in the impact of jumps on continuous volatility which is generally insignificant in our regressions. This is probably due to the prevalence of crisis periods in our sample, when the effect of jumps is overwhelmed by a large continuous volatility component as already documented in the rest of this chapter. Concerning volatility spillovers, we note a strong effect from U.S. to Europe with a lag of 1 day while the weekly component has a weak negative effect and the monthly component is not statistically significant. Importantly we also notice a marked cross-leverage effect between Europe and U.S. that to the best of our knowledge is unprecedented in the literature and suggests a possible direction for future research. The interdependence in volatility

stems also from the cross correlation of the residuals that is over 40% under all jump identification methods.

	ESTX			S&P 500		
	m-LM	s-CPR	m-LM ∩ s-CPR	m-LM	s-CPR	m-LM ∩ s-CPR
c	−1.33***	−1.44***	−1.48***	−1.53***	−1.72***	−1.73***
	(0.20)	(0.21)	(0.20)	(0.20)	(0.20)	(0.20)
IV_{US}	0.22***	0.24***	0.23***	0.44***	0.39***	0.39***
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
$IV_{US}^{(5)}$	−0.10	−0.09	−0.09	0.24***	0.27***	0.27***
	(0.05)	(0.06)	(0.06)	(0.05)	(0.05)	(0.06)
$IV_{US}^{(22)}$	−0.07	−0.10*	−0.10*	0.25***	0.27***	0.27***
	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)
J_{US}	11.21	373.41	475.28	32.60	188.37	526.30
	(367.65)	(355.44)	(381.62)	(312.84)	(606.05)	(609.65)
$J_{US}^{(5)}$	374.63	495.97	299.76	186.61	−28.31	−505.50
	(647.56)	(1204.55)	(1276.56)	(804.68)	(1412.76)	(1472.98)
$J_{US}^{(22)}$	−1309.15	−3032.51	−2223.71	−1299.17	355.15	1418.30
	(1436.40)	(2858.22)	(2953.36)	(1439.30)	(2871.98)	(2913.64)
r_{US}^{-}	−1.05	−0.78	−0.97	−6.70***	−7.26***	−7.44***
	(1.61)	(1.76)	(1.74)	(1.96)	(2.07)	(2.08)
$r_{US}^{(5)-}$	0.79	−0.20	0.53	−12.19	−12.70	−12.81
	(5.61)	(6.06)	(6.01)	(6.41)	(6.72)	(6.57)
$r_{US}^{(22)-}$	25.98	26.96	29.41*	8.89	19.18	22.56
	(15.47)	(15.15)	(14.82)	(17.04)	(17.98)	(18.09)
J_{EU}	−299.77	−288.77	−297.08	−530.12*	−450.34	−400.91
	(264.21)	(265.17)	(273.65)	(259.97)	(311.68)	(339.78)
$J_{EU}^{(5)}$	−657.64	−722.77	−707.31	−1015.27	−1334.81	−1486.38
	(727.60)	(728.93)	(790.90)	(849.13)	(888.68)	(909.46)
$J_{EU}^{(22)}$	−352.97	−264.96	−108.15	−178.40	−569.12	−360.00
	(1252.21)	(1308.38)	(1318.49)	(1367.05)	(1455.09)	(1492.77)
IV_{EU}	0.25***	0.19***	0.19***	−0.01	−0.02	−0.02
	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
$IV_{EU}^{(5)}$	0.35***	0.38***	0.38***	0.00	0.02	0.02
	(0.05)	(0.05)	(0.05)	(0.05)	(0.06)	(0.06)
$IV_{EU}^{(22)}$	0.22***	0.24***	0.23***	−0.07	−0.09	−0.09
	(0.05)	(0.05)	(0.05)	(0.05)	(0.06)	(0.06)
r_{EU}^{-}	−12.60***	−13.60***	−13.26***	−10.52***	−10.28***	−10.05***
	(1.64)	(1.73)	(1.75)	(1.67)	(1.76)	(1.75)
$r_{EU}^{(5)-}$	−23.10***	−22.90***	−23.44***	−14.58*	−16.91**	−16.72**
	(5.01)	(5.20)	(5.19)	(6.18)	(6.32)	(6.31)
$r_{EU}^{(22)-}$	−34.97**	−35.51**	−36.17**	−31.36*	−37.80**	−37.79**
	(12.04)	(12.46)	(12.43)	(13.90)	(14.53)	(14.38)
R^2	0.81	0.79	0.79	0.85	0.84	0.84
obs.	1669	1669	1669	1652	1652	1652
RMSE	0.40	0.43	0.42	0.44	0.46	0.46

***$p < 0.001$, **$p < 0.01$, *$p < 0.05$

Table 1.6: Estimates for the bivariate LHAR-C-CJ, robust to heteroskedasticity and autocorrelation. Standard errors in parentheses.

1.6 Concluding Remarks

We studied the statistical properties of the bivariate jump process involving two stock markets represented by the Euro Stoxx 50 and the S&P 500 index. Our analysis reveals important features characterizing jump events with potentially relevant implications for asset price modeling, derivatives pricing, risk management and portfolio optimization. The empirical evidence is not compatible with a constant intensity Poisson process but reveals significant jump clustering effects that are captured by the self-exciting Hawkes processes. The time persistence of such self-excitations is extremely short (less than 1 hour) and therefore unable to produce measurable effects across different trading days, consistently with what observed by Bajgrowicz et al. (2016). Concerning cross market effect, our estimates exclude the presence of significant spillovers in the jump activity but at the same time they reveal a strong dependence of Euro Stoxx 50 returns from news related to the U.S. economy. Such a dependence is reflected in the shape of the intraday volatility pattern and by the presence of few but statistically significant co-jumps occurring when both markets are simultaneously operating. Our results are complementary to those of Corradi et al. (2012) who document significant spillovers across markets affecting the continuous volatility and confirm the absence of cross-market effects triggered by jumps. Interestingly, our results appear in contrast with those of ADL obtained from daily data on the MSCI indexes. In our view the discrepancy is likely determined by the different time scale at which our analyses are performed and by the specific parametric assumptions of ADL: large daily returns that are captured by price jumps in their can be generated by large volatility shocks. Indeed, our findings suggest that jumps are more likely detected when the continuous volatility is low. The inverse dependence between the jump intensity and the continuous volatility can be related to a mere identification problem: when the volatility is high, the power of the test is reduced because jumps are more difficult to distinguish from continuous returns at a given sampling frequency. However, our results seem to exclude the possibility that jumps played a prominent role during the sub-prime crisis in 2008-2009 and the Euro Sovereign crisis in 2011-2012. These periods are instead characterized by high volatility peaks which overwhelm the jump component in terms of relative contribution to the total price variance. Our analysis provides several directions for future research. Interestingly, our supplementary regres-

sions used to forecast volatility in Section 1.5 show significant cross-leverage effect between Europe and U.S. that suggest a volatility transmission mechanism which, to the best of our knowledge, is still unexplored. Our results point towards a crucial role volatility shocks rather than price jumps in the period 2007-2014, this is another aspect which in our view deserves further investigations.

Appendix A

Technical Details of Chapter 1

A.1 The Intraday Volatility Pattern

It is well established that the stock market volatility tends to be higher at the beginning and at the end of the trading day (seeHarris 1986, Wood et al. 1985 for seminal contributions). Therefore, it is essential to take into account the intraday volatility pattern for the purpose of jumps identification as pointed out by Boudt et al. (2011). Several methods have been proposed in the literature to estimate the intraday volatility correction factor. As an example, Taylor and Xu (1997) use the simple estimator $\hat{\zeta}^2_{TX,i}$ based on the realized volatility measure:

$$\hat{\zeta}^2_{TX,i} = \frac{M \sum_{t=1}^{T} r^2_{t,i}}{\sum_{t=1}^{T}\sum_{i=1}^{M} r^2_{t,i}} \tag{A.1}$$

Andersen and Bollerslev (1997) propose a more sophisticated technique called flexible Fourier function (FFF) that is based on the following regression:

$$\log|r_{t,i}| - c = \mathbf{x}_i'\theta + \epsilon_{t,i} \tag{A.2}$$

where c corresponds to the mean of the log absolute value of a standard normal random variable and

$$\begin{aligned} \mathbf{x}_i'\theta \quad = \quad & \sum_{q=0}^{Q} \sigma_t^q \left[\mu_{0,q} + \mu_{1,q}\frac{i}{N_1} + \mu_{1,q}\left(\frac{i}{N_2}\right)^2 \right. \\ & \left. + \sum_{l=1}^{D} \lambda_{l,q}\, I_{\{i=d_l\}} + \sum_{p=1}^{P} \left(\gamma_{p,q} \cos\frac{p\cdot i\cdot 2\pi}{M} + \kappa_{p,q}\sin\frac{p\cdot i\cdot 2\pi}{M} \right) \right] \end{aligned} \tag{A.3}$$

where $\theta = [\mu_{0,q}, \mu_{1,q}, \lambda_{l,q}, \gamma_{p,q}, \kappa_{p,q}]$ is a parameter vector, σ_t is a measure of the daily volatility level, $N_1 = (N+1)/2$ and $N_2 = (N+1)(N+2)/6$. The FFF parametrization is based on quadratic component and sinusoids, for $Q > 1$ it allows to take into account interactions between the daily volatility level and the intraday periodicity. For $D > 0$ this procedure allows to introduce specific dummies if some intraday intervals do not fit properly within the regular pattern. The regression is estimated by OLS and the intraday volatility corrector is obtained as

$$\hat{\zeta}^2_{FFF,i} = \frac{M \exp(2\mathbf{x}_i'\theta)}{\sum_{i=M}^{M} \exp(2\mathbf{x}_i'\theta)} \tag{A.4}$$

Importantly, neither the estimators described above nor the method proposed by Bormetti et al. (2015) and reported in Equation 1.16 are robust to the presence of price jumps. So, if price discontinuities are concentrated in specific periods within the trading day, they may induce some distortions in the estimate of the intraday volatility corrector and consequently in the instantaneous volatility measurement. The problem is discussed by Boudt et al. who propose alternative parametric and non-parametric estimators. Let us first consider the standardized returns defined as follows:

$$\bar{r}_{t,i} = \frac{r_{t,i}}{\sqrt{BV_t/M}}$$

the shortest half scale estimator is defined as

$$ShortH_i = 0.741 \min\left\{\bar{r}_{(h),i} - \bar{r}_{(1),i}, \ldots, \bar{r}_{(T),i} - \bar{r}_{(T-h+1),i}\right\}$$

where T is the total number of observations, $h = \lfloor T/2 \rfloor$ is the floor of $T/2$ and $\bar{r}_{(j),i}$ are the order statistics of $\bar{r}_{j,i}$. Then the corresponding correction factor is

$$\hat{\zeta}^2_{ShotH,i} = \frac{M\, ShortH_i^2}{\sum_{i=1}^{M} ShortH_i^2}$$

A more efficient non-parametric corrector is the weighted standard deviation estimator that assigns no weight to the largest observations after scaling by $\hat{\zeta}_{ShortH,i}$:

$$WSD_i^2 = 1.081 \frac{M \sum_{t=1}^{T} \omega_{t,i}\, \bar{r}^2_{t,i}}{\sum_{i=1}^{M} \omega_{t,i}}$$

where $\omega_{t,i} = \omega\left(\tilde{r}_{t,i}/\hat{\zeta}_{ShortH,i}\right)$ and $\omega(z) = 1$ if $z^2 \leqslant 6.635$ and 0 otherwise[1]

$$\hat{\zeta}^2_{WSD,i} = \frac{M \cdot WSD_i^2}{\sum_{i=1}^{M} WSD_i^2} \tag{A.5}$$

The parametric method proposed by Boudt et al. represents a modification of the FFF estimator. To see this, consider the residuals

$$e_{t,i}^{WSD} = \log|r_{t,i}| - c - \log\hat{\zeta}_{WSD,i}$$

and define the negative likelihood function as

$$\rho_{ML}(z) = -0.5\,\log\left(\frac{2}{\pi}\right) - z - c + 0.5\,\exp\{2(z+c)\}$$

and the weights as

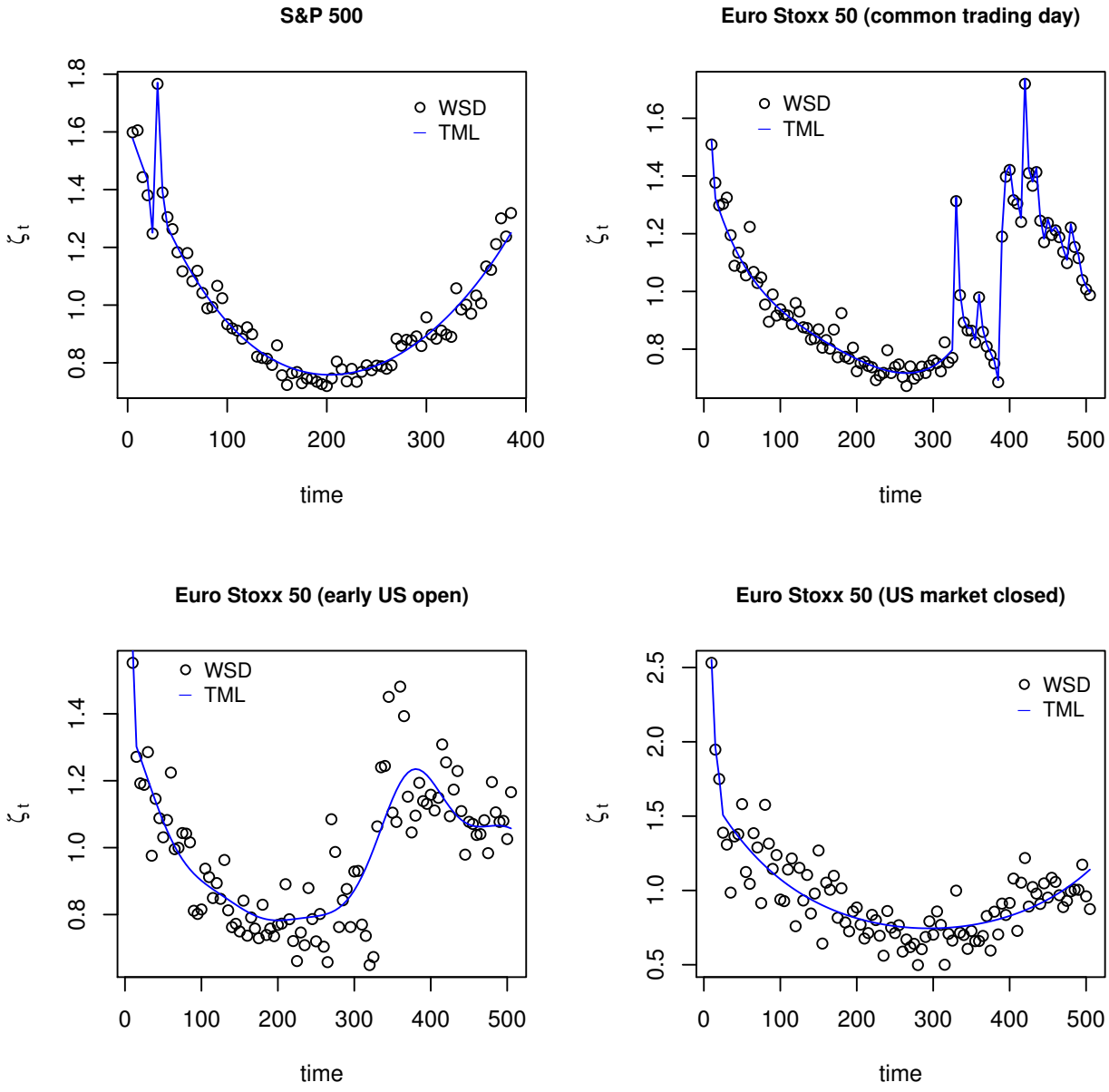
$$\omega_{t,i} = \begin{cases} 1 & \text{if } \rho_{ML}\left(e_{t,i}^{WSD}\right) \leqslant 3.36 \\ 0 & \text{otherwise.} \end{cases}$$

Then the maximum likelihood parameters are estimated as

$$\theta_{ML} = \min_{\theta} \frac{\sum_{t,i}\omega_{t,i}\rho_{ML}(\epsilon_{t,i})}{\sum_{t,i}\omega_{t,i}}$$

where $\epsilon_{t,i}$ is calculated from a regression of the type A.2 and the truncated maximum likelihood (TML) corrector is given by equation A.4.

[1] The threshold for z corresponds to the 99° percentile of a chi squared distribution with one degree of freedom.



FIGUREA.1: The blue line labeled as TML is the correction adopted in this chapter and is constructed using a mixed approach: for the points lying in the neighborhood of discontinuities the TML estimator is substituted by the WSD. On common days the TML approach for the Euro Stoxx 50 is adopted until 14:25 CET. The first spike at 14:30 is related to the beginning of pre-negotiations in the U.S., then the volatility decreases until the beginning of electronic trading at 15:30. The highest level is reached at 16:00 (10:00 EST) which is the usual time of news announcements concerning the U.S. economy. The left bottom panel represents the shape of the volatility correctors for periods when the NYSE opens one hour earlier w.r.t. the CET due to the different adoption of the daylight saving time in the two regions. This anomaly involves a small number of trading days (117 out of 1691) and the WSD estimates are extremely noisy, thus we fully rely on the TML estimates (except for the first 5 minutes interval when volatility is extremely large). When the U.S. market is closed (only 44 days in the whole sample), the intraday volatility in the Euro Stoxx assumes an L shape that is almost entirely fitted by the TML method.

A.1. THE INTRADAY VOLATILITY PATTERN

Figure A.1 shows the WSD and the TML volatility correctors computed on our sample. Generally speaking, the TML estimator is more efficient and also generates smoother patterns, but it fails to capture discontinuities. Therefore, we disregard the TML corrector for some specific time intervals as the one around 10:00 EST where the S&P 500 deviates from a standard U-shaped pattern and exhibits a spike due to the documented news announcement effect mentioned in Section 1.3.2. The intraday volatility of the Euro Stoxx 50 exhibits a strong dependence on the market activity in U.S.: we can identify three different patterns depending on the operating time of the NYSE (see the three Euro Stoxx 50 subplots in Figure A.1). Normally at the Frankfurt Stock Exchange (FSX) the electronic trading starts at 9:00 CET and closes at 17:30, the NYSE opens at 15:30 CET and closes at 22:00 CET (respectively 9:30 and 16:00 EST). The overlapping period between these markets can fluctuate based on changes resulting from the non-simultaneous adoption of the daylight saving time which generates short periods characterized by a different volatility pattern (left bottom panel in Figure A.1). Finally there is a small set of days in which the FSX market is operating normally but the NYSE remains closed (these days usually include some long weekends and national holidays in U.S.). The volatility pattern of the Euro Stoxx appears L-shaped during these periods.

Chapter 2

Non-parametric Jump Testing Based on Optimal Combinations of Robust Volatility Estimators

2.1 Introduction

The importance of price jumps to provide an accurate description of the evolution of asset prices has been discussed in Chapter 1 where the main references on this topic have been also provided (Section 1.1). The ability to disentangle jumps from continuous diffusion is extremely useful for risk management, portfolio optimization and derivatives pricing. The identification of the size and the time of the jumps allows to analyze their statistical properties whose knowledge can be extremely useful not only for investors and risk managers but also for regulators who are primarily interested in the orderly functioning of financial markets. In the last decade, several non-parametric tests constructed on high-frequency data have been developed to detect whether, in addition to the continuous diffusion, asset prices are also affected by jumps. The availability of high frequency data makes jumps identification an important and active area of research. Importantly, the detection of jumps can be performed in principle with an extremely high accuracy provided that the return process can be sampled at a arbitrarily high frequency. Due to the presence of microstructure imperfections unfortunately the exact identification of discontinuities in asset prices is practically unfeasible. For this reason, several alternative techniques have been proposed in the literature. The seminal work in this area is due to Barndorff-Nielsen and Shephard (2004, 2006) (BNS), who developed a non-parametric technique which relies on the comparison of two realized measures of volatility, one determined by continuous price changes and another including also jumps. At a later stage several additional jump robust volatility measures have been designed: Mancini (2009)

introduces a threshold based estimator, Corsi, Pirino, and Renò (2010) (CPR henceforth) define a modified version of the bipower variation of BNS with a reduced finite sample bias, Andersen, Dobrev, and Schaumburg (2012) (ADS) develop new measures based on nearest neighbor truncation with suitable finite sample properties. Christensen, Oomen, and Podolskij (2010) generalize the proposals of ADS introducing quantile based estimators that are also consistent in presence of market microstructure noise.

Jump testing heavily relies on the ability to measure the continuous return volatility. With reference to Chapter 1, two main families of tests can be conveniently distinguished. The BNS family is based on the same idea of BNS but possibly replacing the bipower variation with some alternative robust to jumps measure of price variability. This category includes the contributions of CPR, ADS and Podolskij and Ziggel (2010) (PZ henceforth). The second family is introduced by Lee and Mykland (2008) (LM henceforth) and is referred as the LM family (for more details we refer to Section 1.3). Outside these two families, it is important to mention also the "Swap Variance" tests of Jiang and Oomen (2008) (looking at the difference between relative returns and log-returns) and the test of Aït-Sahalia and Jacod (2009) based on absolute return moments calculated at different sampling frequencies.

The availability of a variety of tools devoted to the same purpose, poses some practical questions on their efficiency, their robustness to microstructure effects and on the coherency of their outcomes. The study of Schwert (2010) for instance highlights a large variability across jumps detected under alternative methods in real data, raising therefore questions about the kind of jumps that are possibly identified or, more radically, on the validity of the theoretical assumptions underlying the construction of the tests[1]. The Monte Carlo simulations of Huang and Tauchen (2005), Dumitru and Urga (2011) (DU hereafter) and Gilder et al. (2014) among others, show that the performances of the various tests are driven by the features of the data generating process and the sampling frequency of price observations but there is no general recipe to select an appropriate method. The recent studies of Christensen et al. (2014) and Bajgrowicz et al. (2016) reveal that many jumps detected at the common sampling frequency of 5 minutes are indeed generated by large price variations with an underlying continuous path that is discoverable analyzing tick-by-tick data. Christensen

[1] Particularly we refer to the assumption of finite jump activity.

et al. (2016) argue that such extremely rapid price changes are characterized by "drift bursts" and develop e formal tests for their detection from high frequency data. While the distinction between true price jumps and different kind of phenomena like drift burst can be extremely useful to regulators and to high frequency traders who can largely benefit from a deeper understanding of the market microstructure, for long term investors as well as for risk management and option pricing purposes such a distinction could be less insightful.

This work contributes to the high frequency econometrics literature adopting a new approach aimed at collecting information from multiple robust measures of realized volatility to construct new efficient estimators. The underlying idea is similar to Kolokolov and Renó (2016) (KS henceforth) who combine multipower variation measures to design efficient quarticity estimators. Anyway our study differs from KS in several ways: first we construct optimal combinations not only for the realized quarticity but also and more importantly for the realized volatility. Second we use measures that involve at most blocks of three consecutive returns as recommended by ADS to reduce the bias generated by stochastic volatility. Third, whenever we use thresholds to reduce the bias due to jumps, the measures involved depend weakly on the threshold level whose choice is unavoidably judgmental.

Although we allow the coefficients of our combinations to take negative values, our newly introduced estimators remain always positive within our analysis either when applied to simulated or to real data. Patton and Sheppard (2009) show that a linear combination of realized variance measures can perform better than each single estimator because it is able to collect more information and to average across differences generated by finite sample effects or imperfections in the market microstructure. Differently, this chapter focuses on the maximization of efficiency under robustness to jumps rather than to microstructure noise. To the best of our knowledge, the only existing measures that are able to reach a low asymptotic variance and to conjugate robustness to market microstructure noise with robustness to jumps are the realized volatility estimators (QRV) of Christensen et al. (2010). ADS show that unfortunately the performances of QRV measures in finite samples can be largely affected by the presence of a rapidly changing stochastic volatility, because they involve long blocks of consecutive intraday returns. On the contrary, the combined realized volatility estimators (CRV) proposed in this chapter, achieve comparable levels of asymptotic efficiency involving at

most three consecutive returns. The "locality" guarantees a high robustness to stochastic volatility (see Monte Carlo results in ADS) but at the cost of losing asymptotic consistency in presence of microstructure noise.

The second important contribution of this chapter is the creation a new class of tests belonging the BNS family and to analyze their performances in finite samples. This new tests can achieve substantially more power than other tests of the same type, with the possible exception of the PZ test which exhibits a very high power in numerical simulations (see PZ and DU). The reason is that the latter is constructed using the threshold estimator of Mancini (2009) which removes returns larger then a certain threshold. Importantly the variance of this estimator attains the Kramer's - Rao lower bound but it is extremely sensitive to the threshold level and entails therefore a serious risk of retaining jumps when the threshold is too high or to generate a negative bias if the threshold is too low. Such a risk becomes even more concrete if the accuracy of instantaneous volatility estimates is jeopardized by volatility jumps. Moreover the PZ test requires an external perturbation to allow the development of the necessary asymptotic theory, an element that introduces additional arbitrariness related to the features of this perturbation that must be set a priori. The CRV estimators proposed in this chapter instead involve the corrected threshold bipower and tripower variation that are weakly depending on the truncation level (see CPR), as it is desirable in empirical applications. At the 5 and 10 minutes frequency, the CRV-based tests can achieve the same power of the test of LM in numerical simulations without requiring the same restrictive assumptions to the volatility process. This aspect is extremely relevant because the method of LM is not robust to volatility jumps, whose importance is increasingly documented by the recent literature (see Todorov and Tauchen 2011, Jacod and Todorov 2010, Corsi and Renò 2012, Wei 2012, Christensen et al. 2014, Bandi and Renò 2016). Remarkably, our framework allows to establish a trade-off between the efficiency in jump detection and the robustness to different features of the data generating process by setting the criteria to drive the selection of the optimal weights.

The third contribution of this work is to the study the finite sample properties of different tests in a Monte Carlo simulation that starts from the baseline setting of a geometric Brownian motion and gradually introduces more realistic features like stochastic volatility with a U shaped intraday pattern, contamination by market microstructure noise and jumps with different characteristics.

Such a simulation exercise is to some extent complementary to the contribution of Dumitru and Urga (2011) because it takes into account the effects of the intraday periodicity of the diffusive price volatility and introduces also jump clustering through self-exciting Hawkes processes, coherently with what is observed in the empirical analysis of Chapter 1. Both aspects are particularly relevant in empirical applications: the extensive discussion of Rognlie (2010) shows that rapid variations of the intraday volatility may generate a negative bias in jump robust measures, pushing towards the spurious detection of jumps. Our Monte Carlo study suggests that removing the periodical intraday effects can significantly improve the performances of all the tests we analyse[2]. Concerning jump clustering we observe that this feature of the data generating process increases the probability to have multiple jumps occurring within the same block of consecutive returns, compromising therefore the finite sample properties of our realized measures. We propose a truncation method that allows to overcome these issue without altering the asymptotic theory.

The chapter proceeds as follows: Section 2.2 introduces the theoretical setting providing the main definitions and the theoretical results, Section 2.4 discusses some possible criteria to construct optimal combinations of measures, in Section 2.5 we introduce the class jump tests based on the CRV measure. Sections 2.6 and 2.7 present respectively our Monte Carlo study and an application to empirical data. Section 2.8 concludes.

2.2 Jump Robust Volatility Measures

The theoretical setting is the same of Section 1.3.1 based on the usual assumption that prices follow a continuous-time semi-martingale (which ensures no arbitrage) with finite activity jumps (equation 1.1). Each trading day t has duration 1 and $M+1$ log-prices $p_{t,0}, \ldots, p_{t,M+1}$ are observed at equally spaced times. The intraday log-returns are $r_{t,i} = p_{t_{i+1}} - p_{t,i}$ for $i = 1, \ldots, M$ and the realized variance on day t is just the sum of intraday squared returns according to equation 1.2. It is well known that for $M \to \infty$ the realized variance RV_t converges in probability to the quadratic variation which in turn equals the integrated variance in absence of jumps. Several jump robust estimators for the integrated variance have been proposed in the literature: the threshold estimator of Mancini (2009), the bipower and multipower variation of BNS, the threshold bipower variation of

[2] Boudt et al. (2011) provide this kind of evidence for some LM type test but do not study the impact of the intraday volatility pattern on tests of the BNS type.

CPR, the minimum realized variance (MinRV) and the median realized variance (MedRV) of ADS, the quantile realized variance of Christensen et al. (2010). I recall here the main definitions and properties for the measures that will be used in this chapter. The multipower variation generalizes expression 1.5:

$$MPV_t(\gamma_1,\ldots,\gamma_k) = \frac{M^{(\gamma_1+\ldots+\gamma_k)/2}}{\mu_{\gamma_1}\ldots\mu_{\gamma_k}(M-k+1)}\sum_{i=k}^{M}\prod_{j=1}^{k}|r_{t,i-j+1}|^{\gamma_j} \tag{2.1}$$

where $\gamma_1,\ldots,\gamma_k > 0$, $\mu_\gamma = \mathbb{E}(|u|^\gamma)$ given $u \sim N(0,1)$. The bipower variation (BPV) and the tripower variation (TPV) are the two important special cases obtained posing respectively $k = 2$, $\gamma_1 = \gamma_2 = 1$ and $k = 3$, $\gamma_1 = \gamma_2 = \gamma_3 = 2/3$. Another example is the staggered bipower introduced by Huang and Tauchen (2005) and based on staggered returns to achieve robustness also to market market microstructure noise:

$$BPVs_t = \mu_1^{-2}\sum_{i=3}^{M}|r_{t,i-2}|\,|r_{t,i}| \tag{2.2}$$

The $MinRV$ and the $MedRV$ are od ADS are instead based on nearest neighbor truncation:

$$MinRV_t = \frac{\pi}{\pi-2}\frac{M}{M-1}\sum_{i=2}^{M}\min(|r_{t,i-1}|,|r_{t,i}|)^2 \tag{2.3}$$

$$MedRV_t = \frac{\pi}{6-4\sqrt{3}+\pi}\frac{M}{M-2}\sum_{i=3}^{M}\mathrm{med}(|r_{t,i-2}|,|r_{t,i-1}|,|r_{t,i}|)^2 \tag{2.4}$$

In presence of jumps with finite activity, all the measures above converge in probability to the integrated volatility as $M \to \infty$. Under some technical conditions on the volatility process σ_t, ADS show that the a central limit theorem holds for the BPV, the $MinRV$, the $MedRV$ and the TPV in absence of jumps implying that their asymptotic distribution is a multivariate normal conditionally on the integrated quarticity. Nevertheless, in presence of jumps, the central limit theorem fails to hold for the bipower variation: the reason is that jumps introduce in this measure a bias of order $M^{-1/2}$; for the the MinRV and the MedRV instead the corresponding rate of decay for is M^{-1} while for the TPV is $M^{-2/3}$. The impossibility to develop an asymptotic theory represents a relevant limitation for the BPV both from the theoretical and the empirical perspective. The threshold

bipower of CPR is designed to overcame this problem introducing a truncation that reduces the upward bias in presence of jumps. More generally, the threshold multipower variation is defined as:

$$TMPV_t(\gamma_1,\ldots,\gamma_k) = \frac{M^{(\gamma_1+\ldots+\gamma_k)/2}}{\mu_{\gamma_1}\ldots\mu_{\gamma_k}(M-k+1)}\sum_{i=k}^{M}\prod_{j=1}^{k}|r_{t,i-j+1}|^{\gamma_j}\mathbf{1}_{\{|r_{t,i-j+1}|^2\leqslant\theta_{t,i-j+1}\}} \tag{2.5}$$

where θ must go to zero slower then $\sqrt{M^{-1}\log M}$ (i.e. the modulus of continuity of the Brownian motion). In practice authors recommend an adaptive threshold that depends on the instantaneous volatility: $\theta_{t,i} = c_\theta^2\,\hat{v}_{t,i}$. The TBPV estimator allows the extension of the asymptotic result 2.12 also in presence of jumps at the cost of requiring the choice of an appropriate threshold level which is to some extent arbitrary[3]. In absence of jumps the $TMPV$ is downward biased in finite samples since absolute returns that are larger then the threshold are systematically removed. This flaw is eliminated in the corrected version of this estimator:

$$C-TMPV_t(\gamma_1,\ldots,\gamma_k) = \frac{M^{(\gamma_1+\ldots+\gamma_k)/2}}{\mu_{\gamma_1}\ldots\mu_{\gamma_k}(M-k+1)}\sum_{i=k}^{M}\prod_{j=1}^{k}Z_{\gamma_j}(r_{t,i-j+1},\theta_{t,i-j+1}) \tag{2.6}$$

where the function Z_γ is defined by expression 1.11. In the$C-TMPV$ absolute returns exceeding the threshold are replaced by their conditional expected value under the normality assumption:

$$\mathbb{E}\left[|r_{t,i}|^\gamma|\,r_{t,i}^2 > c_\theta^2\right] = Z_\gamma\left(r_{t,i},c_\theta^2\hat{v}_{t,i};c_\theta\right)$$

In presence of jumps the threshold multipower variation measures are usually more accurate in finite samples then the multipower variation of Barndorff-Nielsen, Shephard, and Winkel.

2.2.1 Introducing New Measures

ADS observe that the existing jump robust measures exhibit a strong positive correlation and conclude that there is a little room for improving the asymptotic efficiency. In this section we enlarge therefore the set of volatility estimators introducing some new measures. On a stand alone basis, the estimators we are going to define are less accurate then the existing ones, but they are useful to form combinations with superior properties. In Section 2.4 we will show that the optimal combinations

[3] Anyway CPR show that the multipower variation is much less sensitive to the threshold level compared to the threshold estimator of Mancini (2009)

including existing and newly defined measures reach an asymptotic variance that is much closer to the Cramér–Rao lower bound then the bipower. Importantly, all the estimators we consider in this section involve blocks of at most three consecutive returns to ensure that locality is preserved. As pointed out by ADS, this property is extremely important to have a small bias in presence of stochastic volatility by ensuring that the volatility on average does not change significantly within the same block. The first two estimators we introduce are respectively the the threshold version of the staggered bipower variation of Huang and Tauchen (2005) and the staggered counterpart of the $MinRV$ labeled $MinRVs$:

$$TBPVs_t = \mu_1^{-2} \frac{M}{M-2} \sum_{i=3}^{M} |r_{t,i-2}| \, |r_{t,i}| \, \mathbf{1}_{\left\{|r_{t,i-2}|^2 \leqslant \theta_{t,i-2}\right\}} \mathbf{1}_{\left\{|r_{t,i}|^2 \leqslant \theta_{t,i}\right\}} \tag{2.7}$$

$$MinRVs_t = \frac{\pi}{\pi-2} \frac{M}{M-2} \sum_{i=3}^{M} \min\left(|r_{t,i-2}|, |r_{t,i}|\right)^2 \tag{2.8}$$

Then we consider a natural extension of the MinRV taking the minimum of three consecutive squared returns. Like the $MinRV$ and the $MedRV$, also this estimator belongs to the QRV class of Christensen et al. (2010):

$$Min3RV_t = \left[1 - \frac{6}{\pi}\left(1 - \frac{1}{\sqrt{3}}\right)\right]^{-1} \sum_{i=3}^{M} \min\left(|r_{t,i-2}|, |r_{t,i-1}|, |r_{t,i}|\right)^2 \tag{2.9}$$

The last measure is built on the product between the minimum and the median of three consecutive absolute returns:

$$\begin{aligned} MinMedRV_t \quad = \quad & \left[\frac{6 - 6\sqrt{2} + 6\sqrt{3}}{\pi}\right]^{-1} \sum_{i=3}^{M} \min\left(|r_{t,i-2}|, |r_{t,i-1}|, |r_{t,i}|\right) \\ & \qquad \operatorname{med}\left(|r_{t,i-2}|, |r_{t,i-1}|, |r_{t,i}|\right) \end{aligned} \tag{2.10}$$

The following propositions mirror propositions 1 and 2 of Andersen et al. (2012):

Proposition 1. Let the log-price p_t be represented by the jump-diffusion process 1.1 with finite activity. Assume further that μ_t is adapted and locally bounded, σ_t is adapted, càdlàg and larger than zero a.s. Then as $M \to \infty$, $TBPVs_t$, $MinRVs_t$, $Min3RV_t$ and $MinMedRV_t$ converge in

probability to the integrated variance $\int_{t-1}^{t} \sigma_s^2 ds$

Importantly, it can be easily checked that the bias generated by finite activity jumps decreases at the rate M^{-1} for all measures allowing the development of the necessary asymptotic theory.

2.2.2 Asymptotic Theory

The construction of optimal combinations of jump robust measures is based on the asymptotic properties stated by propositions 2 and 3 which extend the results of ADS

Proposition 2. Under the same assumptions of Proposition 1 except that jumps are ruled out, plus assumption A1 of Andersen et al. (2012) Appendix A on σ_t, as $M \to \infty$

$$\sqrt{M} \begin{pmatrix} RV_{t,M} - IV \\ BPV_{t,M} - IV \\ BPVs_{t,M} - IV \\ TPV_{t,M} - IV \\ MinRV_{t,M} - IV \\ MinRVs_{t,M} - IV \\ MedRV_{t,M} - IV \\ Min3RV_{t,M} - IV \\ MinMedRV_{t,M} - IV \end{pmatrix} \xrightarrow{stable\ D} MN\left(0, \begin{bmatrix} 2 & 2\mathbf{i}_8' \\ 2\mathbf{i}_8 & S_{IV} \end{bmatrix} IQ_t\right) \tag{2.11}$$

where MN denotes a mixed normal distribution, i.e. a normal distribution conditional on the realization of the (random) volatility path and $IQ_t = \int_{t-1}^{t} \sigma_s^4 ds$ is the integrated quarticity, $\mathbf{i}_k$ is a k-dimensional vector of ones and

$$S_{IV} = \begin{bmatrix} 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 \\ & 2.609 & 2.283 & 2.737 & 2.976 & 2.448 & 2.534 & 3.287 & 3.056 \\ & & 2.609 & 2.575 & 2.448 & 2.976 & 2.420 & 2.998 & 2.821 \\ & & & 3.061 & 3.138 & 2.885 & 2.609 & 3.867 & 3.504 \\ & & & & 3.810 & 2.751 & 3.048 & 4.204 & 3.773 \\ & & & & & 3.810 & 2.795 & 3.667 & 3.352 \\ & & & & & & 2.959 & 2.972 & 2.960 \\ & & & & & & & 5.950 & 4.862 \\ & & & & & & & & 4.252 \end{bmatrix} \tag{2.12}$$

As noted by Huang and Tauchen (2005) the covariance matrix 2.12 reflects a situation similar to the test of Hausman (1978): in absence of jumps, log-returns are asymptotically normal and the RV is the maximum likelihood estimator with a variance equal to the Cramér-Rao lower bound. The remaining measures are less efficient but are robust to the presence of jumps. The joint asymptotic distribution of the eight estimators in 2.12 is known also in presence of jumps provided that we replace the bipower with the threshold bipower (or equivalently with its corrected version defined

by equation 2.6).

Proposition 3. In presence of finite activity jumps the asymptotic joint distribution of $TBPV$, $TBPVs$, $TTPV$, $MinRV$, $MinRVs$, $MedRV$, $Min3RV$ and $MinMedRV$ is the same as in absence of jumps, i.e. a mixed multivariate normal with the covariance matrix obtained from 2.12.

Proof The proof of propositions 2 and 2.12 parallels the proof of propositions 1,2 and 3 in ADS and therefore is not reported here in order to save space. The first step is to consider the case with no jumps. The main idea is to approximate the sequence of standardized high frequency returns using an alternative sequence with piece-wise constant volatility across one or more consecutive returns. The proof for the no-jumps case proceeds showing that: i) the measures calculated on the approximating process converge in probability to those calculated on the exact process at the rate $M^{-1/2}$ as the number of intraday observations M goes to infinity[4]; ii) the corresponding measures calculated on the approximating process converge in distribution to a mixed multivariate normal centered on the integrated volatility with variance-covariance matrix given by expression 2.12 The extension to the general case including jumps follows directly from the fact that the bias they generate decays faster than $1/\sqrt{M}$. The technical details relative to the computation of the covariance matrix are deferred to the Appendix B.1.

Note that a similar result can be obtained also for the jump robust quarticity estimators. To this purpose we consider the following set of measures:

$$TTPQ_t = \mu_1^{-2} \frac{M}{M-2} \sum_{i=3}^{M} \left| r_{t,i-2} r_{t,i-1} r_{t,i} \right|^{3/2} \mathbf{1}_{\left\{ |r_{t,i-2}|^2 \leqslant \theta_{t,i-2}; |r_{t,i-1}|^2 \leqslant \theta_{t,i-1}; |r_{t,i}|^2 \leqslant \theta_{t,i} \right\}} \quad (2.13)$$

$$MinRQ_t \equiv \frac{3}{3\pi - 8} \frac{M^2}{M-1} \sum_{i=2}^{M} \min \left(|r_{t,i-1}|, |r_{t,i}| \right)^4 \quad (2.14)$$

$$MedRQ_t = \frac{3\pi}{9\pi + 72 - 52\sqrt{3}} \frac{M^2}{M-2} \sum_{i=3}^{M} \text{med} \left(|r_{t,i-2}|, |r_{t,i-2}|, |r_{t,i}| \right)^4 \quad (2.15)$$

[4] Our notation slightly differs from ADS to maintain consistency across chapters: here M is the number of intraday observations that they denote with N, while they use M to denote the number of consecutive returns with constant volatility in the approximating process.

$$MinRQs_t = \frac{3}{3\pi - 8} \frac{M^2}{M-2} \sum_{i=3}^{M} \min\left(|r_{t,i-2}|, |r_{t,i}|\right)^4 \tag{2.16}$$

$$Min3RQ_t = \frac{\pi}{3\pi - 24 + 26/\sqrt{3}} \frac{M^2}{M-2} \sum_{i=1}^{M-2} \min\left(|r_{t,i}|, |r_{t,i+1}|, |r_{t,i+2}|\right)^4 \tag{2.17}$$

$$\begin{aligned} MinMedRQ_t \quad \equiv \quad & \left[1 - \frac{4}{\pi\sqrt{3}}\right]^{-1} \frac{M^2}{M-2} \sum_{i=3}^{M} \min\left(|r_{t,i-2}|, |r_{t,i-1}|, |r_{t,i}|\right)^2 \\ & \qquad \text{med}\left(|r_{t,i-2}|, |r_{t,i-1}|, |r_{t,i}|\right)^2 \end{aligned} \tag{2.18}$$

The estimators 2.13, 2.14 and 2.15 are respectively the truncated threshold tripower quarticity of CPR, the minimum and the median realized quarticity of ADS. The others are constructed by analogy with the integrated volatility measures defined in 2.8, 2.9 and 2.10. The bias in presence of jumps for all these estimators decays again at the rate M^{-1}, thus their joint asymptotic distribution in presence of jumps converges to a mixed normal with the same covariance matrix that is obtained in case of a continuous diffusion with constant volatility:

$$\sqrt{M} \begin{pmatrix} TTPQ_{t,M} - IQ \\ MinRQ_{t,M} - IQ \\ MinRQs_{t,M} - IQ \\ MedRQ_{t,M} - IQ \\ Min3RQ_{t,M} - IQ \\ MinMedRQ_{t,M} - IQ \end{pmatrix} \xrightarrow{stable\ D} MN\left(0,\ S_{IQ} \int_{t-1}^{t} \sigma_s^8 ds\right) \tag{2.19}$$

where

$$S_{IQ} = \begin{bmatrix} 13.65 & 13.07 & 10.67 & 11.05 & 17.02 & 15.65 \\ & 18.54 & 8.53 & 14.80 & 16.78 & 15.78 \\ & & 18.54 & 11.63 & 12.83 & 12.25 \\ & & & 14.16 & 12.09 & 12.55 \\ & & & & 28.62 & 22.60 \\ & & & & & 19.72 \end{bmatrix} \tag{2.20}$$

The elements of the covariance matrix 2.20 are calculated following the same method to those of matrix 2.12 (analytical expressions have been used whenever available while we resorted to numerical integration in other cases).

2.3 Combining Multiple Measures

In this section we study the problem of constructing optimal combinations of robust estimators according to some criteria summarized in a specific loss function. Let $\mathbf{w}$ be a vector of weights with real unbounded components such that $\sum_{i=1}^{n} w_i = 1$ where n is the number of different measures involved (in our case $n = 8$ for the integrated volatility and $n = 6$ for the integrated quarticity).

$$\begin{aligned} CRV_t(\mathbf{w}) &= w_1 TBPV_t + w_2 TBPVs_t + w_3 TTPV_t + w_4 MinRV_t + \\ &\quad w_5 MinRV_t + w_6 Min3RV_t + w_7 Min3RV_t + w_8 MinMedRV_t \end{aligned} \tag{2.21}$$

$$\begin{aligned} CRQ_t(\mathbf{w}) &= w_1 TTPQ_t + w_2 MinRQ_t + w_3 MinRQs_t + \\ &\quad w_4 MedRQ_t + w_5 Min3RQ_t + w_6 MinMedRQ_t \end{aligned} \tag{2.22}$$

Proposition 4. As $M \to \infty$, under the same assumptions of Proposition 3 we have:

$$\sqrt{M}\left(CRV_{M,t} - IV_t\right) \xrightarrow{stable\,D} MN\left(0, \mathbf{w}^T S_{IV} \mathbf{w}\, IQ_t\right) \tag{2.23}$$

Proof The results follow directly from Proposition 3.

Analogously it can be proved that

$$\sqrt{M}\left(CRQ_{M,t} - IQ_t\right) \xrightarrow{stable\,D} MN\left(0, \mathbf{w}^T S_{IQ} \mathbf{w} \int_{t-1}^{t} \sigma_s^8 ds\right) \tag{2.24}$$

A natural criterion to determine the optimal weights is the minimization of the asymptotic variance under some linear constraints $A\mathbf{w} = \mathbf{b}$ where A is $k \times n$ and b is $k \times 1$. The minimum requirement is $\sum_i w_i = 1$ for consistency, but additional constraints can be useful in different circumstances: for instance one may exclude a subset of measures in order to save computational time or may suppress a specific finite sample bias estimated from a numerical simulation. The optimal estimator under the selected criterion is:

$$\mathbf{w}^* = S^{-1} A^T \left(A\, S^{-1}\, A^T\right)^{-1} \mathbf{b} \tag{2.25}$$

To highlight the gain in terms of asymptotic efficiency that comes from combining multiple measure,

the optimal weights have been computed for all possible combinations of realized volatility and realized quarticity estimators that we are considering. The asymptotic variance scaled by M is plotted in Figure 2.2 while the optimal weights are reported on Tables 2.4 and 2.3. Note that the best combination of two realized volatility estimators, which involves the threshold bipower and the $MinRV$, achieves an asymptotic variance that is remarkably smaller than the bipower alone. The improvements become less pronounced as the number of estimators involved increases and we get closer to the maximum likelihood limit. Note that such a limit is attained asymptotically by the threshold estimator of Mancini (2009) but as already mentioned, in practice this estimator is extremely sensitive to the threshold level and to the error affecting the estimates of the local volatility. The CRV instead depends weakly on the threshold through $TBPV$ and the $TTPV$ (see CPR for further details). For the realized quarticity we also observe significant improvements as the number of measures increases, although the optimal estimator obtained with 6 measures has a variance that is still well above the lower bound attained by the efficient multipowers of KR. The variance of CRQ could be further decreased including estimators that involve four or more consecutive returns but this would compromise the locality incrementing the bias in presence of stochastic volatility. Importantly, both the CRV and the CRQ type of estimators can be unbounded and are allowed to assume also negative values in finite samples. Anyway we will show that this happens extremely rarely in practice and moreover is not problematic for the construction of efficient jump tests which is the main focus of this work.

The minimization of the asymptotic variance is just one possible criterion. In empirical applications it can be also useful to take into account also some finite characteristics. The weights in that case can depend also on the specific features of the data generating process (the variability of the instantaneous volatility, the jump intensity and the jump size distribution, the noise to signal ratio) but the asymptotic consistency of the estimators is always guaranteed. A concrete example will be provided later.

Threshold adjustment for quantile based estimators

When combining multiple measures with possibly negative coefficients, we face the problem of obtaining negative estimates for quantities, like the realized volatility, that are non-negative by

definition. Theoretically, the lower is the variance of our measures, the smaller is the probability to obtain negative values. In practice we must also consider some possible distortions emerging in finite samples. A crucial problem is the occurrence of multiple jumps within the same block of consecutive returns. Such events, referred as "gradual jumps" in the terminology of Barndorff-Nielsen et al. (2009), are non negligible in real data as pointed out also by CPR. Christensen et al. (2014) and Christensen et al. (2016) find that their are substantially different from jumps because they are generated by a continuous path dominated by a strong drift component. These events are usually indistinguishable from price jumps unless we use tick-by-tick data. For the purposes of our analysis we need to control the large bias generated in case of gradual jumps. While for the truncated and corrected versions of the multipower variation (equations 2.5 and 2.6) this is guaranteed by the presence of a threshold, the measures based on nearest neighbor truncation can be largely affected by the presence of consecutive abnormal returns. To mitigate this problem we set a correction with no effects on the asymptotic results. Let us generally write our measure as

$$F_k(\mathbf{r}) = (M-k)^{-1} \sum_{i=k}^{M} \Theta\left(|r_{i-k,t}|, \ldots, |r_{i,t}|\right)$$

where $\mathbf{r} = (r_{1,t}, \ldots, r_{M,t})^T$. We substitute this expression with

$$\begin{aligned} F_k(\mathbf{r}) \quad = \quad & \frac{1}{M-k} \sum_{i=k}^{M} \Big[\Theta\left(|r_{i-k,t}|, \ldots, |r_{i,t}|\right) \mathbf{1}_{\left\{\Theta\left(|r_{i-k,t}|,\ldots,|r_{i,t}|\right) \leqslant \tau(\Theta) \max\left(v_{i-k,t},\ldots,v_{i,t}\right)\right\}} + \\ & \max\left(v_{i-k,t}, \ldots, v_{i,t}\right) \mathbf{1}_{\left\{\Theta\left(|r_{i-k,t}|,\ldots,|r_{i,t}|\right) \leqslant \tau(\Theta) \max\left(v_{i-k,t},\ldots,v_{i,t}\right)\right\}} \Big] \end{aligned} \tag{2.26}$$

where $\tau(\Theta)$ is a threshold that goes to zero slower than the modulus of continuity of the Brownian motion as $M \to \infty$. In practice we set M at the 99.99% of the distribution of $\Theta\left(|U_1|, \ldots, |U_k|\right)$ where $U_1, \ldots, U_k$ are i.i.d. standard normals and it is determined numerically. As $M \to \infty$ the variance $v_{i,t}$ goes to zero at the rate M^{-1} which is enough to leave the asymptotic properties of our estimators unaffected. The correction has a negligible effect in absence of jumps but is extremely useful in case of jump clustering.

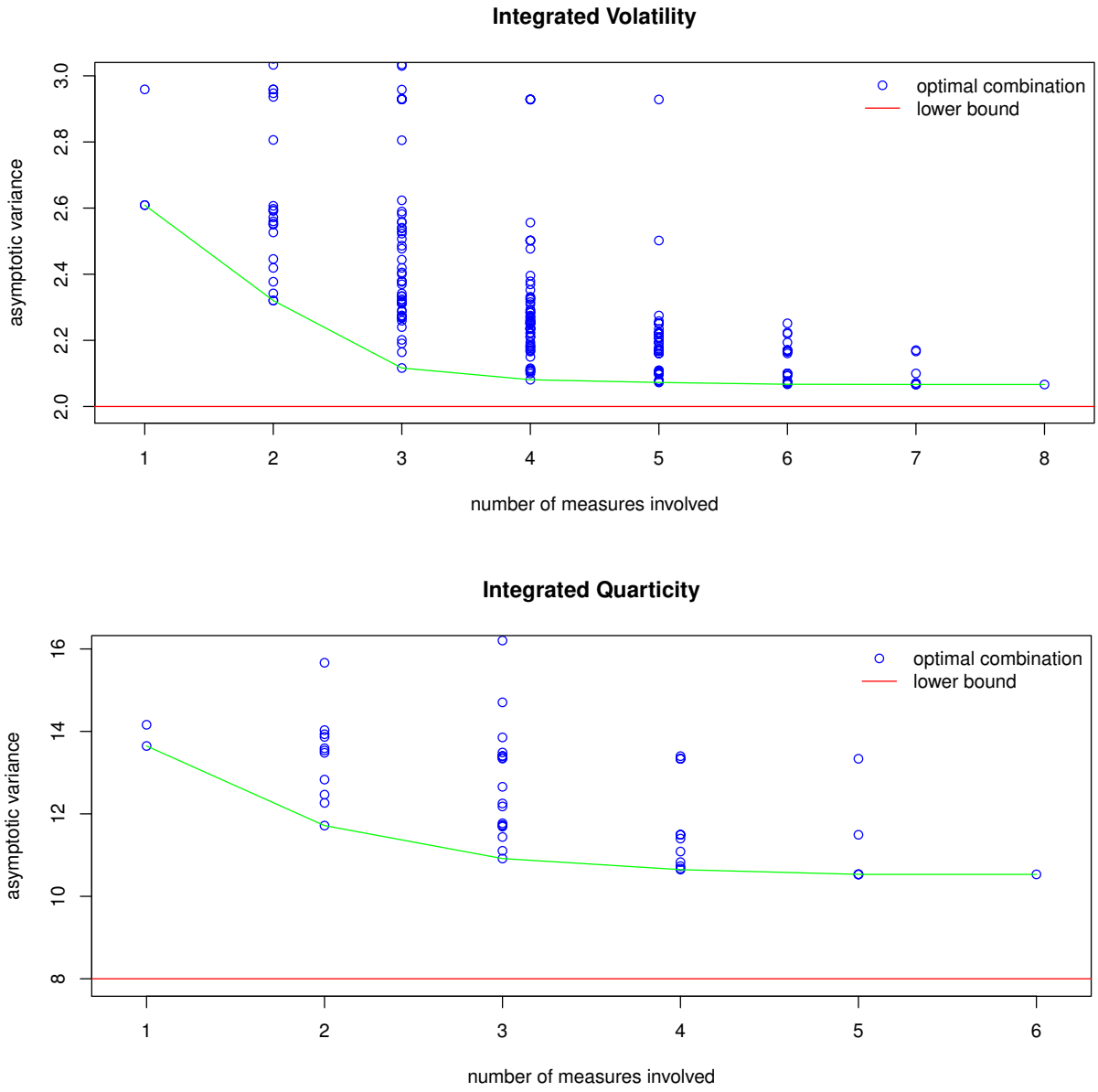


FIGURE2.1: Asymptotic variance of the optimal CRV (top panel) and CRQ (bottom panel) estimators obtained forming all the possible combinations between alternative measures.

n	$TBPV$	$TBPVs$	$TTPV$	$MinRV$	$MinRV$s	$MedRV$	$Min3RV$	$MinMedRV$	asym.
	w_1	w_2	w_3	w_4	w_5	w_6	w_7	w_8	var.
1	10	-	-	-	-	-	-	-	**2.609**
2	1.786	-	-	-0.786	-	-	-	-	**2.320**
3	1.076	0.593	-	-	-	-	-	-0.669	**2.116**
4	2.076	1.077	-1.672	-	-	-0.481	-	-	**2.081**
5	1.932	10	-1.359	-	-	-0.466	-0.107	-	**2.073**
6	2.212	1.295	-2.120	-0.688	-0.442	-	-	0.743	**2.067**
7	2.173	1.274	-2.015	0.193	-	-0.820	-0.394	0.591	**2.066**
8	2.173	1.274	-2.015	1.663	0.735	-2.245	-1.174	0.591	**2.066**

Table 2.1: Asymptotically efficient combinations of integrated volatility estimators (points connected by the green line in Figure 2.2).

	$TTPQ$	$MinRQ$	$MinRQ$s	$MedRQ$	$Min3RQ$	$MinMedRQ$	Asym.
n	w_1	w_2	w_3	w_4	w_5	w_6	var.
1	10	-	-	-	-	-	**13.65**
2	1.965	-	-	-	-	-0.965	**11.72**
3	1.495	-	-	0.378	-	-0.874	**10.92**
4	1.445	0.285	0.276	-	-	-16	**10.65**
5	1.578	0.327	0.292	-	0.248	-1.444	**10.53**
6	1.578	0.216	0.237	0.133	0.282	-1.444	**10.53**

Table 2.2: Asymptotically efficient combinations of integrated quarticity estimators (points connected by the green line in Figure 2.2).

2.4 Combining Multiple Measures

In this section we study the problem of constructing optimal combinations of robust estimators according to some criteria summarized in a specific loss function. Let $\mathbf{w}$ be a vector of weights with real unbounded components such that $\sum_{i=1}^{n} w_i = 1$ where n is the number of different measures involved (in our case $n = 8$ for the integrated volatility and $n = 6$ for the integrated quarticity).

$$\begin{aligned} CRV_t(\mathbf{w}) &= w_1 TBPV_t + w_2 TBPVs_t + w_3 TTPV_t + w_4 MinRV_t + \\ &\quad w_5 MinRV_t + w_6 Min3RV_t + w_7 Min3RV_t + w_8 MinMedRV_t \end{aligned} \tag{2.27}$$

$$\begin{aligned} CRQ_t(\mathbf{w}) &= w_1 TTPQ_t + w_2 MinRQ_t + w_3 MinRQs_t + \\ &\quad w_4 MedRQ_t + w_5 Min3RQ_t + w_6 MinMedRQ_t \end{aligned} \tag{2.28}$$

Proposition 5. As $M \to \infty$, under the same assumptions of Proposition 3 we have:

$$\sqrt{M}\left(CRV_{M,t} - IV_t\right) \xrightarrow{stable\,D} MN\left(0, \mathbf{w}^T S_{IV} \mathbf{w}\, IQ_t\right) \tag{2.29}$$

Proof The results follow directly from Proposition 3.

Analogously it can be proved that

$$\sqrt{M}\left(CRQ_{M,t} - IQ_t\right) \xrightarrow{stable\,D} MN\left(0, \mathbf{w}^T S_{IQ} \mathbf{w} \int_{t-1}^{t} \sigma_s^8 ds\right) \tag{2.30}$$

A natural criterion to determine the optimal weights is the minimization of the asymptotic variance under some linear constraints $A\mathbf{w} = \mathbf{b}$ where A is $k \times n$ and b is $k \times 1$. The minimum requirement is $\sum_i w_i = 1$ for consistency, but additional constraints can be useful in different circumstances: for instance one may exclude a subset of measures in order to save computational time or may suppress a specific finite sample bias estimated from a numerical simulation. The optimal estimator under the selected criterion is:

$$\mathbf{w}^* = S^{-1} A^T \left(A\, S^{-1}\, A^T\right)^{-1} \mathbf{b} \tag{2.31}$$

To highlight the gain in terms of asymptotic efficiency that comes from combining multiple measure, the optimal weights have been computed for all possible combinations of realized volatility and realized quarticity estimators that we are considering. The asymptotic variance scaled by M is plotted in Figure 2.2 while the optimal weights are reported on Tables 2.4 and 2.3. Note that the best combination of two realized volatility estimators, which involves the threshold bipower and the $MinRV$, achieves an asymptotic variance that is remarkably smaller than the bipower alone. The improvements become less pronounced as the number of estimators involved increases and we get closer to the maximum likelihood limit. Note that such a limit is attained asymptotically by the threshold estimator of Mancini (2009) but as already mentioned, in practice this estimator is extremely sensitive to the threshold level and to the error affecting the estimates of the local volatility. The CRV instead depends weakly on the threshold through $TBPV$ and the $TTPV$ (see CPR for further details). For the realized quarticity we also observe significant improvements as the number of measures increases, although the optimal estimator obtained with 6 measures has a variance that is still well above the lower bound attained by the efficient multipowers of KR.

The variance of CRQ could be further decreased including estimators that involve four or more consecutive returns but this would compromise the locality incrementing the bias in presence of stochastic volatility. Importantly, both the CRV and the CRQ type of estimators can be unbounded and are allowed to assume also negative values in finite samples. Anyway we will show that this happens extremely rarely in practice and moreover is not problematic for the construction of efficient jump tests which is the main focus of this work.

The minimization of the asymptotic variance is just one possible criterion. In empirical applications it can be also useful to take into account also some finite characteristics. The weights in that case can depend also on the specific features of the data generating process (the variability of the instantaneous volatility, the jump intensity and the jump size distribution, the noise to signal ratio) but the asymptotic consistency of the estimators is always guaranteed. A concrete example will be provided later.

Threshold adjustment for quantile based estimators

When combining multiple measures with possibly negative coefficients, we face the problem of obtaining negative estimates for quantities, like the realized volatility, that are non-negative by definition. Theoretically, the lower is the variance of our measures, the smaller is the probability to obtain negative values. In practice we must also consider some possible distortions emerging in finite samples. A crucial problem is the occurrence of multiple jumps within the same block of consecutive returns. Such events, referred as "gradual jumps" in the terminology of Barndorff-Nielsen et al. (2009), are non negligible in real data as pointed out also by CPR. Christensen et al. (2014) and Christensen et al. (2016) find that their are substantially different from jumps because they are generated by a continuous path dominated by a strong drift component. These events are usually indistinguishable from price jumps unless we use tick-by-tick data. For the purposes of our analysis we need to control the large bias generated in case of gradual jumps. While for the truncated and corrected versions of the multipower variation (equations 2.5 and 2.6) this is guaranteed by the presence of a threshold, the measures based on nearest neighbor truncation can be largely affected by the presence of consecutive abnormal returns. To mitigate this problem we

set a correction with no effects on the asymptotic results. Let us generally write our measure as

$$F_k(\mathbf{r}) = (M-k)^{-1} \sum_{i=k}^{M} \Theta\left(|r_{i-k,t}|, \ldots, |r_{i,t}|\right)$$

where $\mathbf{r} = (r_{1,t}, \ldots, r_{M,t})^T$. We substitute this expression with

$$\begin{aligned} F_k(\mathbf{r}) \quad = \quad & \frac{1}{M-k} \sum_{i=k}^{M} \Big[\Theta\left(|r_{i-k,t}|, \ldots, |r_{i,t}|\right) \mathbf{1}_{\left\{\Theta\left(|r_{i-k,t}|,\ldots,|r_{i,t}|\right) \leqslant \tau(\Theta) \max\left(v_{i-k,t},\ldots,v_{i,t}\right)\right\}} + \\ & \max\left(v_{i-k,t}, \ldots, v_{i,t}\right) \mathbf{1}_{\left\{\Theta\left(|r_{i-k,t}|,\ldots,|r_{i,t}|\right) \leqslant \tau(\Theta) \max\left(v_{i-k,t},\ldots,v_{i,t}\right)\right\}} \Big] \end{aligned} \tag{2.32}$$

where $\tau(\Theta)$ is a threshold that goes to zero slower than the modulus of continuity of the Brownian motion as $M \to \infty$. In practice we set M at the 99.99% of the distribution of $\Theta\left(|U_1|, \ldots, |U_k|\right)$ where $U_1, \ldots, U_k$ are i.i.d. standard normals and it is determined numerically. As $M \to \infty$ the variance $v_{i,t}$ goes to zero at the rate M^{-1} which is enough to leave the asymptotic properties of our estimators unaffected. The correction has a negligible effect in absence of jumps but is extremely useful in case of jump clustering.

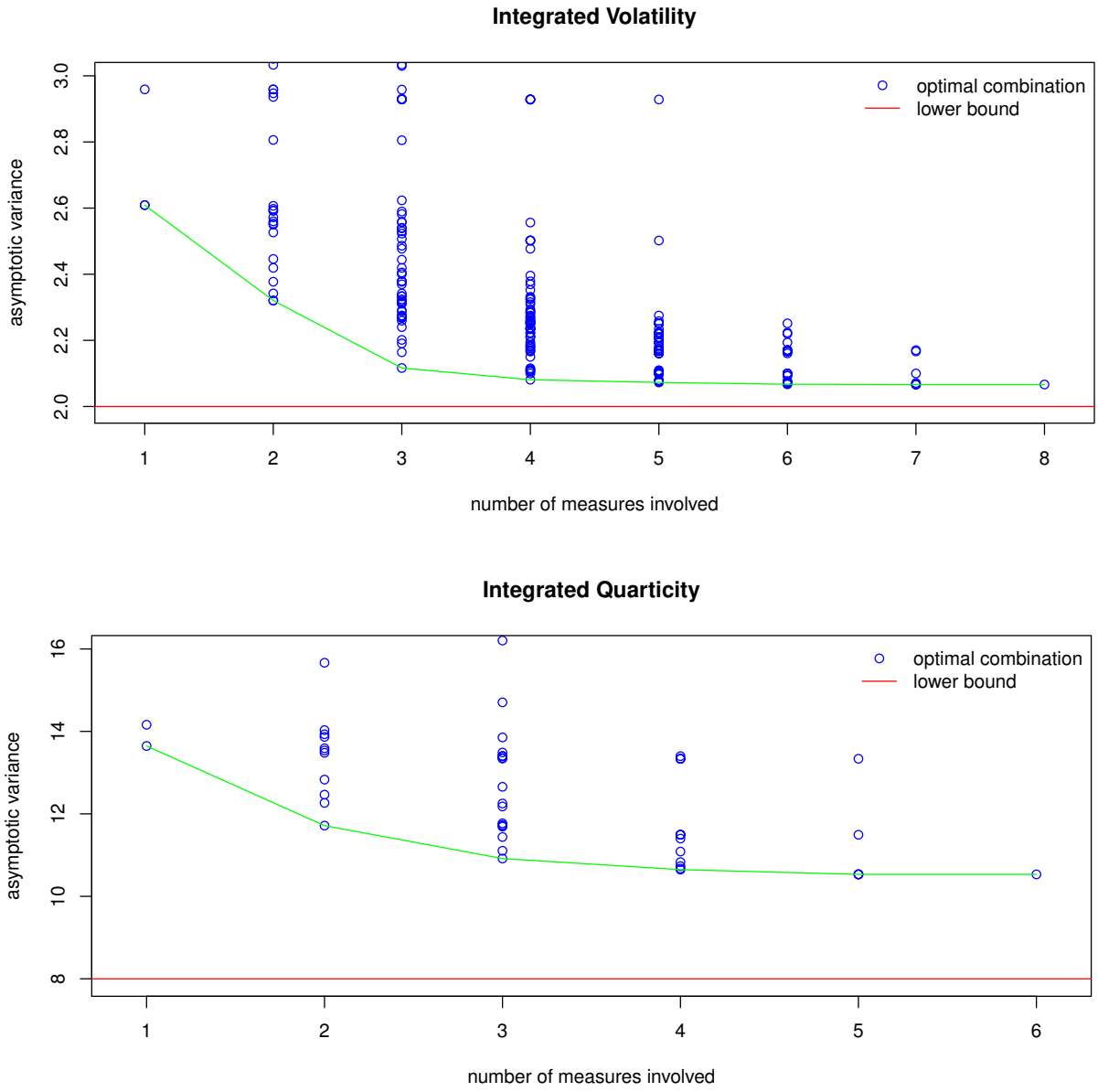


FIGURE2.2: Asymptotic variance of the optimal CRV (top panel) and CRQ (bottom panel) estimators obtained forming all the possible combinations between alternative measures.

n	$TBPV$ w_1	$TBPVs$ w_2	$TTPV$ w_3	$MinRV$ w_4	$MinRV$s w_5	$MedRV$ w_6	$Min3RV$ w_7	$MinMedRV$ w_8	asym. var.
1	10	-	-	-	-	-	-	-	**2.609**
2	1.786	-	-	-0.786	-	-	-	-	**2.320**
3	1.076	0.593	-	-	-	-	-	-0.669	**2.116**
4	2.076	1.077	-1.672	-	-	-0.481	-	-	**2.081**
5	1.932	10	-1.359	-	-	-0.466	-0.107	-	**2.073**
6	2.212	1.295	-2.120	-0.688	-0.442	-	-	0.743	**2.067**
7	2.173	1.274	-2.015	0.193	-	-0.820	-0.394	0.591	**2.066**
8	2.173	1.274	-2.015	1.663	0.735	-2.245	-1.174	0.591	**2.066**

Table 2.3: Asymptotically efficient combinations of integrated volatility estimators (points connected by the green line in Figure 2.2).

n	$TTPQ$ w_1	$MinRQ$ w_2	$MinRQ$s w_3	$MedRQ$ w_4	$Min3RQ$ w_5	$MinMedRQ$ w_6	Asym. var.
1	10	-	-	-	-	-	**13.65**
2	1.965	-	-	-	-	-0.965	**11.72**
3	1.495	-	-	0.378	-	-0.874	**10.92**
4	1.445	0.285	0.276	-	-	-16	**10.65**
5	1.578	0.327	0.292	-	0.248	-1.444	**10.53**
6	1.578	0.216	0.237	0.133	0.282	-1.444	**10.53**

Table 2.4: Asymptotically efficient combinations of integrated quarticity estimators (points connected by the green line in Figure 2.2).

2.5 Testing For Jumps

The tests that we consider in this section are based on the idea of Barndorff-Nielsen and Shephard (2006): two measures of the integrated variance are compared, one is the realize variance that includes the contribution of jumps (if present) while the second is robust to jump and isolates the contribution of the continuous price changes. Among the variety of test statistics proposed by Barndorff-Nielsen and Shephard (2006), Huang and Tauchen (2005) find that the ratio statistics, constructed on the relative contribution of jumps to the total intraday price variability $\hat{RJ}$ defined in expression 1.7, is favored in terms of finite sample properties. The main reason for privileging this statistics is that in absence of jumps and under the assumption of constant volatility, the finite sample bias is of order $M^{-3/2}$. Let ϵ_{RV} and $\epsilon_{I\hat{V}}$ denote the relative measurement errors in absence

of jumps and assuming that $\hat{IV}$ is unbiased:

$$\epsilon_{RV,t} = \frac{RV_t}{IV_t} - 1 \qquad \epsilon_{\hat{IV},t} = \frac{\hat{IV}_t}{IV_t} - 1 \tag{2.33}$$

we have

$$\mathbb{E}\left[\hat{RJ}_t\right] = \mathbb{E}\left[\frac{\epsilon_{RV,t} - \epsilon_{\hat{IV},t}}{1 + \epsilon_{RV,t}}\right] = \mathbb{E}\left[\left(\epsilon_{RV,t} - \epsilon_{\hat{IV},t}\right) \sum_{k=0}^{\infty} \left(-\epsilon_{RV,t}\right)^k\right] \tag{2.34}$$

note that the first term on the RHS is zero because $\mathbb{E}\left[\epsilon_{RV}^2\right] = \mathbb{E}\left[\epsilon_{RV}\epsilon_{\hat{IV},t}\right] = \frac{2}{M}\frac{IQ_t}{IV_t^2}$, the bias is therefore of order $O\left(M^{-3/2}\right)$ as it is well known.

To obtain a pivotal quantity Z_t under the null hypothesis of pure diffusion, $\hat{RJ}_t$ is scaled according to equation 1.8. Under the null of no jumps Z_t converges asymptotically to a standard normal random variable. A jump is detected with the confidence level $1-\alpha$ when $Z_t > \Phi_{1-\alpha}^{-1}$ being $\Phi_{1-\alpha}^{-1}$ the inverse standard normal evaluated at $1-\alpha$. We recall that the constants $v_{\hat{IV}}$ and v_{RV} appearing in 1.8 are respectively the asymptotic variance of $\hat{IV}$ and RV multiplied by M, therefore $v_{RV} = 2$ and $v_{\hat{IV}} > 2$. The crucial role played by the asymptotic variance of $\hat{IV}_t$ is emphasized in the definition of Z_t making clear the advantage of using measures with a small value of $v_{\hat{IV}}$. Small jumps, that generate modest relative differences between RV and IV, can be distinguished from the noise only if the integrated variance is measured with a high precision. Also the estimates of the integrated quarticity play a crucial role in determining the size and the power of the test as pointed out by KR. This motivates a particular interest on the also on the possible advantages coming from the use of combined quarticity measures. While the gain in efficiency determined by a reduction of the asymptotic variance of $\hat{RJ}$ is quite evident, the role of the quarticity measure is more difficult to identify and the best of our knowledge not deeply analyzed by the literature. Let us again assume that there are no jumps and let $\epsilon_{\hat{IQ}}$ denote the relative error affecting our quarticity estimates which are also characterized by a relative bias $\beta_{\hat{IQ}}$

$$\epsilon_{\hat{IQ}} = \frac{\hat{IQ}_t}{IQ_t} - 1 - \beta_{\hat{IQ}} \tag{2.35}$$

The following result holds

Proposition 6. Assuming no jumps and that $\hat{IV}_t$ is an unbiased and consistent estimator of IV_t while $\hat{IQ}_t$ is a possibly biased consistent estimator for IQ_t. In this case Z_t has a bias of order $M^{-1/2}$ given by

$$\mathbb{E}\left[Z_t\right] = -IV_t\sqrt{\frac{M}{\left(v_{\hat{IV}}-2\right)IQ_t}}\left\{\frac{v_{\hat{IV}}-2}{M}\frac{IQ}{IV^2}+\frac{1}{2}\mathbb{E}\left[\left(\epsilon_{RV,t}-\epsilon_{\hat{IV},t}\right)\epsilon_{\hat{IQ},t}\right]\right\} \tag{2.36}$$

Proof Substituting expressions 2.33 and 2.35 into definition 1.8 we obtain

$$\begin{aligned}
\mathbb{E}\left[Z_t\right] &= IV_t\sqrt{\frac{M}{\left(v_{\hat{IV}}-2\right)IQ_t}}\,\mathbb{E}\left[\frac{\left(\epsilon_{RV,t}-\epsilon_{\hat{IV},t}\right)\left(1+\epsilon_{\hat{IV},t}\right)}{\left(1+\epsilon_{RV,t}\right)\left(1+\beta_{\hat{IQ}}+\epsilon_{\hat{IQ},t}\right)^{1/2}}\right] \\
&= IV_t\sqrt{\frac{M}{\left(v_{\hat{IV}}-2\right)IQ_t}}\,\mathbb{E}\left[\left(\epsilon_{RV,t}-\epsilon_{\hat{IV},t}\right)\left(1+\epsilon_{\hat{IV},t}-\epsilon_{RV,t}-\frac{1}{2}\epsilon_{\hat{IQ},t}-\frac{1}{2}\beta_{\hat{IQ}}\right)\right]+O\left(M^{-1}\right) \\
&= -IV_t\sqrt{\frac{M}{\left(v_{\hat{IV}}-2\right)IQ_t}}\left\{\frac{v_{\hat{IV}}-2}{M}\frac{IQ}{IV^2}+\frac{1}{2}\mathbb{E}\left[\left(\epsilon_{RV,t}-\epsilon_{\hat{IV},t}\right)\epsilon_{\hat{IQ},t}\right]\right\}+O\left(M^{-1}\right)
\end{aligned}$$

Note that the bias on Z_t is not only determined by the variance of $\hat{IV}_t$ in excess to the Kramer's Rao bound but depends also on the correlation between $\epsilon_{\hat{IQ},t}$ and $\epsilon_{RV,t}-\epsilon_{\hat{IV},t}$. When such a correlation is negative the it provides a positive contribution to the bias which may eventually increase the spurious detection rate. For this reason less accurate measures of the integrated quarticity can be sometime lead to a superior size-adjusted power as we notice in our numerical simulation.

Finite sample correction for the covariance matrix

Note that the covariance matrices 2.12 and 2.20 in finite samples are biesed also in the ideal setting characterized by constant volatility and i.i.d. normal returns. The bias in this case can be calculated from expressions B.1, B.2 and B.3:

$$S_{IV}^{M} = S_{IV}+\frac{1}{M}\Delta S_{IV}+O\left(M^{-2}\right) \tag{2.37}$$

where

$$\Delta S_{IV} = \begin{bmatrix} 1.47 & 1.14 & 1.93 & 1.75 & 1.22 & 1.76 & 2.38 & 2.19 \\ & 2.93 & 2.74 & 1.22 & 3.50 & 2.53 & 3.29 & 3.06 \\ & & 3.63 & 2.26 & 3.14 & 2.98 & 4.74 & 4.23 \\ & & & 2.43 & 1.38 & 2.20 & 3.16 & 2.80 \\ & & & & 4.87 & 3.05 & 4.20 & 3.77 \\ & & & & & 3.60 & 3.31 & 3.41 \\ & & & & & & 7.83 & 6.14 \\ & & & & & & & 5.26 \end{bmatrix}$$

The corrections is generally negligible for the computation of the common tests statistics: for the test of BNS for instance it is smaller then 1.5% at ten minutes. As the asymptotic variance of RJ decreases, the finite sample distortion becomes more relevant in relative terms: for the optimal combination of all measures reported on Table 2.3 the corrected variance is 2.26 at ten minutes, implying that the variance of RJ is almost four times larger than its asymptotic limit. Of course other finite sample effects related to stochastic volatility, like market microstructure noise and jumps and stale quotes, cannot be calculated in closed form but may also have relevant effects on the test statistics. Such effects will be studied through an extensive Monte Carlo simulation in the next section.

2.6 Simulation Study

A Monte Carlo study has been conducted to investigate the performances of the new tests when realistic features like stochastic volatility and market microstructure noise are introduced in the the data generating process. These elements can affect the performances of IV estimators in finite samples and consequently the outcomes of the tests. The following set of models is considered:

Model 1: Driftless Brownian motion (BM). Is the baseline model with i.i.d. Gaussian homoskedastic log-returns that allows to isolate finite sample effects without altering any other feature of the data generating process with respect to the "ideal" asymptotic setting

Model 2: Stochastic volatility and intraday U-shaped volatility pattern

Model 3: Stochastic volatility, intraday U-shaped volatility pattern and market microstructure noise

Model 4: Stochastic volatility plus rare big jumps

Model 5: Stochastic volatility plus small self-exciting jumps

Model 6: Stochastic volatility, microstructure noise and rare big jumps

Model 7: Stochastic volatility , microstructure noise and frequent self-exciting small jumps

2.6.1 Simulation Design

For each model 20.500 trading days have been simulated using the first 500 as a burn-in period. Each day has a duration of 6 hours and 30 minutes, consistently with usual operating time of the NYSE. In all models except model 1, stochastic volatility is introduced through the log-volatility specification which consists in the following dynamics of log-prices:

$$dY(t) = \mu\, dt + \sigma_U(t)\, \sqrt{v(t)}\, dW_1(t) \tag{2.38}$$

$$d\log v(t) = (\beta_0 - \beta_1 v(t))\, dt + \eta\, dW_2(t) \tag{2.39}$$

where W_1 and W_2 are correlated Brownian motions with constant correlation ρ to generate the well known leverage effect. This setting is common to many other simulation experiments including Huang and Tauchen (2005), CPR, Dumitru and Urga (2011), Gilder et al. (2014) and KS. We set the parameter $\rho = -0.61$ according to the estimates of Andersen et al. (2002) and $\mu = 3\%$, while for the evolution of the stochastic volatility we prefer to calibrate the model on our S&P data described in Section 1.2 that will be used also in the empirical application of Section 2.7. To this purpose we run the simple OLS regression

$$\log IV_t = \beta_0 + (1 - \beta_1) \log IV_{t-1} + \eta\epsilon_t$$

where IV is the continuous realized volatility component that is constructed subtracting jumps identified using the sequential version of the test proposed by CPR (see Section 1.3 for details) and we assume $\epsilon_t \sim N(0,1)$. From the regression coefficients we find $\beta_0 = -0.0829$, $\beta_1 = 0.128$,$\eta = 0.55$, where time is measured in days and log-returns in percentage. The volatility is unconditionally distributed as a log-normal with location and scale parameters respectively equal to $\beta_0/\beta_1 = -0.648$ and $\eta^2/4\beta_1 = 0.591$, the average daily volatility is $\exp\left(\beta_0/\beta_1 + \eta^2/4\beta_1\right) = 0.95\%$ corresponding to an annualized value of 15.52% (assuming 252 trading days per year). Note that the mean reversion is very strong being characterized by an half-life time of 7.81 trading days. The model is clearly

unable to capture the long term volatility persistence, probably due to the strong volatility shocks characterizing the sub-prime crisis. Anyway it captures quite well the the unconditional mean and variance of the continuous realized volatility, which is probably more relevant for the purposes of our study.

The intraday pattern is introduced following Andersen et al. (2012) according to the functional form of Hasbrouck (1999):

$$\sigma_U(t) = C + A\,e^{-at} + B\,e^{-b(1-t)} \quad t \in [0,1] \tag{2.40}$$

parameters are $A = 0.75$, $B = 0.25$, $B = 0.89$, $a = 10$, $b = 10$. The U-shape associated with this values is very pronounced: the volatility at the opening and at the closing time on each day are respectively 3 and 1.5 times larger than the mid-day volatility.

In models 4, 5, 6 and 7 the log-price process is augmented with i.i.d. normally distributed jumps having zero mean and variance σ_J^2:

$$p_t = Y_t + \int_0^t J_s dN_s$$

In models 4 and 6 N_t is an independent counting process having constant intensity $\lambda = 1/8$ and $\sigma_J = 1.5$. In models 5 and 7 we introduce jump clustering effects comparable to those documented in Chapter 1 and featured by a self-exciting Hawkes process:

$$d\lambda_t = \gamma\,(\lambda_0 - \lambda_t)\,dt + \kappa dN_t$$

where the parameters values are $\lambda_0 = 0.25$, $\gamma = 8$ and $\kappa = 2.5$ while the jump size is characterized by $\sigma_J = 0.75$. The self-excitation effect is very strong, generating an upward shock to the jump intensity that is ten times larger than its baseline level λ_0 whenever a jump occurs. The jump intensity is quickly reverting to λ_0 with an half lifetime of about 55 minutes. This jumps dynamics is consistent with the empirical findings of the first chapter and can have important effects on realized volatility measures given that they are constructed assuming that at most a single jump can occur in a block of consecutive returns sampled at high frequency. Under this specification, the probability to observe jumps occurring very close in time is much higher compared to a Poisson process. While the threshold bipower is designed to be robust in presence of consecutive jumps, the measured based on nearest neighbor truncation can be strongly upward biased. The correction

introduced in equation 2.32 helps to reduce this bias.

The last ingredient introduced in the simulation is the market microstructure noise, which is present in models 5 and 7 where we assume that prices are observed with an error:

$$\tilde{p}_t = p_t + \epsilon_t \tag{2.41}$$

following Gilder et al. (2014), ϵ_t is normal with zero mean and variance equal to 10^{-3} times the daily realized variance, to obtain a noise to signal ratio consistent with the empirical findings of Hansen and Lunde (2006). Equations 2.38 and 2.39 and the jump processes are simulated according to the Euler scheme with a time increment of one second. The simulated data for all models are then sampled at coarser frequencies of 1 min, 5 min and 10 min following the recommendation of Huang and Tauchen (2005). Each trading day contains respectively 390, 78 and 34 intraday log-returns. Model 1 coincides with the Brownian motion W_1 driving the log-price in equation 2.38.

2.6.2 Simulation Results

For each different simulated model we compute all the measures of Section 2.2.1. To form optimal combinations we must first of all define a criterion that should clearly depend on the purpose for which the new measures are defined. Though the main objective of this work is the construction of new non-parametric jump tests, it is also interesting to explore the possible gain of precision in measuring volatility and quarticity that can be obtained in finite samples combining different measures. To set a trade-off between asymptotic efficiency and finite sample biases we follow KR minimizing the relative squared error averaged over all simulated trading days and across the different simulated scenarios at the reference sampling frequency of 5 minutes[5]. The results reported on Tables 2.5 and 2.6 show that the newly defined measures (denoted as CRV^* and CRQ^*) have a finite sample bias comparable to the others but the average relative squared error is clearly smaller. Note that the weights defining CRV^* and CRQ^* (Table 2.7) differ from the asymptotically optimal weights (Tables 2.3 and 2.4), confirming that the finite sample effects are non negligible at 5 minutes. Importantly, though the CRV^* and $CRQ^*$are in principle unbounded, the zero lower bound is never violated in practice.

[5] This sampling frequency is widely considered as a good compromise between the desire to sample data as finely as possible and the need to avoid an excessive contamination from market microstructure noise.

The improved accuracy of the combined measures is expected to provide some advantages also for the construction of a new jump test. Nevertheless the minimizing the variance of the measurement errors is not necessarily the best choice for the construction of a test statistics with suitable properties. As already mentioned, there is no generally accepted criterion to compare the performances of different jump tests. Interestingly, DU build a comparative analysis based on the adjusted power, i.e. the power of the tests calculated after adjusting the confidence levels to match the nominal size in absence of jumps. Although such an adjustment guarantees a fair comparison, it is based exclusively on the assessment of the power neglecting other relevant finite sample characteristics. An important feature of the ratio statistics of BNS based on the bipower variation is that the distribution in finite samples is quite close to a normal. Pronounced deviations from the asymptotic normality indicate important distortions the finite samples. As a consequence, the actual size of the test may differ significantly from the nominal one and requires some specific adjustments quantified usually through numerical methods. In empirical applications this aspect is particularly relevant since the data generating process is unknown. If instead the distribution of the test statistics is close to the normal in finite samples and under realistic features of the data generating process, it is much easier to control the occurrence of type 1 errors.

To investigate the advantages of combining multiple measures, we construct a large set of alternative statistics obtained forming all the asymptotically efficient combinations of the eight available measures for the stochastic volatility and the six estimators of the realized quarticity. For each of the 16,065 combinations the weights are determined minimizing the asymptotic covariance matrix. Again we perform the analysis at the sampling frequency of 5 minutes. For each test we compute the size adjusted power in the different simulated scenarios as suggested by DU. The adjustment to the confidence level is designed to match the effective size of the test in a context with no jumps: the adjustment for models 4 and 5 is therefore calibrated on model 2 while the adjustment for models 6 and 7 is calibrated on model 3. This procedure allows to take into account the different effects of stochastic volatility, a marked U-shaped intraday volatility pattern and market microstructure noise on the effective size of the tests. To better control for the role of stochastic volatility, we consider also the proposal of Rognlie (2010) to scale log-returns by an appropriate factor in order to eliminate the intraday periodicity of volatility. The correction is based on the Truncated Maximum

Likelihood estimator of Boudt et al. (2011) described in Appendix A.1. The first panel in Figure 2.3 shows the dependence between the power of the test and the precision of the IV estimator calculated as the inverse of the asymptotic variance. As expected, the efficiency of the IV measure is a fundamental driver for the power of the corresponding test. Importantly, the test based on CRV^* and CRQ^* is not the most powerful in absolute terms.

The second dimension that we explore is the distance from normality measured using the loss function of the efficient GMM based on the uncentered moments up to the fourth order[6]. The second panel of Figure 2.3 reports the location of the different tests along the selected dimensions: efficiency in jump detection, ranking in terms of average distance from the normal. We note that tests with more power are also deviating significantly form, suggesting the need to find a balance between efficiency and robustness in finite samples. Moving from right to left the finite sample properties improve without affecting the power. The upper bound of the populated region can be therefore regarded as a sort of "efficient frontier". Importantly, the red points representing commonly used statistics based on the ratios of known measures, are located below this frontier. We propose therefore some alternative combinations which provide significant improvements either in terms of power or normality or both. Starting from the statistics with the maximum power (green point, denoted also as MP henceforth), we note that the finite sample properties can be substantially improved without compromising the power if we move to the blue point labeled as "High Power" (HP henceforth). The test based on efficient measures (yellow point) can be conveniently replaced by the magenta point labeled as "Balanced", which has a slightly lower power in absolute terms but exhibits a suitable balance between power and normality. The weights defining the new tests are reported on Table 2.7. Importantly, among the statistics based on a single measure for IV and IQ, the traditional association between bipower variation and tripower quarticity is not the optimal choice. The red points represent common tests based on known measures and importantly none of them lies on the efficient frontier. The best combination based on a single measure for the IV and a single measure for the IQ is constructed using the corrected threshold bipower and the staggered minimum realized quarticity (orange point in the figure). This can be understood looking at the bias that affects the tests statistics (expression 2.36) which depends on the covariance between $\hat{IQ}_t$

[6] Note that the distance from the normal is calculated only in absence of jumps.

and $RV_t - \hat{IV}_t$. A large negative covariance between these quantities increases the probability of type 1 errors and requires more severe adjustments to the confidence level. Therefore it is not surprising that the tests with the highest power can be constructed also using quarticity estimators that are not extremely accurate.

Figures 2.4, 2.5 and 2.6 show the quantile-quantile plot for the ratio test statistics based on the different estimators and under different features of the data generating process. The test based on the corrected threshold bipower and the staggered minimum realized variance is generally closer to the normal than the others in almost all simulated scenarios. All tests are strongly affected by the presence of market microstructure noise at 1 minute with the only exception of the HP test that exhibits much smaller effects. This suggests the possibility to design appropriate combinations of measures to improve the robustness to market microstrucure noise at high frequency. The distributional properties of the Balanced statistics in finite samples are comparable to those of the CPR test while the HP test clearly exhibits fatter tails with respect to the normal, even when the data generating process is a geometric Brownian motion.

A complete comparative analysis for the different tests is presented on tables 2.8 and 2.9 where the size adjusted power is reported under each simulated process. Two benchmarks have been considered: the test of Lee and Mykland (2008) and its modified version presented in Section 1.3.1 (denoted as m-LM henceforth). The latter is based on a different estimator for the local volatility designed to reduce the bias in presence jumps and to be more responsive in presence of fast volatility changes. To highlight the role of time-changing volatility, the results are reported with and without correcting for the intraday periodicity of volatility. In agreement with Rognlie (2010) and Boudt et al. (2011) our, taking into account the intraday volatility pattern is extremely beneficial to reduce the spurious detection rates and therefore difference between the actual and the nominal size of the tests. In the end, this determines a higher level of the size-adjusted power.

Importantly, under all simulated processes, the m-LM test has always the highest power that is clearly superior also to the standard LM tests. It is anyway important to to remark that the detection procedure of Lee and Mykland and its variants require severe restrictions to the stochastic volatility process for consistency which rule out volatility jumps. In light of the growing empirical evidence in this regard, we believe that this a relevant limitation to the empirical use of these

detection methods. Remarkably, the HP tests achieves a level of efficiency that is on average higher then the standard LM test (especially at 5 and 10 minutes) remaining asymptotically consistent also in presence of a càdlàg volatility dynamics. The high level of power comes at the cost of a pronounced sensitivity to specific features of the data generating process in finite samples, requiring substantial corrections to the nominal significance level. The Balanced tests exhibits a jump detection ability that is comparable to the LM test but at the same time minor corrections to the significance level are required. Overall it seems to providing a good trade-off between robustness and efficiency. Interestingly, the modified version of the CPR test based on the $MinRQs$ estimator for the realized quarticity performs always better than the standard CPR test.

Figures 2.7, 2.8 and 2.9 show the dependence of the size adjusted power on the confidence level at different sampling frequencies. The m-LM test dominates in all circumstances, while the standard LM test has more power than the HP test only at the highest sampling frequency of 1 minute.

Frequency	Measure	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
1 min	$C-TBPV$	**-0.020**	0.791	**-0.019**	**-0.015**	0.795	0.801
	$C-TBPVs$	-0.022	**0.745**	-0.020	-0.017	0.749	0.755
	$C-TTPV$	-0.023	0.783	-0.021	-0.018	0.786	0.791
	$MinRV$	-0.022	0.820	-0.021	-0.019	0.822	0.825
	$MinRVs$	-0.024	0.743	-0.023	-0.021	**0.745**	0.748
	$MedRV$	-0.024	0.787	-0.023	-0.021	0.789	0.792
	$Min3RV$	-0.022	0.809	-0.021	-0.020	0.811	0.813
	$MinMedRV$	-0.024	0.798	-0.023	-0.022	0.799	0.802
	CRV*	-0.022	0.758	**-0.019**	**-0.015**	0.762	**0.769**
	$C-TTPQ$	-0.038	1.997	**-0.033**	**-0.025**	2.010	2.035
	$MinRQ$	-0.040	2.095	-0.038	-0.033	2.103	2.115
	$MinRQs$	-0.049	1.756	-0.047	-0.042	1.763	1.777
	$MedRQ$	-0.045	1.970	-0.042	-0.037	1.978	1.992
	$Min3RQ$	**-0.037**	2.080	-0.051	-0.047	2.015	2.026
	$MinMedRQ$	-0.048	29	-0.046	-0.042	2.015	2.027
	CRQ*	-0.043	**1.919**	-0.040	-0.033	**1.929**	**1.948**
5 min	$C-TBPV$	**-0.030**	0.126	**-0.021**	**-05**	0.136	0.154
	$C-TBPVs$	-0.041	0.112	-0.032	-0.017	0.122	0.139
	$C-TTPV$	-0.041	0.114	-0.033	-0.020	0.123	0.138
	$MinRV$	-0.032	0.126	-0.027	-0.020	0.131	0.140
	$MinRVs$	-0.043	**0.109**	-0.039	-0.031	**0.114**	**0.122**
	$MedRV$	-0.042	0.114	-0.037	-0.029	0.120	0.129
	$Min3RV$	-0.042	0.115	-0.038	-0.030	0.119	0.128
	$MinMedRV$	-0.043	0.114	-0.038	-0.031	0.118	0.127
	CRV*	-0.037	0.116	-0.027	-09	0.128	0.148
	$C-TTPQ$	-0.091	0.190	-0.071	**-0.031**	0.217	0.268
	$MinRQ$	**-0.072**	0.219	**-0.059**	-0.041	0.234	0.261
	$MinRQs$	-0.106	**0.165**	-0.095	-0.076	**0.181**	**0.204**
	$MedRQ$	-0.098	0.186	-0.085	-0.063	0.203	0.231
	$Min3RQ$	-0.092	0.194	-0.098	-0.083	0.188	0.208
	$MinMedRQ$	-0.101	0.183	-0.091	-0.071	0.196	0.221
	CRQ*	-0.101	0.178	-0.085	-0.055	0.200	0.237
10 min	$C-TBPV$	**-0.039**	0.039	**-0.022**	07	0.057	0.087
	$C-TBPVs$	-0.057	0.020	-0.040	-0.012	0.039	0.068
	$C-TTPV$	-0.058	0.020	-0.044	-0.020	0.035	0.060
	$MinRV$	-0.041	0.037	-0.032	-0.019	0.047	0.061
	$MinRVs$	-0.061	**0.015**	-0.052	-0.040	**0.025**	**0.038**
	$MedRV$	-0.059	0.019	-0.049	-0.035	0.030	0.045
	$Min3RV$	-0.058	0.019	-0.051	-0.031	0.027	0.047
	$MinMedRV$	-0.060	0.018	-0.052	-0.037	0.027	0.042
	CRV*	-0.052	0.026	-0.032	**01**	0.047	0.082
	$C-TTPQ$	-0.136	-04	-0.100	**-0.021**	0.037	0.126
	$MinRQ$	**-0.094**	0.039	-0.071	-0.034	0.064	0.109
	$MinRQs$	-0.156	-0.026	-0.133	-0.105	**-01**	**0.035**
	$MedRQ$	-0.141	-09	-0.115	-0.077	0.021	0.067
	$Min3RQ$	-0.131	**00**	-0.132	-0.104	03	0.034
	$MinMedRQ$	-0.142	-0.010	-0.126	-0.085	0.010	0.054
	CRQ*	-0.149	-0.018	-0.119	-0.065	0.017	0.080

Table 2.5: Relative bias of the different measures under the alternative simulated processes.

Frequency	Measure	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
1 min	$C-TBPV$	0.007	0.008	0.651	0.008	0.008	0.657	0.667
	$C-TBPVs$	0.007	0.008	**0.579**	0.008	0.008	**0.584**	**0.593**
	$C-TTPV$	0.008	0.009	0.642	0.009	0.009	0.647	0.655
	$MinRV$	0.010	0.011	0.710	0.011	0.011	0.712	0.717
	$MinRVs$	0.010	0.011	0.585	0.011	0.011	0.588	**0.593**
	$MedRV$	0.008	0.009	0.647	0.009	0.009	0.650	0.656
	$Min3RV$	0.015	0.018	0.709	0.017	0.017	0.712	0.717
	$MinMedRV$	0.011	0.013	0.676	0.013	0.013	0.678	0.683
	CRV^*	**0.006**	**0.007**	0.595	**0.007**	**0.007**	0.601	0.613
	$C-TTPQ$	0.035	0.067	4.483	0.067	0.069	4.543	4.660
	$MinRQ$	0.047	0.088	5.040	0.088	0.089	5.074	5.136
	$MinRQs$	0.048	0.083	**3.576**	0.083	0.083	**3.603**	**3.663**
	$MedRQ$	0.036	0.065	4.359	0.066	0.067	4.393	4.457
	$Min3RQ$	0.074	0.139	5.346	0.126	0.126	4.945	4.997
	$MinMedRQ$	0.051	0.090	4.676	0.090	0.091	4.704	4.760
	CRQ^*	**0.031**	**0.057**	4.088	**0.057**	**0.058**	4.131	4.216
5 min	$C-TBPV$	0.034	0.037	0.063	0.038	0.042	0.068	0.078
	$C-TBPVs$	0.034	0.036	0.058	0.037	0.040	0.062	0.071
	$C-TTPV$	0.040	0.042	0.067	0.043	0.045	0.070	0.078
	$MinRV$	0.049	0.053	0.085	0.054	0.055	0.087	0.091
	$MinRVs$	0.050	0.052	0.078	0.053	0.053	0.081	0.084
	$MedRV$	0.038	0.041	0.065	0.041	0.042	0.067	0.071
	$Min3RV$	0.078	0.080	0.119	0.081	0.087	0.121	0.130
	$MinMedRV$	0.055	0.057	0.087	0.058	0.059	0.089	0.093
	CRV^*	**0.029**	**0.031**	**0.052**	**0.032**	**0.037**	**0.057**	**0.069**
	$C-TTPQ$	0.178	0.261	0.415	0.283	0.403	0.463	0.607
	$MinRQ$	0.242	0.376	0.595	0.395	0.418	0.627	0.686
	$MinRQs$	0.249	0.326	0.502	0.341	0.359	0.536	0.571
	$MedRQ$	0.185	0.257	0.413	0.267	0.287	0.440	0.489
	$Min3RQ$	0.382	0.552	0.856	0.512	0.536	0.796	0.841
	$MinMedRQ$	0.260	0.353	0.554	0.363	0.384	0.579	0.621
	CRQ^*	**0.159**	**0.222**	**0.353**	**0.235**	**0.273**	**0.387**	**0.452**
10 min	$C-TBPV$	0.069	0.073	0.084	0.078	0.094	0.092	0.115
	$C-TBPVs$	0.070	0.071	0.078	0.074	0.089	0.086	0.108
	$C-TTPV$	0.082	0.083	0.093	0.086	0.098	0.098	0.116
	$MinRV$	0.102	0.106	0.122	0.109	0.114	0.126	0.135
	$MinRVs$	0.102	0.101	0.114	0.103	0.106	0.118	0.123
	$MedRV$	0.080	0.080	0.089	0.082	0.085	0.092	**0.100**
	$Min3RV$	0.160	0.157	0.181	0.159	0.281	0.184	0.309
	$MinMedRV$	0.114	0.113	0.128	0.114	0.120	0.131	0.140
	CRV^*	**0.059**	**0.061**	**0.067**	**0.065**	**0.086**	**0.076**	0.106
	$C-TTPQ$	0.372	0.458	0.550	0.527	0.902	0.635	1.106
	$MinRQ$	0.512	0.710	0.838	0.774	0.881	0.915	1.087
	$MinRQs$	0.502	0.533	0.655	0.588	0.620	0.723	0.803
	$MedRQ$	0.387	0.437	0.527	0.487	0.565	0.598	0.760
	$Min3RQ$	0.812	0.984	1.207	0.889	0.991	1.166	1.258
	$MinMedRQ$	0.550	0.638	0.774	0.662	0.798	0.810	0.964
	CRQ^*	**0.330**	**0.376**	**0.454**	**0.422**	**0.544**	**0.520**	**0.708**

Table 2.6: Variance of the relative measurement error under the alternative simulated processes.

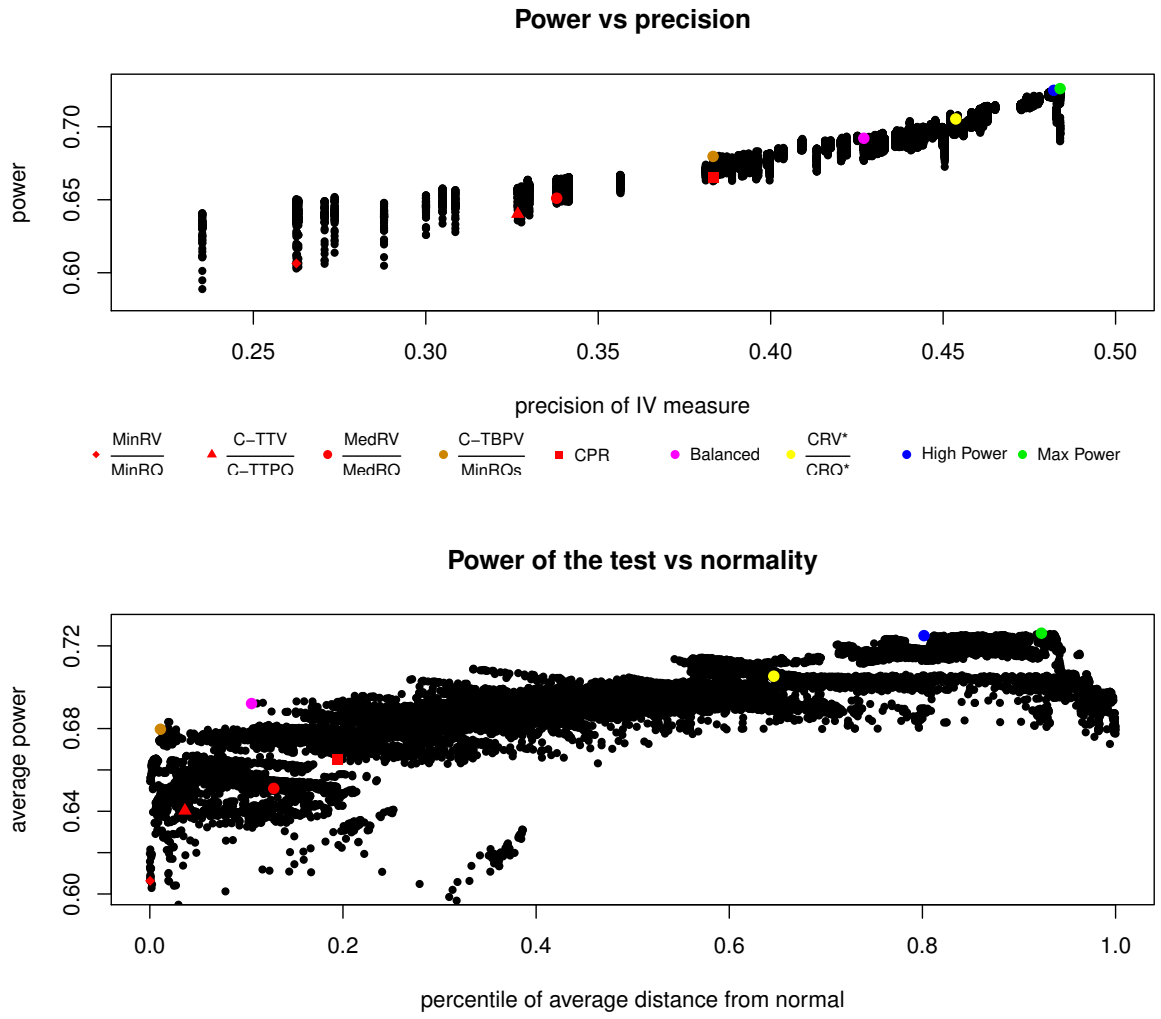


FIGURE2.3: In the top panel the size adjusted power of the test is plotted against the precision of the integrated volatility measure. The bottom panel shows the position of the different test in terms of power and average distance from the asymptotic normal distribution (measured using the efficient GMM loss function calculated on the first four uncentered moments).

	Maximum Power Combination	High Power Combination	$\frac{CRV*}{CRQ*}$	Balanced Combination	$\frac{C-TBPV}{MinRQs}$
$C-TBPV$	2.173	1.830	0.512	1.597	1.000
$C-TBPVs$	1.274	1.224	0.869	0.000	0.000
$C-TTPV$	-2.015	-1.503	-0.139	0.000	0.000
$MinRV$	-0.550	0.000	-0.337	0.000	0.000
$MinRVs$	-0.371	-0.169	-0.359	0.000	0.000
$MedRV$	-0.100	-0.383	0.493	0.000	0.000
$Min3RV$	0.000	0.000	0.055	0.000	0.000
$MinMedRV$	0.591	0.000	-0.095	-0.597	0.000
Asymptotic Variance	2.066	2.074	2.204	2.342	2.609
$C-TTPQ$	0.000	0.000	0.418	0.000	0.000
$MinRQ$	0.000	0.000	-0.170	0.000	0.000
$MinRQs$	0.000	1.000	0.180	1.000	1.000
$MedRQ$	0.000	0.000	0.635	0.000	0.000
$Min3RQ$	1.000	0.000	-0.003	0.000	0.000
$MinMedRQ$	0.000	0.000	-0.060	0.000	0.000
Asymptotic Variance	28.632	18.542	12.099	18.542	18.542
Avarage Power (%)	72.6	72.5	70.5	69.2	68.0
Distance from normal	19805.161	7625.280	1905.870	93.485	41.584

Table 2.7: Weights for the new tests reported on Figure 2.3. The distance from the normal is calculated using the efficient GMM loss function calculated on the first four uncentered moments.

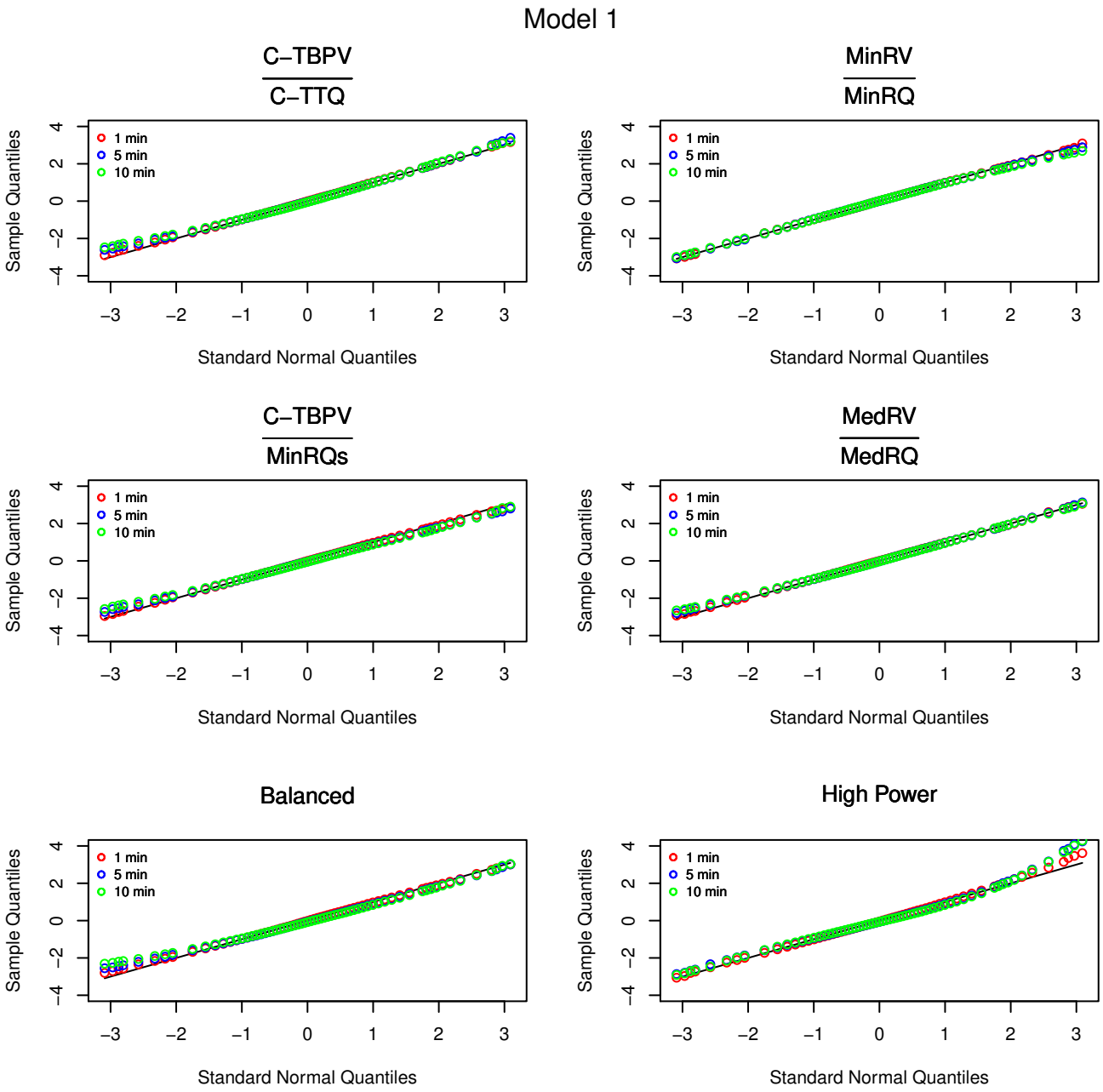


FIGURE2.4: Quantile - quantile plot of the alternative statistics against the normal distribution. The simulated price process is a standard Brownian motion.

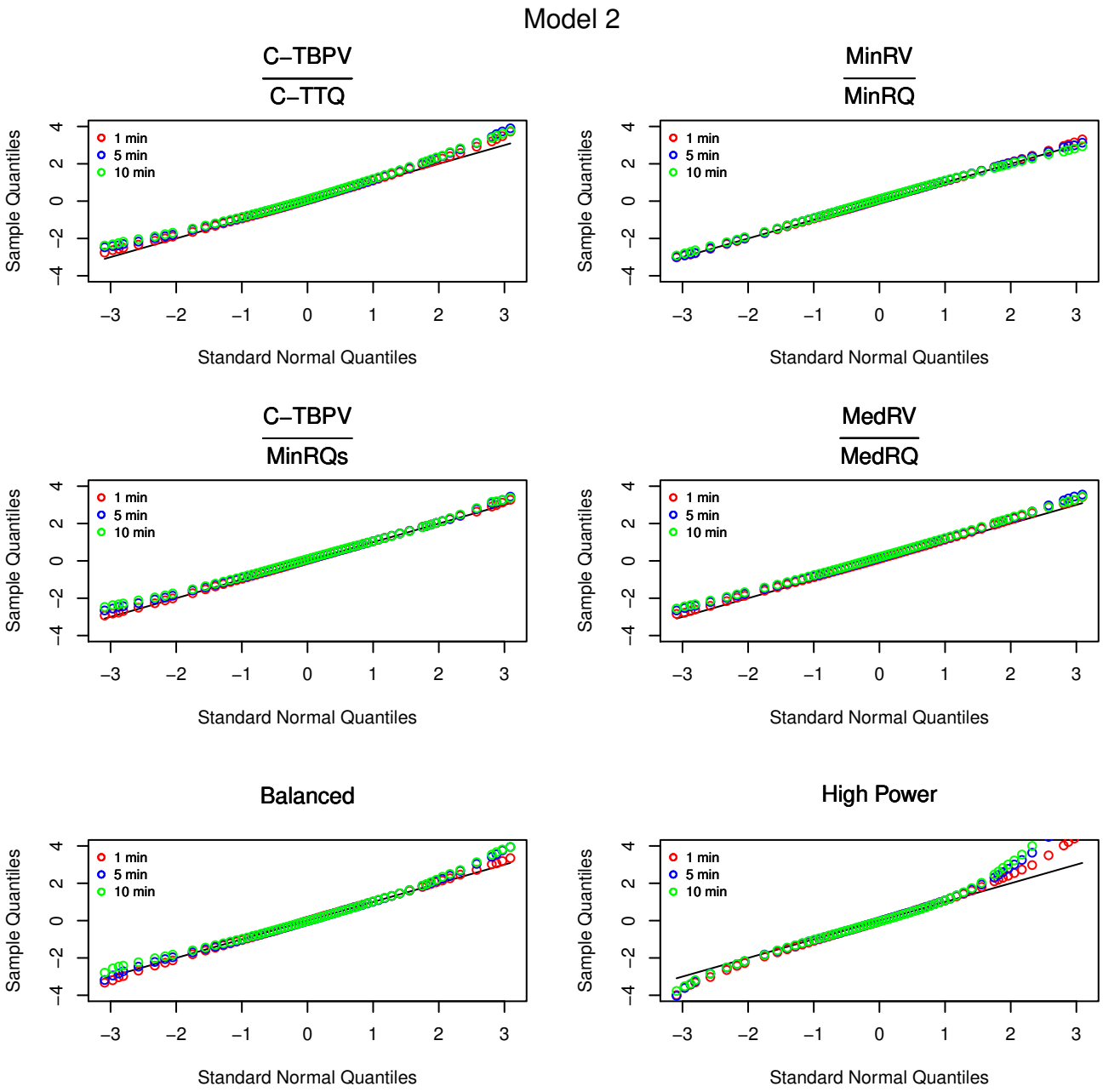


FIGURE2.5: Quantile - quantile plot of the alternative statistics against the normal distribution. The simulated price process is characterized by stochastic volatility and no jumps.

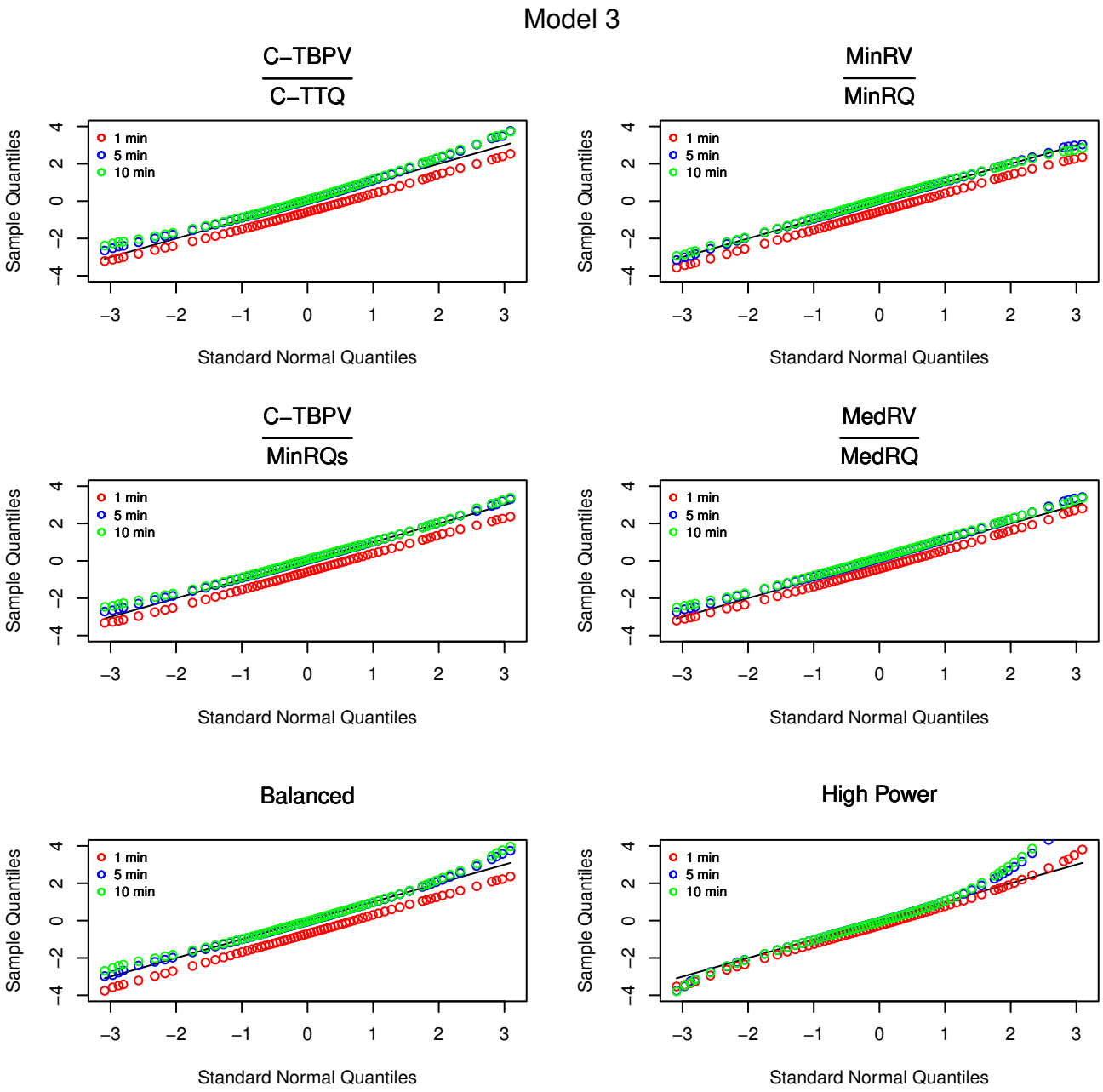


FIGURE2.6: Quantile - quantile plot of the alternative statistics against the normal distribution. The simulated price process is characterized by stochastic volatility, market microstructure noise and no jumps.

		No intraday volatility correction						Intraday volatility correction					
		Model 2		Model 4		Model 5		Model 2		Model 4		Model 5	
freq.	IV	actual	adj.	adj.	spurious	adj.	spurious	actual	adj.	adj.	spurious	adj.	spurious
(min)	measure	size	s.l.	power	jumps	power	jumps	size	s.l.	power	jumps	power	jumps
	m-LM	5.84	0.086	90.2	1	84.5	1.04	1.08	0.917	91.6	1.03	86.9	1.05
	High Power	2.87	0.147	89.2	0.99	83.6	0.98	2.09	0.340	90.0	1	84.1	1.01
	LM	17.16	04	87.9	0.96	81.5	0.95	3.58	0.234	90.6	0.99	85.3	0.99
	Balanced	1.33	0.687	87.1	1	79.5	1.06	0.97	1.055	88.1	0.99	80.5	0.97
1	C-TBPV/MinRQs	1.20	0.810	86.2	1.01	77.5	1.02	0.90	1.149	86.9	1	78.7	1.01
	CPR	1.80	0.496	85.6	0.98	76.8	0.97	1.31	0.752	86.3	0.99	78.2	0.99
	MedRV/MedRQ	1.76	0.525	84.8	1.03	75.5	1.02	1.18	0.785	85.9	1.02	77.0	1.03
	C-TTV/C-TTPQ	1.44	0.644	84.0	0.96	74.5	0.96	0.98	1.015	84.7	0.99	76.0	0.94
	MinRV/MinRQ	1.30	0.755	82.1	1.02	71.5	1.01	0.92	1.080	83.3	1.03	72.9	1.04
	m-LM	4.33	0.123	78.3	1.05	66.2	1.05	1.06	0.922	81.8	1	71.0	1.02
	High Power	3.91	0.014	78.9	1.02	66.6	1.02	2.30	0.167	79.7	1.04	68.7	1.04
	LM	8.52	0.025	75.4	0.89	60.7	0.80	3.91	0.181	78.4	0.95	65.7	0.89
	Balanced	1.70	0.386	75.8	1.01	62.6	1.04	0.92	1.129	77.3	1.04	64.9	1.07
5	C-TBPV/MinRQs	1.26	0.727	75.0	0.99	61.3	1.04	0.66	1.433	76.3	1	63.5	1.04
	CPR	2.31	0.279	73.7	1.01	59.2	1.04	1.44	0.645	75.3	1.01	61.4	1.01
	MedRV/MedRQ	2.01	0.408	72.3	1	58.0	1.01	1.24	0.781	74.3	0.99	60.1	1.02
	C-TTV/C-TTPQ	1.60	0.587	71.5	0.96	56.7	0.98	0.84	1.188	73.1	0.96	59.1	0.95
	MinRV/MinRQ	1.12	0.857	68.0	1.01	53.4	1.04	0.81	1.204	70.0	1.01	55.3	1.02
	m-LM	3.24	0.179	71.3	1.04	55.5	0.97	1.10	0.915	74.9	1	61.3	0.94
	High Power	4.59	03	70.0	1	54.7	1	2.35	0.112	72.6	1.03	58.0	1
	LM	6.59	0.048	67.3	0.86	48.3	0.76	3.86	0.172	69.9	0.86	53.9	0.69
	Balanced	2.06	0.337	67.7	1.03	50.9	1.01	0.94	1.099	69.8	1.05	54.1	1.07
10	C-TBPV/MinRQs	1.41	0.646	66.3	1.05	49.9	1.08	0.67	1.599	68.9	1.08	53.3	1.10
	CPR	2.53	0.234	65.0	1.03	47.1	0.99	1.52	0.555	67.0	1.04	50.2	1.03
	MedRV/MedRQ	2.20	0.426	63.8	1.01	46.0	1.05	1.16	0.905	66.1	1.02	49.2	1.06
	C-TTV/C-TTPQ	1.62	0.641	62.4	1	43.9	1.02	0.83	1.215	64.5	1.03	46.6	1.08
	MinRV/MinRQ	0.86	1.133	59.3	1.02	41.2	1.01	0.57	1.526	61.6	1.02	43.3	1.02

Table 2.8: Actual size, adjusted significance level, size-adjusted power and spurious jump detection rates of the non-parametric tests in presence of stochastic volatility with an intraday U-shaped pattern. Results are expressed in percentage and are presented with and without applying the intraday periodicity correction, in presence of large Poisson jumps (model 4) and small Hawkes jumps (model 5).

		No intraday volatility correction						Intraday volatility correction					
		Model 3		Model 6		Model 7		Model 3		Model 6		Model 7	
freq.	IV	actual	adj.	adj.	spurious	adj.	spurious	actual	adj.	adj.	spurious	adj.	spurious
(min)	measure	size	s.l.	power	jumps	power	jumps	size	s.l.	power	jumps	power	jumps
	m-LM	1.74	0.523	87.8	0.99	80.7	1.03	1.08	0.917	91.6	1.03	86.9	1.05
	High Power	1.17	0.755	86.8	1.02	78.4	1.04	2.09	0.340	90.0	1	84.1	1.01
	LM	6.04	0.061	85.9	0.91	77.6	0.95	3.58	0.234	90.6	0.99	85.3	0.99
	Balanced	0.11	5.433	83.8	1.04	73.6	1.02	0.97	1.055	88.1	0.99	80.5	0.97
1	C-TBPV/MinRQs	0.11	4.499	82.0	1.01	72.0	0.96	0.90	1.149	86.9	1	78.7	1.01
	CPR	0.18	3.870	81.8	1.01	71.4	0.96	1.31	0.752	86.3	0.99	78.2	0.99
	MedRV/MedRQ	0.36	2.656	80.3	0.99	69.5	0.99	1.18	0.785	85.9	1.02	77.0	1.03
	C-TTV/C-TTPQ	0.29	2.893	79.3	1.02	68.8	0.96	0.98	1.015	84.7	0.99	76.0	0.94
	MinRV/MinRQ	0.12	4.148	77.5	0.96	65.0	0.91	0.92	1.080	83.3	1.03	72.9	1.04
	m-LM	3.28	0.238	77.9	1.02	65.6	1.02	1.06	0.922	81.8	1	71.0	1.02
	High Power	3.67	0.016	77.3	1.02	64.0	1.06	2.30	0.167	79.7	1.04	68.7	1.04
	LM	6.88	0.043	74.6	0.91	59.4	0.86	3.91	0.181	78.4	0.95	65.7	0.89
	Balanced	1.56	0.528	74.7	1.01	60.4	1.04	0.92	1.129	77.3	1.04	64.9	1.07
5	C-TBPV/MinRQs	1.23	0.776	73.3	1.02	59.0	1.08	0.66	1.433	76.3	1	63.5	1.04
	CPR	2.25	0.361	72.2	1.02	57.5	1.06	1.44	0.645	75.3	1.01	61.4	1.01
	MedRV/MedRQ	1.88	0.457	70.4	1.01	55.7	1.02	1.24	0.781	74.3	0.99	60.1	1.02
	C-TTV/C-TTPQ	1.60	0.643	70.1	1	54.2	1.02	0.84	1.188	73.1	0.96	59.1	0.95
	MinRV/MinRQ	1.07	0.928	66.7	0.97	50.7	0.98	0.81	1.204	70.0	1.01	55.3	1.02
	m-LM	3.01	0.228	70.7	1.02	54.8	1	1.10	0.915	74.9	1	61.3	0.94
	High Power	4.30	06	70.1	1.02	53.4	1.02	2.35	0.112	72.6	1.03	58.0	1
	LM	6.14	0.055	66.3	0.84	47.5	0.69	3.86	0.172	69.9	0.86	53.9	0.69
	Balanced	1.90	0.391	66.5	1.05	49.7	1.04	0.94	1.099	69.8	1.05	54.1	1.07
10	C-TBPV/MinRQs	1.29	0.721	65.4	1.05	48.9	1.04	0.67	1.599	68.9	1.08	53.3	1.10
	CPR	2.38	0.279	64.8	1.02	46.3	1.02	1.52	0.555	67.0	1.04	50.2	1.03
	MedRV/MedRQ	1.99	0.464	62.5	1.04	45.1	1.09	1.16	0.905	66.1	1.02	49.2	1.06
	C-TTV/C-TTPQ	1.62	0.602	61.2	1.05	42.2	1.04	0.83	1.215	64.5	1.03	46.6	1.08
	MinRV/MinRQ	0.87	1.127	58.9	0.99	39.5	1.04	0.57	1.526	61.6	1.02	43.3	1.02

Table 2.9: Actual size, adjusted significance level, size-adjusted power and spurious jump detection rates of the non-parametric tests in presence of stochastic volatility with an intraday U-shaped pattern and market microstructure noise. Results are expressed in percentage and are presented with and without applying the intraday periodicity correction, in presence of large Poisson jumps (model 6) and small Hawkes jumps (model 7).

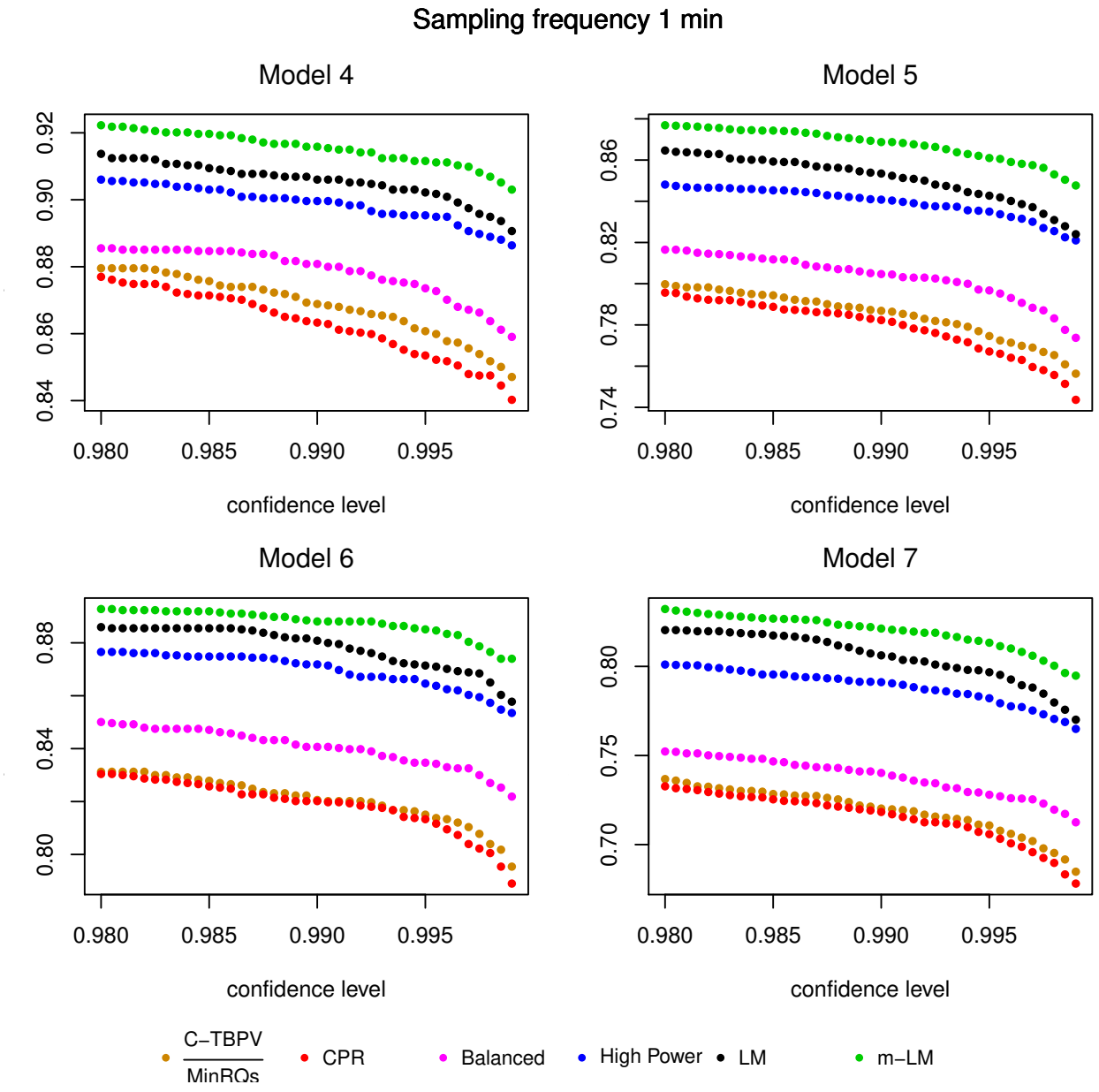


FIGURE2.7: Size adjusted power of the alternative tests under the different data generating processes.

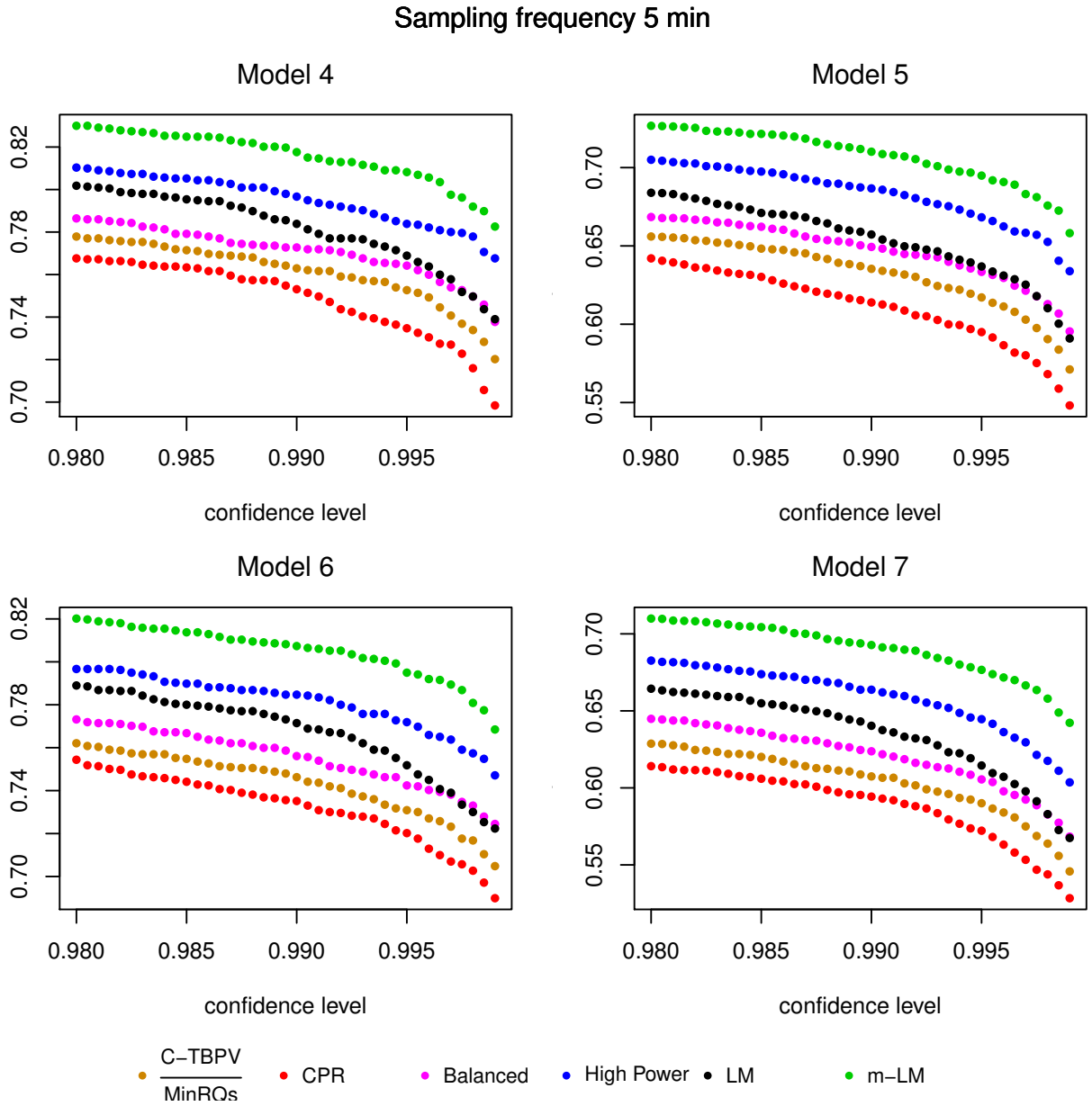


FIGURE2.8: Size adjusted power of the alternative tests under the different data generating processes.

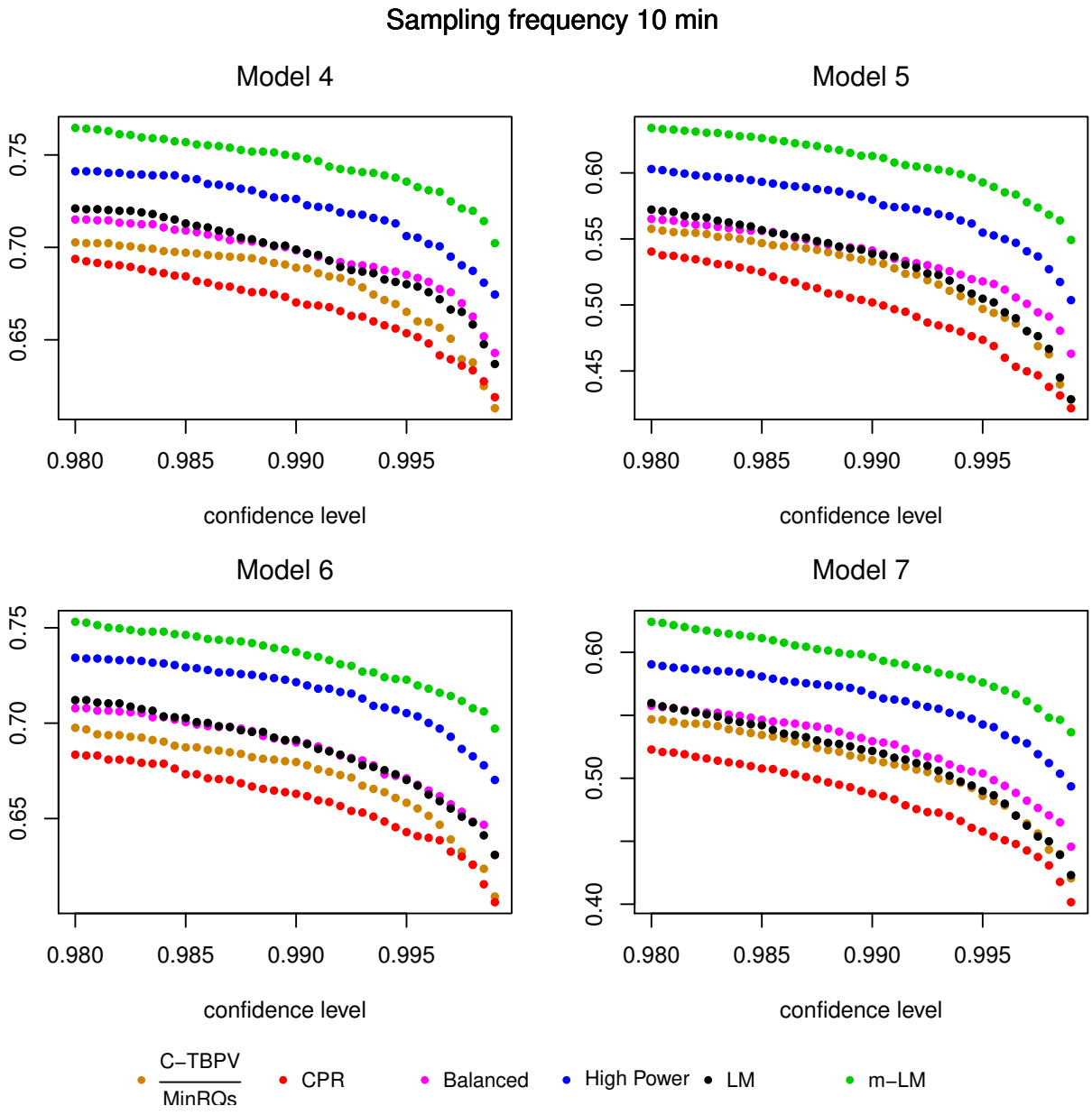


FIGURE2.9: Size adjusted power of the alternative tests under the different data generating processes.

2.7 Empirical Analysis

In this section we apply the non-parametric tests previously developed to empirical data. Our data set is the same described in Chapter 1 which contains intraday market quotes reported every 5 minutes from September 2007 to April 2014 for the S&P 500 and the Euro Stoxx 50. Before computing all the measures of interest, stale quotes have been removed from the data set. This

operation leaves the realized quadratic variation unchanged but reduces the impact of zero returns on the jump robust measures: while an isolated stale quote generates a single null contribution to the realized quadratic variation, it can indeed cause two or three consecutive blocks to provide a zero contribution to the corresponding measure, introducing therefore a positive bias on RJ that can increase the false detection rate. If the intraday stale quotes are removed, their impact on is on average the same as the impact on the quadratic variation and the downward bias on RJ is is consequently removed.

A genuine comparison between the performances of the different tests is of course extremely problematic in an empirical applications since the real data generating process is unknown. In this section we adopt a different approach whose purpose is to show that the outcomes on real data are very similar to those obtained from the numerical simulation. We therefore compute the fraction of trading days containing at least one jump at a given confidence level according to the different tests. To take into account the significant deviations from normality that some statistics exhibit in finite samples, we adjust the nominal confidence levels to match the actual size of the test at 5 minutes under model 3 (which incorporates stochastic volatility with a U-shaped intraday pattern and market microstructure noise). In doing this we must consider that the number of intraday returns on real data is generally different from the numerical simulation: on standard trading days we have 77 returns for the S&P 500 and 100 returns for the Euro Stoxx 50 but of course these numbers can reduce due to the elimination of stale quotes or to the fact that some trading days are shorter. To properly adjust the significance level we adopt the following formula:

$$\alpha^{*}\left(M\right)=1-\phi\left(P_{2}^{\alpha}\left(M^{-1/2}\right)\right)$$

where α is the nominal significance level and P_2^{α} is a polynomial of order which interpolates the quantiles corresponding to the adjusted confidence levels calculated on simulated data for $M = 39, 78, 390$. Despite the simulated return process is designed to possibly mimic the salient features observed on real data, we are not guaranteed that the adjustment is well suited for the empirical application because the actual size of finite sample effects is indeed unknown. This can therefore potentially affect the outcomes of our analysis. Figure 2.10, reporting the fraction of detected jump days on real data by each test, is very similar to the power comparison of Figure 2.8. This suggests that the proposed tests are probably able to detect more jumps also on real data compared to the CPR test, though our ignorance on the true data generating process and the unavailability of tick-by-tick data prevents us from drawing sound conclusions in this regard. Importantly the number

of jumps discovered by the new tests remains larger also when the confidence level is more severe. Note that the variant of the CPR test constructed with the staggered MinRQ is detecting more jumps after adjusting the confidence level. At the nominal confidence level of 99% and without applying any correction, the number of jumps detected by the CPR test is actually larger. This happens also on simulated data where the additional jumps detected by the CPR test are spurious. As already mentioned, this depends probably due to different bias generated in the test statistics by different choices of the quarticity estimator (expression 2.36). Using the staggered MinRQ such a bias is apparently smaller and the adjustment required to the confidence level to match the actual size is less severe.

To check the consistency of the outcomes across the different tests, in analogy with Schwert (2010), we report on Tables 2.10, 2.11 and 2.12 the contingency matrices calculated at the 99% confidence level (applying the finite sample corrections mentioned above) for the S&P 500, the Euro Stoxx 50 and the simulated data from model 7. The contingency matrices contain on the main diagonal the fraction of jump days detected by each tests while the off-diagonal elements are the intersections between sets of jump events detected by different tests. On simulated data a large fraction of jumps detected by one test is also confirmed by other tests. On real data the outcomes are less consistent suggesting that the features of the simulated process may differ from those characterizing real data (e.g. distribution of the jump size, presence of volatility jumps, different noise to signal ratio, etc.).

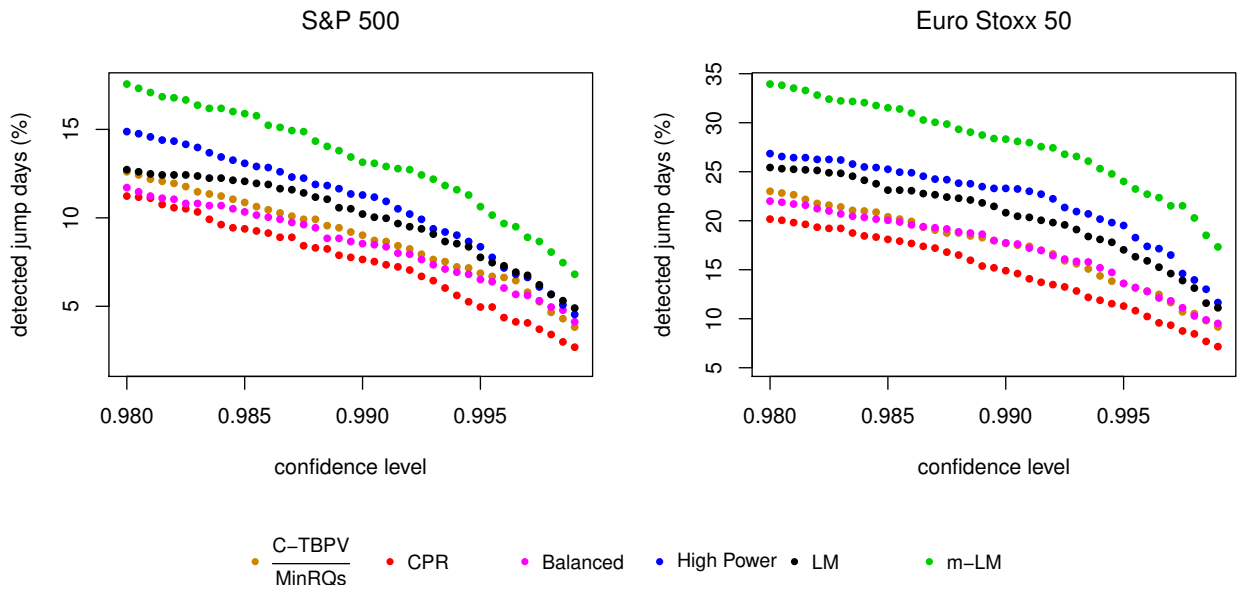


FIGURE2.10: Fraction of trading days when the null hypothesis of continuous price evolution is rejected.

	$\frac{MinRV}{MinRQ}$	$\frac{C-TTV}{C-TTPQ}$	$\frac{MedRV}{MedRQ}$	CPR	$\frac{C-TBPV}{MinRQs}$	Balanced	HP	LM	m-LM
$\frac{MinRV}{MinRQ}$	13.5	8.2	9.4	9.9	10.3	9.6	9.4	8.0	10.4
$\frac{C-TTV}{C-TTPQ}$		11.8	7.9	10.5	10.8	9.2	9.5	7.7	9.6
$\frac{MedRV}{MedRQ}$			15.8	10.1	11.9	11.2	12.2	9.7	12.8
CPR				14.9	14.2	13.1	12.1	9.8	12.5
$\frac{C-TBPV}{MinRQs}$					17.7	14.5	14.1	10.9	14.6
Balanced						17.7	14.6	11.8	15.2
HP							23.3	15.3	20.3
LM								20.8	17.9
m-LM									28.3

Table 2.10: Contingency matrix for the different tests applied to the Euro Stoxx 50.

	$\frac{MinRV}{MinRQ}$	$\frac{C-TTV}{C-TTPQ}$	$\frac{MedRV}{MedRQ}$	CPR	$\frac{C-TBPV}{MinRQs}$	Balanced	HP	LM	m-LM
$\frac{MinRV}{MinRQ}$	7.6	4.1	5.3	5.0	5.3	4.8	4.3	2.6	4.3
$\frac{C-TTV}{C-TTPQ}$		6.0	3.8	5.4	5.3	4.4	4.4	2.5	3.9
$\frac{MedRV}{MedRQ}$			8.2	4.8	5.6	5.0	4.9	3.0	4.9
CPR				7.6	6.7	6.3	5.6	3.3	5.3
$\frac{C-TBPV}{MinRQs}$					9.0	7.0	6.8	4.1	6.2
Balanced						8.5	6.7	4.5	6.5
HP							11.3	6.0	8.8
LM								10.2	7.8
m-LM									13.1

Table 2.11: Contingency matrix for the different tests applied to the S&P 500.

	$\frac{MinRV}{MinRQ}$	$\frac{C-TTV}{C-TTPQ}$	$\frac{MedRV}{MedRQ}$	CPR	$\frac{C-TBPV}{MinRQs}$	Balanced	HP	LM	m-LM
$\frac{MinRV}{MinRQ}$	13.1	12.3	12.5	12.6	12.6	12.5	12.3	12.0	12.3
$\frac{C-TTV}{C-TTPQ}$		13.9	12.8	13.4	13.4	13.1	13.0	12.7	13.1
$\frac{MedRV}{MedRQ}$			14.3	13.3	13.6	13.4	13.5	13.1	13.4
CPR				14.6	14.3	14.1	13.9	13.3	13.8
$\frac{C-TBPV}{MinRQs}$					14.9	14.4	14.2	13.6	14.2
Balanced						15.3	14.6	14.0	14.6
HP							16.2	14.6	15.6
LM								15.6	15.0
m-LM									16.9

Table 2.12: Contingency matrix for the different tests applied to simulated data in presence of stochastic volatility and market microstructure noise (model 7) at the sampling frequency of 5 min.

2.8 Concluding Remarks

We have introduced a new class of asymptotically consistent estimators for the integrated volatility and the integrated quarticity, obtained as weighted combinations of multiple measures which are robust to price jumps. The weights can be determined minimizing the asymptotic variance but it is also possible control some specific features in finite samples. Remarkably, the maximum asymptotic efficiency that can be reached combining multiple measures of the integrated volatility which involve at most three consecutive returns is much closer to the Maximum Likelihood limit than other estimators proposed in the literature.

We have shown that the gain in asymptotic variance provides substantial advantages for the construction of non-parametric tests devoted to detect the presence of jumps. The new tests proposed in this chapter achieve, in realistic scenarios, a detection power comparable to the test of Lee and Mykland (2008) while preserving consistency under more general features of the volatility process, including also volatility jumps. The recent paper of Bajgrowicz et al. (2016) highlights that most of the jumps detected by standard non-parametric tests are indeed spurious. The solution they propose to overcome this issue is to use extremely severe confidence levels poses the problem of drastically reducing the power of the tests and possibly neglect jumps with a small size. The method presented in this chapter allows to define new non-parametric detection methods possessing more power than standard once also when the confidence level is particularly severe. The application to real data leads to the detection of a larger number of presumed jumps in comparison to the BNS and the CPR tests, even after correcting for finite sample effects and setting a severe confidence level. Importantly, the large number of degrees of freedom at our disposal when combining multiple volatility and quarticity measures, allows not only to increment the power of our tests but also to improve their properties in finite samples and setting a trade off between these features. Surprisingly we find that the maximizing the accuracy of the integrated quarticity estimates is not necessarily the optimal choice to increase the power of our tests. In conclusion, the present work suggests that combining different jump robust measures is a promising technique for the development of new tools in high frequency financial econometrics.

Appendix B

Technical Details of Chapter 2

B.1 Covariance Matrix of the Jump Robust Volatility Estimators

The elements of the covariance matrix 2.12 can be calculated for a driftless Brownian motion with constant volatility and no jumps. Al the finite sample effects due to stochastic volatility and jumps decay faster than $M^{-1/2}$ where M is the number of equally spaced intraday quotes. To compute the variance-covariance matrix, let us consider a sequence of i.i.d. standard normal random variables U_i with $i = 1, \ldots, N$. Since our jump-robust measures involve blocks containing at most three consecutive returns, they are characterized by two possible kinds of functional forms:

$$F_2\left(U_1, \ldots, U_M\right) = \frac{1}{M-1} \sum_{i=2}^{M} \Theta_{F_2}\left(\left|U_{i-1}\right|, \left|U_i\right|\right)$$

$$F_3\left(U_1, \ldots, U_M\right) = \frac{1}{M-2} \sum_{i=3}^{M} \Theta_{F_3}\left(\left|U_{i-1}\right|, \left|U_{i-1}\right|, \left|U_i\right|\right)$$

The covariance matrix between two generic measures F and G can therefore take three distinct forms

1. Covariance between 2 estimators based on blocks of 2 returns:

$$\begin{aligned} Cov\left[F_2, G_2\right] &= \frac{1}{M-1} Cov\left[\Theta_{F_2}\left(\left|U_1\right|, \left|U_2\right|\right), \Theta_{G_2}\left(\left|U_1\right|, \left|U_2\right|\right)\right] + \\ & \quad 2\,\frac{M-2}{\left(M-1\right)^2} Cov\left[\Theta_{F_2}\left(\left|U_1\right|, \left|U_2\right|\right), \Theta_{G_2}\left(\left|U_2\right|, \left|U_3\right|\right)\right] \end{aligned} \tag{B.1}$$

2. Covariance between estimators based on blocks of 2 and 3 returns respectively

$$\begin{aligned} Cov\left[F_2, G_3\right] &= \frac{2}{M-1} Cov\left[\Theta_{F_2}\left(|U_1|, |U_2|\right), \Theta_{G_3}\left(|U_1|, |U_2|, |U_3|\right)\right] + \\ & \frac{2\left(M-3\right)}{\left(M-1\right)\left(M-2\right)} Cov\left[\Theta_{F_2}\left(|U_1|, |U_2|\right), \Theta_{G_3}\left(|U_2|, |U_3|, |U_4|\right)\right] \end{aligned} \tag{B.2}$$

3. Covariance between estimators based on blocks of 3 returns

$$\begin{aligned} Cov\left[F_3, G_3\right] &= \frac{1}{M-2} Cov\left[\Theta_{F_3}\left(|U_1|, |U_2|, |U_3|\right), \Theta_{G_3}\left(|U_1|, |U_2|, |U_3|\right)\right] + \\ & 2\frac{M-3}{\left(M-2\right)^2} Cov\left[\Theta_{F_3}\left(|U_1|, |U_2|, |U_3|\right), \Theta_{G_3}\left(|U_2|, |U_3|, |U_4|\right)\right] + \\ & 2\frac{M-4}{\left(M-2\right)^2} Cov\left[\Theta_{F_3}\left(|U_1|, |U_2|, |U_3|\right), \Theta_{G_3}\left(|U_3|, |U_4|, |U_5|\right)\right] \end{aligned} \tag{B.3}$$

The expected values required to calculate these expressions are reported below. The computations involve multi-dimensional integrals which in many cases can be solved analytically, in the other cases we resort to numerical methods. Further details are discussed in the second part of the appendix.

$$\mathbb{E}\left[|U|^{\gamma}\right] = \frac{2^{\gamma/2}}{\sqrt{\pi}}\Gamma\left(\frac{\gamma+1}{2}\right) \tag{B.4}$$

$$\mathbb{E}\left[\min\left(|U_1|, |U_2|\right)^4\right] = 3 - \frac{8}{\pi} \tag{B.5}$$

$$\mathbb{E}\left[\min\left(|U_1|, |U_2|\right)^2 \min\left(|U_2|, |U_3|\right)^2\right] = \frac{5}{3} - \frac{8}{\pi} + \frac{6}{\pi\sqrt{3}} \tag{B.6}$$

$$\mathbb{E}\left[\mathrm{med}\left(|U_1|, |U_2|, |U_3|\right)^2 \mathrm{med}\left(|U_1|, |U_2|, |U_3|\right)^2\right] = \frac{9\pi + 72 - 52\sqrt{3}}{3\pi} \tag{B.7}$$

$$\mathbb{E}\left[\mathrm{med}\left(|U_1|, |U_2|, |U_3|\right)^2 \mathrm{med}\left(|U_2|, |U_3|, |U_4|\right)^2\right] = \frac{5}{3} + \frac{16}{\pi} - \frac{28}{\pi\sqrt{3}} - \frac{8}{\pi^2} \tag{B.8}$$

$$\mathbb{E}\left[\mathrm{med}\left(|U_1|, |U_2|, |U_3|\right)^2 \mathrm{med}\left(|U_3|, |U_4|, |U_5|\right)^2\right] = 0.6304889 \tag{B.9}$$

$$\mathbb{E}\left[\min\left(|U_1|, |U_2|, |U_3|\right)^2 \min\left(|U_1|, |U_2|, |U_3|\right)^2\right] = 3 - \frac{24}{\pi} + \frac{26}{\pi\sqrt{3}} \tag{B.10}$$

$$\mathbb{E}\left[\min\left(|U_1|, |U_2|, |U_3|\right)^2 \min\left(|U_2|, |U_3|, |U_4|\right)^2\right] = 0.081154612 \tag{B.11}$$

$$\mathbb{E}\left[\min\left(|U_1|,|U_2|,|U_3|\right)^2 \min\left(|U_3|,|U_4|,|U_5|\right)^2\right] = 0.052979325 \tag{B.12}$$

$$\mathbb{E}\left[\min\left(|U_1|,|U_2|,|U_3|\right)^2 \operatorname{med}\left(|U_1|,|U_2|,|U_3|\right)^2\right] = 1 - \frac{4}{\pi\sqrt{3}} \tag{B.13}$$

$$\begin{aligned}\mathbb{E}\big[&\min\left(|U_1|,|U_2|,|U_3|\right)\operatorname{med}\left(|U_1|,|U_2|,|U_3|\right)\\ &\min\left(|U_2|,|U_3|,|U_4|\right)\operatorname{med}\left(|U_2|,|U_3|,|U_4|\right)\big] \quad = \quad 0.18479863\end{aligned} \tag{B.14}$$

$$\begin{aligned}\mathbb{E}\big[&\min\left(|U_1|,|U_2|,|U_3|\right)\operatorname{med}\left(|U_1|,|U_2|,|U_3|\right)\\ &\min\left(|U_3|,|U_4|,|U_5|\right)\operatorname{med}\left(|U_3|,|U_4|,|U_5|\right)\big] \quad = \quad 0.13180674\end{aligned} \tag{B.15}$$

$$\mathbb{E}\left[\min\left(|U_1|,|U_2|\right)^2 |U_1|\,|U_2|\right] = \frac{2}{\pi} \tag{B.16}$$

$$\mathbb{E}\left[\min\left(|U_1|,|U_2|\right)^2 |U_2|\right] = \frac{2^{3/2}-2}{\sqrt{\pi}} \tag{B.17}$$

$$\mathbb{E}\left[\operatorname{med}\left(|U_1|,|U_2|,|U_3|\right)^2 |U_1|\right] = 0.77817813 \tag{B.18}$$

$$\mathbb{E}\left[\operatorname{med}\left(|U_1|,|U_2|,|U_3|\right)^2 |U_1|\,|U_2|\right] = \frac{4}{\pi}\left(\frac{1}{\sqrt{3}} + \frac{3}{2} - \sqrt{2}\right) \tag{B.19}$$

$$\mathbb{E}\left[\min\left(|U_1|,|U_2|,|U_3|\right)^2 |U_1|\right] = 0.22326864 \tag{B.20}$$

$$\mathbb{E}\left[\min\left(|U_1|,|U_2|,|U_3|\right)^2 |U_1|\,|U_2|\right] = \frac{2}{\pi}\left(1 - \frac{1}{\sqrt{3}}\right) \tag{B.21}$$

$$\mathbb{E}\left[\min\left(|U_1|,|U_2|,|U_3|\right)\operatorname{med}\left(|U_1|,|U_2|,|U_3|\right)|U_1|\,|U_2|\right] = \frac{1}{3} - \frac{2}{\pi\sqrt{3}} + \frac{\sqrt{2}}{\pi} \tag{B.22}$$

$$\mathbb{E}\left[\min\left(|U_1|,|U_2|,|U_3|\right)\operatorname{med}\left(|U_1|,|U_2|,|U_3|\right)|U_1|\right] = 0.35575564 \tag{B.23}$$

$$\mathbb{E}\left[\min\left(|U_1|,|U_2|\right)^2 \operatorname{med}\left(|U_1|,|U_2|,|U_3|\right)^2\right] = \frac{5}{3} - \frac{20}{\pi\sqrt{3}} + \frac{8}{\pi} \tag{B.24}$$

$$\mathbb{E}\left[\min\left(|U_1|,|U_2|\right)^2 \operatorname{med}\left(|U_2|,|U_3|,|U_4|\right)^2\right] \quad = \quad 0.36453612 \tag{B.25}$$

$$\mathbb{E}\left[\min\left(|U_1|,|U_2|\right)^2 |U_1|^{2/3}\,|U_2|^{2/3}\right] \quad = \quad \frac{2^{8/3}}{\pi}\left[\frac{5}{6}\,\Gamma\left(\frac{5}{6}\right)^2 - \frac{1}{2^{5/3}}\Gamma\left(\frac{5}{3}\right)\right] \tag{B.26}$$

$$\mathbb{E}\left[\min\left(|U_1|,|U_2|\right)^2 |U_2|^{2/3}\right] = 0.41952385 \tag{B.27}$$

$$\mathbb{E}\left[\min\left(|U_1|,|U_2|\right)^2\min\left(|U_1|,|U_2|,|U_3|\right)^2\right]=0.18079593 \tag{B.28}$$

$$\mathbb{E}\left[\min\left(|U_1|,|U_2|\right)^2\min\left(|U_2|,|U_3|,|U_4|\right)^2\right]=0.1065999 \tag{B.29}$$

$$\mathbb{E}\left[\min\left(|U_1|,|U_2|\right)^2\min\left(|U_1|,|U_2|,|U_3|\right)\operatorname{med}\left(|U_1|,|U_2|,|U_3|\right)\right]=0.27146358 \tag{B.30}$$

$$\mathbb{E}\left[\min\left(|U_1|,|U_2|\right)^2\min\left(|U_2|,|U_3|,|U_4|\right)\operatorname{med}\left(|U_2|,|U_3|,|U_4|\right)\right]=0.16853585 \tag{B.31}$$

$$\mathbb{E}\left[\operatorname{med}\left(|U_1|,|U_2|,|U_3|\right)^2\left|U_1U_2U_3\right|^{2/3}\right]=0.69275742 \tag{B.32}$$

$$\mathbb{E}\left[\operatorname{med}\left(|U_1|,|U_2|,|U_3|\right)^2\left|U_1U_2\right|^{2/3}\right]=0.71862343 \tag{B.33}$$

$$\mathbb{E}\left[\operatorname{med}\left(|U_1|,|U_2|,|U_3|\right)^2\left|U_1\right|^{2/3}\right]=0.71697316 \tag{B.34}$$

$$\mathbb{E}\left[\min\left(|U_1|,|U_2|,|U_3|\right)^2\operatorname{med}\left(|U_2|,|U_3|,|U_4|\right)^2\right]=0.23168651 \tag{B.35}$$

$$\mathbb{E}\left[\min\left(|U_1|,|U_2|,|U_3|\right)^2\operatorname{med}\left(|U_3|,|U_4|,|U_5|\right)^2\right]=0.17730031 \tag{B.36}$$

$$\mathbb{E}\left[\min\left(|U_1|,|U_2|,|U_3|\right)\operatorname{med}\left(|U_1|,|U_2|,|U_3|\right)^3\right]=\frac{4}{\pi}\left(\frac{9}{2}-\frac{15}{2\sqrt{2}}+\frac{2}{\sqrt{3}}\right) \tag{B.37}$$

$$\mathbb{E}\left[\min\left(|U_1|,|U_2|,|U_3|\right)\operatorname{med}\left(|U_1|,|U_2|,|U_3|\right)\operatorname{med}\left(|U_2|,|U_3|,|U_4|\right)^2\right]=0.36626489 \tag{B.38}$$

$$\mathbb{E}\left[\min\left(|U_1|,|U_2|,|U_3|\right)\operatorname{med}\left(|U_1|,|U_2|,|U_3|\right)\operatorname{med}\left(|U_3|,|U_4|,|U_5|\right)^2\right]=0.28365967 \tag{B.39}$$

$$\mathbb{E}\left[\min\left(|U_1|,|U_2|,|U_3|\right)^2\left|U_1U_2U_3\right|^{2/3}\right]=0.25338328 \tag{B.40}$$

$$\mathbb{E}\left[\min\left(|U_1|,|U_2|,|U_3|\right)^2\left|U_1U_2\right|^{2/3}\right]=0.22646547 \tag{B.41}$$

$$\mathbb{E}\left[\min\left(|U_1|,|U_2|,|U_3|\right)^2\left|U_1\right|^{2/3}\right]=0.20682189 \tag{B.42}$$

$$\mathbb{E}\left[\min\left(|U_1|,|U_2|,|U_3|\right)\operatorname{med}\left(|U_1|,|U_2|,|U_3|\right)\left|U_1U_2U_3\right|^{2/3}\right]=0.37964422 \tag{B.43}$$

$$\mathbb{E}\left[\min\left(|U_1|,|U_2|,|U_3|\right)\operatorname{med}\left(|U_1|,|U_2|,|U_3|\right)\left|U_1U_2\right|^{2/3}\right]=0.35209542 \tag{B.44}$$

$$\mathbb{E}\left[\min\left(|U_1|,|U_2|,|U_3|\right)\operatorname{med}\left(|U_1|,|U_2|,|U_3|\right)\left|U_1\right|^{2/3}\right]=0.32932346 \tag{B.45}$$

$$\mathbb{E}\left[\min\left(|U_1|,|U_2|,|U_3|\right)^3\operatorname{med}\left(|U_1|,|U_2|,|U_3|\right)\right]=\frac{4}{\pi}\left(\frac{5}{\sqrt{3}}-3\sqrt{2}+\frac{3}{2}\right) \tag{B.46}$$

$$\mathbb{E}\left[\min\left(|U_1|,|U_2|,|U_3|\right)\operatorname{med}\left(|U_1|,|U_2|,|U_3|\right)\min\left(|U_2|,|U_3|,|U_4|\right)^2\right]=0.1211352 \tag{B.47}$$

$$\mathbb{E}\left[\min\left(|U_1|,|U_2|,|U_3|\right)\operatorname{med}\left(|U_1|,|U_2|,|U_3|\right)\min\left(|U_3|,|U_4|,|U_5|\right)^2\right]=0.083314876 \tag{B.48}$$

Further details

The following integral is fundamental for the analytical computation of the expressions above:

$$I_n\left(\alpha_1,\ldots,\alpha_n;\beta_1,\ldots,\beta_n\right)=\int_0^\infty dy_1\ldots dy_n\,\mathbf{1}_{\{y_n\leqslant\ldots\leqslant y_1\}}\exp\left(-\sum_{j=1}^n\alpha_j y_j\right)\prod_{j=1}^n y_j^{\beta_j} \tag{B.49}$$

where $\alpha_j>0$. To guarantee the integrability around the origin the conditions $\sum_{j=1}^n\beta_j+n-1>0$ and $\beta_n>-1$ are necessary. Note that

$$\mathbb{E}\left[u_1^{\gamma_1}\ldots u_n^{\gamma_n}\mathbf{1}_{(|u_1|\leqslant\ldots\leqslant|u_n|)}\right]=\frac{2^{\sum_{i=1}^n\gamma_i/2}}{\pi^{n/2}}I_n\left(1,\ldots,1;\frac{\gamma_1-1}{2},\ldots,\frac{\gamma_n-1}{2}\right) \tag{B.50}$$

Integrating by parts the following property can be easily proved for $\beta_i>-1$, $a_i>0$ for $i=1,\ldots,n$ and $\beta_n>0$[1]:

$$\begin{aligned}I_n\left(\alpha_1,\ldots,\alpha_n;\beta_1,\ldots,\beta_n\right)=\frac{\beta_j}{\alpha_j}\,&I_n\left(\alpha_1,\ldots,\alpha_n;\beta_1,\ldots,\beta_j-1,\ldots,\beta_n\right)+\\&\frac{\delta_{j,1}-1}{\alpha_j}I_{n-1}\left(\alpha_1,\ldots,\alpha_{j-1}+\alpha_j,\alpha_{j+1},\ldots,\alpha_n;\right.\\&\qquad\left.\beta_1,\ldots,\beta_{j-1}+\beta_j,\beta_{j+1},\ldots,\beta_n\right)+\\&\frac{1-\delta_{j,n}}{\alpha_j}I_{n-1}\left(\alpha_1,\ldots,\alpha_j+\alpha_{j+1},\alpha_{j+2},\ldots,\alpha_n;\right.\\&\qquad\left.\beta_1,\ldots,\beta_j+\beta_{j+1},\beta_{j+2},\ldots,\beta_n\right)\end{aligned} \tag{B.51}$$

where $\delta_{j,1}$ denotes the Kroneker's delta. Let $\sigma\left(1,\ldots n\right)$ denote the set of all possible permutations

[1] In the case $\beta_n=0$ the integral on y_n gives $(1-e^{\alpha_n y_{n-1}})/\alpha_n$.

of the indexes $1, \ldots, n$:

$$\begin{aligned}
\sum_{\sigma} I_1\left(\alpha_{\sigma_1}, \ldots, \alpha_{\sigma_n}; \beta_{\sigma_1}, \ldots, \beta_{\sigma_n}\right) &= \int_0^{\infty} dy_1 \ldots dy_n \, \exp\left(-\sum_{j=1}^{n} \alpha_j y_j\right) \prod_{j=1}^{n} y_j^{\beta_j} \sum_{\sigma} \mathbf{1}_{\left(y_{\sigma_1} \leqslant \ldots \leqslant y_{\sigma_n}\right)} \\
&= \prod_{j=1}^{n} \int_0^{\infty} e^{-\alpha_j y} y^{\beta_j} dy \\
&= \prod_{j=1}^{n} I_1\left(\alpha_j; \beta_j\right) \qquad \text{(B.52)}
\end{aligned}$$

In the special case of equal arguments the function I_n can be calculated explicitly:

$$I_n\left(\alpha, \ldots, \alpha; \beta, \ldots, \beta\right) = \frac{1}{n!} \left[\frac{\Gamma\left(\beta+1\right)}{\alpha^{\beta+1}}\right]^n \qquad \text{(B.53)}$$

In addition to relation B.51, the following properties are enjoyed by I_2:

$$\begin{aligned}
I_2\left(\alpha_1, \alpha_2; \beta_1, \beta_2\right) &= \int_0^{+\infty} dx_1 \, e^{-\alpha_1 x_1} x_1^{\beta_1} \int_0^{x_1} dx_2 e^{-\alpha_2 x_2} x_2^{\beta_2} \\
&= \frac{1}{\alpha_1^{\beta_1+1}} \int_0^{+\infty} dy_1 \, e^{-y_1} y_1^{\beta_1} \int_0^{y_1/\alpha_1} dx_2 e^{-\alpha_2 x_2} x_2^{\beta_2} \\
&= \frac{1}{\alpha_1^{\beta_1+\beta_2+2}} \int_0^{+\infty} dy_1 \, e^{-y_1} y_1^{\beta_1} \int_0^{y_1} dy_2 e^{-y_2 \alpha_2/\alpha_1} y_2^{\beta_2} \\
&= \frac{1}{\alpha_1^{\beta_1+\beta_2+2}} I_2\left(1, \frac{\alpha_2}{\alpha_1}; \beta_1, \beta_2\right) \qquad \text{(B.54)}
\end{aligned}$$

$$\begin{aligned}
I_2\left(\alpha_1, \alpha_2; \beta_1, \beta_2\right) &= \frac{1}{\alpha_2^{\beta_2+1}} \int_0^{+\infty} dx_1 \, e^{-\alpha_1 x_1} x_1^{\beta_1} \int_0^{x_1 \alpha_2} dy_2 e^{-y_2} y_2^{\beta_2} \\
&= \frac{1}{\alpha_2^{\beta_1+\beta_2+2}} \int_0^{+\infty} dy_1 \, e^{-y_1 \alpha_1/\alpha_2} y_1^{\beta_1} \int_0^{y_1} dy_2 e^{-y_2} y_2^{\beta_2} \\
&= \frac{1}{\alpha_2^{\beta_1+\beta_2+2}} I_2\left(\frac{\alpha_1}{\alpha_2}, 1; \beta_1, \beta_2\right) \qquad \text{(B.55)}
\end{aligned}$$

these relations allow to reduce the dimensionality of the integral and to apply a series expansion:

$$\begin{aligned}
I_2(\alpha_1,\alpha_2;\beta_1,\beta_2) &= \frac{1}{\alpha_2^{\beta_1+\beta_2+2}}\int_0^{+\infty} dy_1\, e^{-y_1\alpha_1/\alpha_2} y^{\beta_1}\gamma(\beta_2+1,\, y) \\
&= \frac{1}{\alpha_2^{\beta_1+\beta_2+2}}\Gamma(\beta_2+1)\int_0^{+\infty} dy\, e^{-(1+\alpha_1/\alpha_2)\,y_1} y^{\beta_1+\beta_2+1}\sum_{k=0}^{\infty}\frac{y^k}{\Gamma(\beta_2+k+2)} \\
&= \frac{1}{\alpha_2^{\beta_1+\beta_2+2}}\Gamma(\beta_2+1)\sum_{k=0}^{\infty}\frac{I_1\left(1+\frac{\alpha_1}{\alpha_2},\, \beta_1+\beta_2+k+1\right)}{\Gamma(\beta_2+k+2)} \\
&= \frac{\Gamma(\beta_2+1)}{(\alpha_1+\alpha_2)^{\beta_1+\beta_2+2}}\sum_{k=0}^{\infty}\frac{\Gamma(\beta_1+\beta_2+k+2)}{\Gamma(\beta_2+k+2)\left(1+\frac{\alpha_1}{\alpha_2}\right)^k}
\end{aligned} \tag{B.56}$$

where we used the relation

$$\gamma(s,x) = \Gamma(s)\, e^{-x}\sum_{k=0}^{\infty}\frac{x^{s+k}}{\Gamma(k+s+1)}$$

being γ the lower incomplete gamma function: $\gamma(s,x) = \int_0^x dt\, e^{-t}t^{s-1}$. This technique can be extended also to the case $n=3$ obtaining a double series expansion:

$$\begin{aligned}
I_3(\alpha_1,\alpha_2,\alpha_2;\beta_1,\beta_2,\beta_2) &= \frac{1}{2}\int_0^{+\infty} dy_1 e^{-\alpha_1 y_1} y_1^{\beta_1}\left(\int_0^{y_1} dy_2 e^{-\alpha_2 y_2} y_2^{\beta_2}\right)^2 \\
&= \frac{1}{2\alpha_2^{2(\beta_2+1)}}\int_0^{+\infty} dy_1 e^{-\alpha_1 y_1} y_1^{\beta_1}\left(\int_0^{y_1/\alpha_2} dz_2\, e^{-z_2} z_2^{\beta_2}\right)^2 \\
&= \frac{\alpha_2^{\beta_1-2\beta_2-1}}{2}\int_0^{+\infty} dz_1 e^{-(\alpha_1\alpha_2+2)z_1} z_1^{\beta_1}\left(\int_0^{z_1} dz_2\, e^{-z_2} z_2^{\beta_2}\right)^2 \\
&= \frac{\alpha_2^{\beta_1-2\beta_2-1}}{2}\sum_{k=0}^{\infty} c_k(\beta_2)\int_0^{+\infty} dz\, e^{-\alpha_1\alpha_2 z} z^{\beta_1+\beta_2+k+1} \\
&= \frac{\alpha_2^{\beta_1-2\beta_2-1}}{2(\alpha_1\alpha_2+2)^{\beta_1+\beta_2+2}}\sum_{k=0}^{\infty}\frac{c_k(\beta_2)}{(\alpha_1\alpha_2+2)^k}\int_0^{+\infty} dz\, e^{-u} u^{\beta_1+\beta_2+k+1} \\
&= \frac{\alpha_2^{\beta_1-2\beta_2-1}}{2(\alpha_1\alpha_2+2)^{\beta_1+\beta_2+2}}\sum_{k=0}^{\infty}\frac{c_k(\beta_2)}{(\alpha_1\alpha_2+2)^k}\Gamma(\beta_1+\beta_2+k+2)
\end{aligned}$$

where

$$c_k(\beta_2) = \sum_{j=0}^{k} \frac{\Gamma(\beta_2+1)^2}{\Gamma(k-j+\beta_2+2)\,\Gamma(j+\beta_2+2)}$$

We provide some concrete examples:

Expression B.5 Let $k \in \mathbb{Z}$ and $k \geqslant 0$:

$$\begin{aligned}\mathbb{E}\left[\min(|u_1|,|u_2|)^{2k}\right] &= 2\mathbb{E}\left[u_1^{2k}\mathbf{1}_{(|u_1|\leqslant|u_2|)}\right] \\ &= \frac{2^{k+1}}{\pi} I_2\left(1,1;-\frac{1}{2},k-\frac{1}{2}\right)\end{aligned}$$

where the substitution $y_i = u_i^2/2$ has been applied. Using relation B.51

$$I_2\left(1,1;-\frac{1}{2},\frac{2k-1}{2}\right) = \frac{2k-1}{2} I_2\left(1,1;-\frac{1}{2},(k-1)-\frac{1}{2}\right) + \frac{(k-1)!}{2^k}$$

thus the following recursive relation is obtained

$$\mathbb{E}\left[\min(|u_1|,|u_2|)^{2k}\right] = (2k-1)\,\mathbb{E}\left[\min(|u_1|,|u_2|)^{2(k-1)}\right] - \frac{2}{\pi}(k-1)!$$

$$\begin{aligned}\mathbb{E}\left[\min(|u_1|,|u_2|)^{2}\right] &= 1-\frac{2}{\pi} \\ \mathbb{E}\left[\min(|u_1|,|u_2|)^{4}\right] &= 3-\frac{8}{\pi} \\ \mathbb{E}\left[\min(|u_1|,|u_2|)^{6}\right] &= 15-\frac{44}{\pi} \\ \mathbb{E}\left[\min(|u_1|,|u_2|)^{8}\right] &= 105-\frac{320}{\pi}\end{aligned}$$

Expression B.6 The expectation can be split as follows:

$$\mathbb{E}\left[\min(|u_1|,|u_2|)^2 \min(|u_2|,|u_3|)^2\right] = 2\,\mathbb{E}\left[u_2^4\mathbf{1}_{(|u_2|\leqslant|u_1|\leqslant|u_3|)}\right] + 4\,\mathbb{E}\left[u_1^2\,u_2^2\mathbf{1}_{(|u_1|\leqslant|u_2|\leqslant|u_3|)}\right]$$

each term is given by

$$\mathbb{E}\left[u_1^4 \mathbf{1}_{(|u_1|\leqslant|u_2|\leqslant|u_3|)}\right] = \frac{4}{\pi^{3/2}} I_3\left(1,1,1;-\frac{1}{2},-\frac{1}{2},\frac{3}{2}\right) \tag{B.57}$$

$$\mathbb{E}\left[u_1^2\, u_2^2 \mathbf{1}_{(|u_1|\leqslant|u_2|\leqslant|u_3|)}\right] = \frac{4}{\pi^{3/2}} I_3\left(1,1,1;-\frac{1}{2},\frac{1}{2},\frac{1}{2}\right) \tag{B.58}$$

applying relations B.51 and B.53 the RHS of B.57 gives

$$\begin{aligned} I_3\left(1,1,1;-\frac{1}{2},-\frac{1}{2},\frac{3}{2}\right) &= \frac{3}{2} I_3\left(1,1,1;-\frac{1}{2},-\frac{1}{2},\frac{1}{2}\right) - I_2\left(1,2;-\frac{1}{2},1\right) \\ &= \pi^{1/2}\left(\frac{8}{\pi} - 1 + \frac{13}{12\sqrt{3}}\right) \end{aligned}$$

where we used the relations

$$\begin{aligned} I_3\left(1,1,1;-\frac{1}{2},-\frac{1}{2},\frac{1}{2}\right) &= \frac{1}{2} I_3\left(1,1,1;-\frac{1}{2},-\frac{1}{2},-\frac{1}{2}\right) - I_2\left(1,2;-\frac{1}{2},0\right) \\ &= \frac{\pi^{3/2}}{12} - \frac{\pi^{1/2}}{2}\left(1 - \frac{1}{\sqrt{3}}\right) \end{aligned}$$

$$I_2\left(1,2;-\frac{1}{2},1\right) = \frac{1}{2} I_2\left(1,2;-\frac{1}{2},0\right) - \frac{1}{2} I_1\left(3;\frac{1}{2}\right) = \frac{\sqrt{\pi}}{4}\left(1 - \frac{4}{3\sqrt{3}}\right)$$

For B.58 we have

$$\begin{aligned} I_3\left(1,1,1;-\frac{1}{2},\frac{1}{2},\frac{1}{2}\right) &= 2\, I_3\left(1,1,1;\frac{1}{2},\frac{1}{2},\frac{1}{2}\right) - 2 I_2\left(2,1;1,\frac{1}{2}\right) \\ &= \frac{\pi^{3/2}}{24} - I_2\left(2,1;0,\frac{1}{2}\right) - I_1\left(3,\frac{3}{2}\right) \\ &= \frac{\pi^{3/2}}{24} - \frac{\pi^{1/2}}{6\sqrt{3}} \end{aligned}$$

where we used

$$I_2\left(2,1;0,\frac{1}{2}\right) = \frac{1}{2} I_1\left(3,\frac{1}{2}\right) = \frac{\pi^{1/2}}{12\sqrt{3}}$$

Finally

$$\mathbb{E}\left[u_2^4 \mathbf{1}_{(|u_2| \leqslant |u_1| \leqslant |u_3|)}\right] = \frac{1}{2} - \frac{4}{\pi} + \frac{13}{\pi 3\sqrt{3}}$$

$$\mathbb{E}\left[u_1^2\, u_2^2 \mathbf{1}_{(|u_1| \leqslant |u_2| \leqslant |u_3|)}\right] = \frac{1}{6} - \frac{2}{\pi 3\sqrt{3}}$$

$$\mathbb{E}\left[\min\left(|u_1|, |u_2|\right)^2 \min\left(|u_2|, |u_3|\right)^2\right] = \frac{5}{3} - \frac{8}{\pi} + \frac{6}{\pi\sqrt{3}}$$

Chapter 3

Price and Volatility Jumps in High Frequency Affine Modeling

3.1 Introduction

Affine modeling has a long history in finance, starting from the seminal papers of Vasicek (1977) and Cox et al. (1985) who first apply the Ornstein–Uhlenbeck and the square root process to interest rates. These stochastic processes are the fundamental building blocks of continuous time affine modeling which is an important tool in modern finance. The tractability characterizing this class of models has been uncovered in a number of relevant contributions including Heston (1993), Duffie and Kan (1996), Dai and Singleton (2000), Duffie et al. (2000), Duffie et al. (2003) and Filipović et al. (2013). The range of empirical applications is extremely wide, encompassing interest rates and credit risk, stock price returns, foreign exchange rates and intra-trade durations with a prominent role in asset pricing, portfolio optimization and risk management.

While the early studies on affine models have been conducted in continuous time, the recent literature highlights the superior flexibility achievable in the discrete time framework. Important contribution in this area are provided by Gourieroux and Jasiak (2006) introducing the discrete time version of the square root process named Autoregressive Gamma (ARG), Darolles et al. (2006) discussing the general properties of compound autoregressive processes (Car) including stationarity and ergodicity, Monfort and Pegoraro (2007) incorporating regime switching into a general multivariate affine framework with empirical applications to the term structure of interest rates and Monfort et al. (2015) introducing a variant of the ARG process (named ARG_0) that is designed to describe long periods of zero interest rates.

In this chapter we contribute to the extant literature on affine models defining a new stochastic process named Autoregressive Generalized Gamma process (*ARGG* henceforth) that is a generalization of the *ARG* and the ARG_0. The single additional parameter introduced in the conditional distribution allows to improve the control over the dispersion and the tail decay while preserving the same level of tractability of the *ARG*. Moreover, the new process can be directly related to an *ARG* (or ARG_0) where the state variable is affected by an idiosyncratic noise component or it is observed with an error. In the process allows to design also approximated filtering algorithms.

We propose an empirical application to stochastic volatility contributing to the large body of literature dedicated to this topic. Particularly our proposed models are designed to exploit the information available at high frequency. The idea to use the realized volatility within a GARCH framework to improve volatility forecasts dates to Engle (2002) and has been subsequently implemented by several authors including Shephard and Sheppard (2010), Hansen et al. (2012), Chen et al. (2011). The use of high frequency data to jointly model volatility and returns for option pricing purposes is a growing research area initiated by Stentoft (2008) who proposes a specification where the volatility is conditionally distributed as a normal inverse Gaussian. Importantly, closed form option pricing formulas are not available for his model, forcing to resort to computationally demanding Monte Carlo methods. The HARG model of Corsi et al. (2013) belongs instead to the affine framework but neglects the leverage effect, which is recovered in the HARGL specification but again at the expenses of the analytical tractability. Christoffersen et al. (2014) first include the leverage effect in a fully tractable way inspired to the affine GARCH of Heston and Nandi (2000). Notably, this active research area has largely benefited from the important developments that high frequency econometrics experienced in the last decade, building on the seminal work of Barndorff-Nielsen and Shephard (2004) and Barndorff-Nielsen and Shephard (2006) who first introduced a new econometric tool able to disentangle the continuous and the discontinuous component of asset price fluctuations. The remarkable advantage of these techniques is to make latent volatility and jumps directly observable.

The present work is related to the recent paper of Majewski et al. (2015), proposing a general affine option pricing framework which includes a wide class of discrete time models with multiple volatility and leverage components while a flexible pricing kernel allows to incorporate risk premia.

The model introduced by Majewski et al. is an improved version of the HARGL model of Corsi (2009) where the stochastic volatility follows an *ARG* process. Alitab et al. (2016) design a more advanced version (J-LHARG) where the volatility is also positively affected by price jumps. Our work shows that *ARG* specification fails to capture the violent volatility shocks observed between September 2007 and April 2014. We propose two new alternative affine models featuring a more flexible volatility dynamics. These models accommodate the well known volatility persistence and lagged as well as contemporaneous leverage effects. At the same time, they are able to account for additional volatility dispersion thanks to the flexibility of the *ARGG* process. More precisely, we extend the LHARG model of Majewski et al. in three distinct ways: i) we move from the *ARG* to the *ARGG* process to describe stochastic volatility, ii) we introduce a second volatility factor featuring volatility jumps, iii) we introduce a negative skewness at the daily time horizon. Though option pricing is fully tractable within our framework, the empirical analysis of this chapter is developed exclusively under the physical measure. Interesting implications to option pricing are left for future research.

Our results, obtained analyzing 5 minutes returns from the S&P 500 from September 2007 to April 2014, show a remarkable improvement of the fit moving from the standard *ARG* process to its generalized version. Importantly, according to the empirical analysis developed in Chapter 1, the peaks of realized quadratic variation observed during the sub-prime crisis are not generated by price jumps. This calls for the inclusion of a second factor able to generate sudden volatility shocks that will be referred as volatility jumps. The need for multiple volatility factors is widely recognized in the literature, some examples are Engle and Lee (1999), Bollerslev and Zhou (2002), Christoffersen et al. (2008), Adrian and Rosenberg (2008), Chernov et al. (2003) among many others. The occurrence of volatility jumps in the stock market is also largely supported by many empirical studies such as Duffie et al. (2000), Bates (2000), Pan (2002), Eraker et al. (2003), Eraker (2004), Bandi and Renò (2016), Corsi and Renò (2012). Non-parametric evidence has been also provided by Jacod and Todorov (2010), Todorov and Tauchen (2011) reporting the simultaneous occurrence of price and volatility jumps. It has been pointed out that jumps in returns and jumps in volatility have a different role in the description of asset returns. Price jumps generate large and rare sudden movements which would otherwise require an extremely high and unrealistic level of volatility to be

justified. Volatility jumps generate instead violent changes in the level of volatility with a persist effect on the size of returns. We find the latter to dominate at the end of 2008 during the period leading to the Leheman bankruptcy.

The inclusion of the contemporaneous leverage is another important innovation in discrete time modeling. The relevance of the leverage effect is well known since Black (1976). The studies of Andersen et al. (2007a) and Andersen et al. (2010a) point out that, even when the realized volatility is accurately measured from high frequency intraday returns, the standardized log-returns purified by the effect of price jumps are still deviating from the normal distribution. This is due to the pronounced asymmetric relationship between return and volatility innovations. We distinguish between two different types of leverage: the contemporaneous leverage generating distributional asymmetries at a short time horizons (say a single trading day) and the lagged leverage stemming from the positive effect of negative returns on future volatility, which reinforce the asymmetries on longer time horizons. The models proposed by the literature usually fail to capture simultaneously both of these features, whose relevance is stressed for instance by Bollerslev et al. (2006). Continuous time models instead are usually designed to feature contemporaneous leverage but often not able to account for its persistence on long horizons; in discrete time models instead leverage is usually introduced with a time lag allowing negative returns to positively affect future volatility, this is the case for example in Heston and Nandi (2000), Christoffersen et al. (2008), Christoffersen et al. (2012), Christoffersen et al. (2015), Majewski et al. (2015), Alitab et al. (2016) where distributional asymmetries at the daily time scale can be generated only by price jumps. The new specification that we propose is able to accommodate both contemporaneous and lagged leverage by introducing a random drift component which depends linearly on volatility. At the same time the volatility risk premium is assumed to be proportional to the expected volatility according to the information available at the end previous trading day. Consistently with Bollerslev et al. (2006), Corsi and Renò (2012), Majewski et al. (2015) and Alitab et al. (2016), we find that the simultaneous and the legged leverage are both statistically significant, even after taking into account the effect of price jumps.

The rest of this chapter is structured as follows: Section 3.2 introduces the generalized non-central gamma distribution and some of its properties, Section 3.3 is dedicated to the *ARGG* process, in

Section 3.4 we develop the empirical application to stochastic volatility and Section 3.5 concludes. Technical aspects are relegated to the appendices.

3.2 The Generalized Non-Central Gamma Distribution

In this section we introduce a new distribution that generalizes the non-central gamma. The latter has a prominent role in discrete time affine models, mostly determined by its direct link with the square root process in continuous time[1]: the autoregressive gamma process of Gourieroux and Jasiak (2006) is the discrete time counterpart of the CIR process Cox et al. (1985). This distribution can be defined as a Poisson mixture of gammas with the same scale and different shape parameters. Its functional form is fully characterized by three free parameters: the shape, the scale and the non-centrality. The p.d.f. and the m.g.f. are respectively

$$\bar{\gamma}\left(x;\,\delta,\theta,\lambda\right) = e^{-\lambda}\sum_{n=0}^{\infty}\frac{\lambda^{n}}{n!}\gamma\left(x;\,\delta+n,\,\theta\right) \tag{3.1}$$

$$\hat{\bar{\gamma}}\left(z;\,\delta,\theta,\lambda\right) = \left(1-\theta z\right)^{-\delta}\exp\left(\lambda\frac{\theta z}{1-\theta z}\right) \tag{3.2}$$

the mean and variance are

$$\mathbb{E}\left(X\right) = \theta\left(\delta+\lambda\right) \tag{3.3}$$

$$Var\left(X\right) = \theta^{2}\left(\delta+2\lambda\right) \tag{3.4}$$

In time series modeling δ and θ are usually kept constant while non-centrality is a time-varying state variable. Importantly, the mean and the variance are linear affine in the non-centrality. To achieve more flexibility in empirical applications it can be useful to control the four affine coefficients independently while in equations 3.6 and 7 they are determined by three independent parameters. There are different ways to introduce additional flexibility: one possibility could be for instance to assume that δ depends λ through a linear affine relation without substantially changing the features of the distribution. Here we choose a different approach working directly on the mixing distribution with the key to have additional control on the tail decay of the density as it will be shown later.

[1] It is well known that the conditional distribution of a continuous random variable following a square root process and observed at any fixed time horizon is a non-central gamma.

An important characteristic of the Poisson distribution is the equivalence between its mean and its variance: the impossibility to set the first and the second moment independently is in fact often regarded as a major limitation of Poisson models in empirical applications. We propose a simple extension of the non-central gamma density based on the same idea introduced in the contest of Poisson regression models to account for over-dispersion (see Hilbe 2011), i.e. we consider the a negative binomial mixing distribution. To define the new density let us consider a negative binomial random variable Z with shape parameter $r > 0$ and probability of success $p \in [0, 1)$ and probability of failure $q = 1-p$. In the limit $q \to 0$ and $r \to \infty$ with rq constant, the mixing distribution converges weakly to a Poisson, it is therefore convenient to adopt the parametrization $r = \lambda (1 - q) / q$.

Definition 1. Let X be a non negative random variable. X follows a generalized non-central gamma distribution with non-centrality $\lambda \geqslant 0$, over-dispersion $q \in [0, 1)$, shape $\delta > 0$ and scale $\theta > 0$ denoted by $\psi (x;\, \delta, \theta, \lambda, q)$ if its distribution conditionally on Z is a gamma with shape parameter $\delta + Z$ and scale parameter θ

$$X | \, Z \sim \gamma (\delta + Z, \theta) \tag{3.5}$$

The p.d.f. and the moment generating function, defined for $Re (z) < (1 - q) / \theta$, are respectively

$$\psi (x;\, \delta, \theta, \lambda, q) = \begin{cases} \sum_{n=0}^{\infty} nb (n; r (\lambda, q), q) \, \gamma (x;\, \delta + n,\, \theta) & \text{if } q \in (0, 1) \\ e^{-\lambda} \sum_{n=0}^{\infty} \frac{\lambda^n}{n!} \gamma (x;\, \delta + n,\, \theta) & \text{if } q = 0 \end{cases} \tag{3.6}$$

where $r (\lambda, q) = \lambda (1 - q) / q$, $nb (n; r, q) = \dfrac{\Gamma (r + n)}{\Gamma (r)\, n!} q^n (1 - q)^r$ and $\gamma (x;\, \nu,\, \theta) = \dfrac{e^{-x/\theta} x^{\nu - 1}}{\theta^{\nu} \Gamma (\nu)}$

Proposition 7. The moment generating function associated to the generalized non-central gamma distribution is

$$\hat{\psi} (z;\, \delta, \theta, \lambda, q) = \begin{cases} (1 - \theta z)^{-\delta} \left(1 - \dfrac{q}{1 - q} \dfrac{\theta z}{1 - \theta z} \right)^{-r(\lambda, q)} & \text{if } q \in (0, 1) \\ (1 - \theta z)^{-\delta} \exp \left(\lambda \dfrac{\theta z}{1 - \theta z} \right) & \text{if } q = 0 \end{cases} \tag{3.7}$$

defined for $Re (z) < (1 - q) / \theta$

Proof: For $q = 0$ the distribution is a non-central gamma, we focus therefore on the more interesting case $q > 0$

$$\begin{aligned}\mathbb{E}\left[e^{zX}\right] &= \sum_{n=0}^{\infty} nb\left(n; r\left(\lambda\right), q\right)\left(1-\theta z\right)^{-(\delta+n)} \\ &= \left(1-\theta z\right)^{-\delta}\left[\frac{1-q}{1-q/\left(1-\theta z\right)}\right]^{r(\lambda,q)}\end{aligned}$$

The Generalized Gamma-zero distribution

The gamma-zero distribution introduced by Monfort et al. (2015) is obtained from the non-central gamma when the shape parameter is zero. Such a distribution has a point mass at zero which makes its use particularly attractive to model non-negative random variables that can stay a zero for prolonged periods of time (e.g. interest rates). The generalized gamma-zero distribution can be analogously obtained from the generalized non-central gamma considering a zero shape parameter δ.

Definition 2. The non negative random variable X follows a generalized gamma-zero distribution with parameters λ, q and θ denoted by $\psi_0\left(x; \theta, \lambda, q\right)$ if its distribution conditionally on Z is a gamma with shape parameter Z and scale parameter θ

$$X|\, Z \sim \gamma\left(Z, \theta\right) \tag{3.8}$$

The p.d.f. is

$$\psi_0\left(x;\, \theta, \lambda, q\right) = \begin{cases} \left(1-q\right)^{r(\lambda)} \mathbb{1}_{\{x=0\}} + \sum_{n=0}^{\infty} nb\left(n; r, q\right)\gamma\left(x;\, n,\, \theta\right) & \text{if } q \in (0,1) \\ e^{-\lambda}\left[\mathbb{1}_{\{x=0\}} + \sum_{n=0}^{\infty} \frac{\lambda^n}{n!}\gamma\left(x;\, n,\, \theta\right)\right] & \text{if } q = 0 \end{cases} \tag{3.9}$$

Proposition 8. The moment generating function of the non-central generalized gamma-zero distribution is

$$\hat{\psi}_0(z;\theta,\lambda,q) = \begin{cases} \left(1 - \frac{q}{1-q}\frac{\theta z}{1-\theta z}\right)^{-r(\lambda,q)} & \text{if } q \in (0,1) \\ \exp\left(\lambda \frac{\theta z}{1-\theta z}\right) & \text{if } q = 0 \end{cases} \tag{3.10}$$

Proof: Analogous to proposition 3.7 with the only difference that $\delta = 0$.

Note that this distribution has a point mass located at $x = 0$ with an associated probability of $(1-p)^{r(\lambda,q)}$ (this it can be immediately checked taking the limit $z \to \infty$ in 3.10).

Some basic properties

Alternative representation The negative binomial distribution can be regarded as a Poisson with random intensity, hence the following proposition provides an alternative definition for the generalized non-central gamma

Proposition 9. Consider the non-negative random variable Λ distributed as a gamma with scale $q/(1-q)$ and shape $r = \lambda(1-q)/q$ with $\lambda > 0$, $q \in (0,1)$. X follows a generalized non-central gamma distribution with parameters $\delta \geqslant 0$, λ, q, $\theta > 0$ when its distribution conditionally on Λ is a non-central gamma with shape δ, non-centrality Λ and scale θ.

Proof: The equivalence between this definition and definitions 1 and 2 can be immediately verified just observing that the moment generating function that arises from 9 coincides with expression 3.7. To recover the Poisson distribution we can consider the limit $q \to 0$ such that $r(\lambda,q) \to \infty$ and $r(\lambda,q)\, q \to \lambda$, thus the gamma degenerates to a Dirac delta with the whole probability mass concentrated at λ.

According to Proposition 9 the generalized non-central gamma admits a simple representation similar to the non-central gamma and descending directly from the gamma mixture interpretation. Such a representation is particularly useful in numerical simulations:

$$X = \varepsilon + \sum_{j=1}^{Z} W_j \tag{3.11}$$

where $\varepsilon \sim \gamma(\delta,\theta)$, $W_j \sim \gamma(1,\theta)$ and Z is extracted from a negative binomial distribution with

parameters $r(\lambda, q)$ and q or equivalently from a Poisson distribution with random non-centrality $\Lambda \sim \gamma(r(\lambda, q), q/(1-q))$. The generalized gamma-zero admits the same representation but with the noise term ε identically null.

Mean and variance The mean and the variance of the generalized non-central gamma distribution are

$$\mathbb{E}(X) = \theta(\delta + \lambda) \tag{3.12}$$

$$Var(X) = \theta^2\left(\delta + \lambda\frac{2-q}{1-q}\right) \tag{3.13}$$

notably, the parameter q affects only the variance of the distribution generating over-dispersion with respect to the non-central gamma that is obtained for $q = 0$. Now the coefficients defining the affine dependence of the mean and the variance from the non-centrality λ can be controlled independently through four distinct parameters.

Tail decay We see from equations 3.7 and 3.10 that the moment generating function has a singularity at $z = (1-q)/\theta$, suggesting that the tail for x is proportional to $e^{-(1-q)/\theta}$ as it is stated by the following proposition:

Proposition 10. As $x \to \infty$, the rate of decay of $\psi(x; \delta, \theta, \lambda, q)$ is $e^{-x(1-q)/\theta}(x/\theta)^{r-1}$.

Proof: See Appendix C.1.1.

Additional properties of the generalized non-central gamma are discussed in Appendix C.1.

Numerical computation of the density

The generalized non-central gamma density can be computed numerically at least in two distinct ways: through the explicit summation of the series in expression 3.6 and by Fourier inversion. The first method is convenient for small values of the argument and requires to control the truncation error. The second is efficient for medium and large values of the argument but it is computationally more complex because requires the simultaneous control of three sources of error: the truncation, the discretization and the round off. Technical details are discussed in Appendix C.3. Figure 3.1

shows the different shape of the density depending on the choice of the parameter p and highlights the way it affects the tail decay.

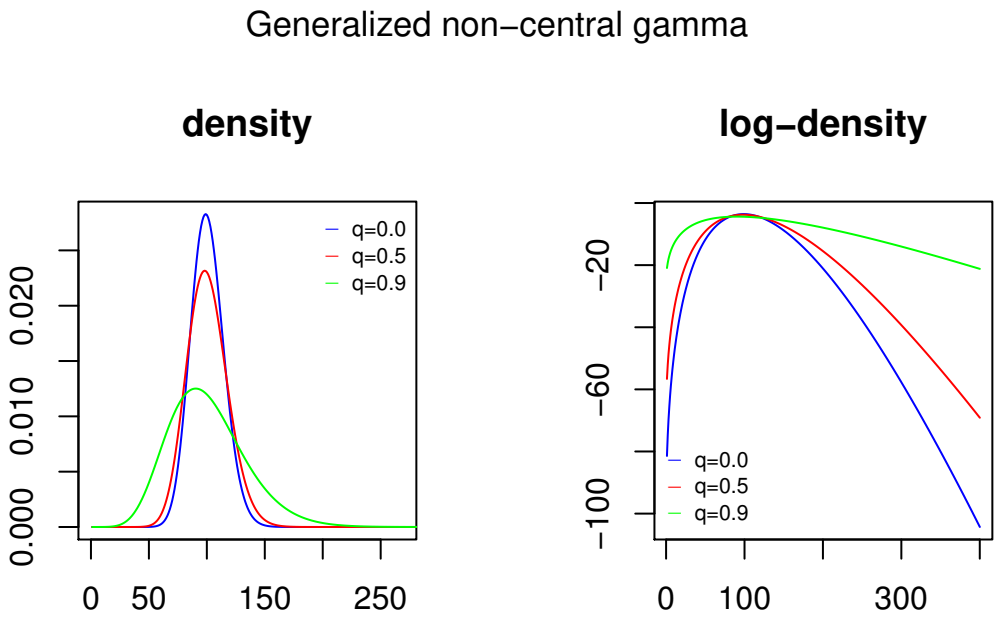


FIGURE3.1: Density and log-density of the generalized non-central gamma with parameters $\delta = 5$, $\lambda = 100$, $\theta = 1$

3.3 The Autoregressive Generalized Gamma Process

The Autoregressive Gamma process of Gourieroux and Jasiak (2006) has multiple applications in financial econometrics like the description of inter-quote durations, interest rate and volatility modeling. The extensive use of this process in various research areas is largely due to its affine properties which allow to perform asset pricing in a tractable way and to calculate the conditional moment generating function at each time horizon in a semi-analytical form. Here we introduce a simple generalization of the *ARG* processes based on the Generalized Non-Central Gamma distribution and therefore named *ARGG* process. Importantly the *ARGG* belongs to the class of compound autoregressive processes (Car) whose properties are extensively discussed by Darolles et al. (2006).

Definition 3. The process (Y_t) is an autoregressive generalized gamma process of order p ($ARGG_\delta(p)$)

when the conditional distribution of Y_t given Y_{t-1} is $\psi\left(\delta, \theta, d + \beta' \underline{Y_{t-1}}, q\right)$, where $\delta > 0$, $q \in (0,1)$, $\beta, d \geqslant 0$, $\theta > 0$ and $\underline{Y_t} = (Y_t, \ldots, Y_{t-p+1})'$.

The definition can be naturally extended to the generalized gamma-zero process if we allow δ to take the value zero provided $d > 0$ to prevent the lower bound from becoming absorbing. Following the interpretation based on a mixture of gamma distributions, the $ARGG_\delta(p)$ process can be represented (and simulated) as follows:

$$Y_t = \sum_{j=1}^{Z_t} W_{j,t} + \epsilon_t$$

where $\varepsilon \sim \gamma(\delta, \theta)$, $W_j \sim \gamma(1, \theta)$ and Z_t is extracted from a Poisson distribution with a conditionally stochastic non-centrality $\Lambda_t \sim \gamma\left(r_t, q/(1-q)\right)$ being $r_t = \left(d + \beta' \underline{Y_{t-1}}\right)(1-q)/q$. The $ARGG_0(p)$ process is obtained setting $\epsilon_t = 0$, while the $ARG_\delta(p)$ is characterized by a deterministic non-centrality $\Lambda_t = \lambda_t$ corresponding to the limiting case $q \to 0$. Note that the $ARGG_\delta(p)$ process is a Car process of order p because its conditional Laplace transform is an exponential affine function of the state variables:

$$\mathbb{E}\left[\exp(z\, Y_t) \middle| \underline{Y_{t-1}}\right] = \exp\left\{\sum_{i=1}^{p} a_i(z)\, Y_{t-i} + b(z)\right\} \quad z < (q-1)/\theta$$

where

$$a_i(z) = -\beta_i \frac{1-q}{q} \log\left(1 - \frac{q}{1-q}\frac{\theta z}{1-\theta z}\right) \tag{3.14}$$

$$b(z) = -\delta \log(1-\theta z) - d\frac{1-q}{q} \log\left(1 - \frac{q}{1-q}\frac{\theta z}{1-\theta z}\right) \tag{3.15}$$

From Darolles et al. (2006) proposition 2, $\underline{Y_t}$ follows a multivariate vector Car process of order 1 with moment generating function

$$\mathbb{E}\left[\exp\left(\mathbf{z}'\underline{Y_t}\right) \middle| \underline{Y_{t-1}}\right] = \exp\left[A(\mathbf{z})'\, \underline{Y_{t-1}} + B(\mathbf{z})\right]$$

where

$$A(\mathbf{z}) = (a_1(z_1) + z_2, \ldots, a_{p-1}(z_1) + z_p, a_p(z_1))' \qquad B(\mathbf{z}) = b(z_1)$$

3.3.1 Stationarity and Ergodicity

Following Darolles et al. (2006), the stationarity condition is determined by the multi-dimensional function $A(\mathbf{z})$ through the condition $\lim_{h\to\infty} A^{\circ h} = 0$ for all $\mathbf{z} \in \mathbb{R}^p$ with non-negative components, where $A^{\circ h}$ denotes the function A compounded h times.

Proposition 11. Under the condition $\theta \sum_{i=1}^{q} \beta_i < 1$, the $ARGG_\delta(p)$ process is stationary and ergodic.

Proof: See Appendix C.2.1

3.3.2 The $ARGG(1)$ process

We focus on the simplest case $p = 1$ to highlight the main features of the process. Assume δ, β, $\theta > 0$, $d \geqslant 0$ and

$$Y_t | Y_{t-1} \sim \psi\left(\delta, \theta, d + \beta \underline{Y_{t-1}}, q\right)$$

Conditional mean and variance

A let $\rho = \theta\beta$, the conditional mean and variance one time interval ahead are

$$\mathbb{E}(Y_t | Y_{t-1}) = \theta(\delta + d) + \rho Y_{t-1} \tag{3.16}$$

$$Var(Y_t | Y_{t-1}) = \theta^2\left(\delta + \frac{2-q}{1-q}d\right) + \theta\rho\frac{2-q}{1-q}Y_{t-1} \tag{3.17}$$

more generally, the following result holds

Proposition 12. The first and the second moment of the $ARGG(1)$ with $\rho = \theta\beta < 1$ and $q \in (0,1)$conditionally on lagged values of the state variable Y are given by the following expressions

$$\mathbb{E}(Y_t | Y_{t-h}) = \theta(\delta + d)\frac{1-\rho^h}{1-\rho} + \rho^h Y_{t-h} \tag{3.18}$$

$$\begin{aligned} Var(Y_t | Y_{t-h}) = {} & \theta^2\left\{\frac{\delta+d}{1-\rho}\frac{2-q}{1-q}\left[\frac{1-\rho^h}{1-\rho} - \frac{1-\rho^{2h}}{1-\rho^2}\right]\right. \\ & \left. + \left(\delta + \frac{2-q}{1-q}d\right)\frac{1-\rho^{2h}}{1-\rho^2}\right\} + \theta\frac{2-q}{1-q}\rho^h\frac{1-\rho^h}{1-\rho}Y_{t-h} \end{aligned} \tag{3.19}$$

the unconditional mean and variance are

$$\begin{aligned}\mathbb{E}(Y_t) &= \theta\frac{\delta+d}{1-\rho}\\ Var(Y_t) &= \frac{\theta^2}{1-\rho^2}\left[\delta+\frac{2-q}{1-q}\left(d+\rho\frac{\delta+d}{1-\rho}\right)\right]\end{aligned}$$

Proof: See AppendixC.2.2.

The weak $AR(1)$ interpretation allows to represent the model in the form

$$Y_t = \theta(\delta+d)+\rho Y_{t-1}+\epsilon_t$$

where ϵ_t is a martingale difference. Thus, not surprisingly, the first order autocorrelogram is the same of the $ARG(1)$ process

$$Corr(Y_t, Y_{t-h}) = \rho^h$$

Long memory

Like the $ARG(1)$ model, also the $ARGG(1)$ is able to feature long memory when $\rho = 1$. Assume for simplicity $\theta = 1$, for $z \leqslant 0$ we have

$$0 \leqslant a_q(z) = \frac{1-q}{q}\log\left(1-\frac{q}{1-q}\frac{z}{1-z}\right) \leqslant a_0(z) \quad \forall z > q-1$$

where $a_0(z) = z/(1-z)$ corresponding to the $ARG(1)$ process. Thus $a_q^{\circ h}(z)$ converges to zero at least the same rate $1/h$ of $a_0^{\circ h}(z)$

$$0 \leqslant a_q^{\circ h}(z) \leqslant a_0^{\circ h}(z) = \frac{z}{1-hz} \to 0 \quad \text{as } h \to \infty$$

that implies weak ergodicity according to Darolles et al. (2006). Moreover

$$\begin{aligned}a_q(z) &= \frac{-z}{1-z}-\frac{1}{2}\frac{q}{1-q}\left(\frac{z}{1-z}\right)^2+O\left(z^3\right)\\ &= -z/\left[1-z\left(1+\frac{1}{2}\frac{q}{1-q}\right)\right]+O\left(z^3\right)\end{aligned}$$

thus the rate of convergence is exactly $1/h$.

Approximated filtering

It ha been pointed out that the *ARGG* process can be regarded as an *ARG* process with a stochastic non-centrality that is conditionally distributed as a gamma. In practice this situation can materialize when the state variable is not directly observed. In this case the approximated filtering algorithm proposed by Bates (2006) for latent affine processes can be applied. This technique generalizes the Kalman Filter using at each step an approximating conditional distribution that is possibly different from the Gaussian. Such a procedure is well suited for uni-dimensional filtering problems. To provide a concrete example we describe the algorithm applied to the famous Heston model where the unobserved volatility follows a square root process. We show that in this case the filtered process for latent volatility is analogous $ARGG\,(1)$ with the only different that the overdispersion parameter is time varying. The Heston model is defined trough the following equations:

$$\begin{aligned} dy_t &= \mu_0 + \left(\mu_1 - \frac{1}{2}\right) V_t + \sqrt{V_t}\left(\rho dW_{1,t} + \sqrt{1-\rho^2} dW_{2,t}\right) \\ dV_t &= \alpha - \beta V_t + \sigma\sqrt{V_t} dW_{2,t} \end{aligned}$$

where y_t is the log-price, V_t is the instantaneous volatility and $W_{1,t}$, $W_{2,t}$ are independent Wiener processes. The joint conditional density of y and V has an exponential affine characteristic function

$$\hat{f}\left(\imath u, \imath z\right) = \mathbb{E}\left[\exp\left(\imath y_{t+h} u + \imath V_{t+h} z\right) \middle| y_t, V_t\right] = \exp\left[\imath y_t u + A\left(u, z; h\right) + V_t B_V\left(u, z; h\right)\right]$$

the functions A and B_V can be calculated in closed form (see Heston 1993 for details). Given that V_t follows a CIR process, the marginal distribution of V_{t+h} conditional on V_t in this model is a non central gamma with shape $2\alpha/\sigma^2$, non centrality $V_t \cdot 2\beta e^{-\beta h}/\sigma^2\left(1 - e^{-\beta h}\right)$ and scale $\left(1 - e^{-\beta h}\right)\sigma^2/2\beta$ (see Alfonsi 2010).

Assume that the price is observed at $t = 1, 2, \ldots T$, therefore $h = 1$. To describe the recursive algorithm we detail a single step starting from time $t-1$ where the density g_{t-1} of the unobserved state variable V_{t-1} conditional on the information available at time $t-1$ is assumed to be known. Let $\phi_{t|t-1}$ be the joint density of y_t and V_t conditional on the information at $t-1$ whose characteristic

function is

$$\begin{aligned}\phi_{t|t-1}(\imath u, \imath z) &= \mathbb{E}\left[\exp(\imath y_t u + \imath V_t z)|\,\mathcal{F}_t\right] \\ &= \mathbb{E}\left[\exp\left[A(u,z) + V_t B_V(u,z)\right]|\,\mathcal{F}_t\right] \\ &= \exp\left[\imath y_t u + A(u,z)\right]\hat{g}_{t-1}\left(B_V(u,z)\right)\end{aligned}$$

From proposition 1 in Bates (2006) the Bayesian updating gives

$$\hat{g}_t(z; y_t) = \frac{\int_{-\infty}^{+\infty}\phi_{t|t-1}(\imath u, \imath z)\, e^{-\imath u y_t} du}{\int_{-\infty}^{+\infty}\phi_{t|t-1}(\imath u, 0)\, e^{-\imath u y_t} du}$$

the idea it to approximate this conditional density to circumvent the complications related to the computation of the exact filtered density. The approximating distribution proposed by Bates for non negative latent variables is a gamma matching the conditional mean and variance

$$m_t(y_{t+1}) = \left.\frac{\partial \hat{g}_t}{\partial z}\right|_{z=0} \qquad v_t(y_t) = \left.\frac{\partial^2 \hat{g}_t}{\partial z^2}\right|_{z=0} - m_t(y_t)^2$$

$$\hat{g}_t(z; y_t) \simeq (1 - \vartheta_t z)^{-\delta_t}$$

where $\vartheta_t = v_t(y_t)/m_t(y_t)$ and $\delta_t = m_t(y_t)^2/v_t(y_t)$. Bates demonstrates that the errors affecting the conditional moments due to this approximation die out geometrically. Under this method, the marginal distribution of V conditional on the information available at the previous time is a generalized non central gamma with shape $2\alpha/\sigma^2$, and over-dispersion parameter $q_t = \Theta_t/(\Theta_t + 1)$ where $\Theta_t = 2\vartheta_t \beta e^{-\beta h}/\sigma^2\left(1 - e^{-\beta h}\right)$ and non-centrality $\lambda_t = \delta_t/\Theta_t$. Note that q_t and λ_t are time dependent non-linear functions of the observed returns capturing respectively the uncertainty and the expected level of the unobserved state variable.

3.4 Modeling S&P 500 Returns and Volatility

In this section we discuss an empirical application of the *ARGG* process to to stochastic volatility. The increasing availability of high frequency data and the advances in high frequency econometrics, stimulated in the last years the development of a new class of stochastic volatility models able to exploit the large amount of intraday information to improve the accuracy of volatility forecasts

(Andersen et al. 2007b, Corsi et al. 2010, Corsi and Renò 2012 among others) and the option pricing performances (Stentoft 2008,Corsi et al. 2013, Alitab et al. 2016, Bormetti et al. 2016). Following the idea of Andersen et al. (2011) among others, we distinguish three components of close-to-close returns and variance: an intraday continuous component, a price jump component and an overnight component. Paralleling many other contributions in this area of research, we assume that the integrated volatility is observed without error and jumps are exactly identified. These simplifying assumptions offer the considerable advantage of making the latent volatility directly observable and lead to a substantial simplification of the statistical inference. Some studies deviating from this assumptions are anyway available in the literature: Christoffersen et al. (2015) consider a specific form of the measurement error leading to a straightforward filtering while Bormetti et al. (2016) rely on Markov Chain Monte Carlo methods for the statistical inference.

This section is structured as follows: in Subsection 3.4.1 we describe our data set discussing the salient stylized facts observed on the three return and volatility components, in Subsection 3.4.2 we propose a model for close-to-close returns obtained as the sum of continuous and discontinuous price changes, in Subsection 3.4.3 we complete the model defining three alternative dynamic specifications for stochastic volatility with increasing complexity: the first specification is the LHARG model of Majewski et al. (2015) representing as a benchmark, the second is a simple generalization where the standard *ARG* process is replaced by the *ARGG* while in the third model we include also an additional volatility factor to feature volatility jumps. In Subsection 3.4.4 we discuss our empirical results. Importantly, though our modeling framework is fully affine and allows the development of tractable option pricing formulas, the present analysis is developed exclusively under the physical measure. Option pricing application are left for future research.

3.4.1 Data and Stylized Facts

The data set used in this section is a part of the data set described in Section 1.2 of Chapter 1. It contains intraday quotes from the S&P 500 reported every 5 minutes from 2007-9-14 to 2014-04-30, for a total number of 1674 trading days. This sample includes multiple high volatility periods like the sub-prime crisis leading to the bankruptcy of Leheman Brothers on September 15^{th} 2008 and the European Sovereign crisis of 2011. For each trading day, the first quote is reported 5 minutes after

the market opening time, thus the first intraday return is incorporated in the overnight return[2]. Let S_t denote the closing price of the index, then $y_t = \log(S_t/S_{t-1})$ is the close-to-close log-return on day t. Intraday log-returns are indicated as $r_{i,t} = \log(S_{i,t}/S_{i-1,t})$ with $i = 1, \dots, M$ (in our case $M = 77$). Open to close log-returns are obtained as the sum of intraday returns: $y_t^{id} = \sum_{i=1}^{M} r_{i,t}$. The closing price reported at the end of each trading day, usually differs from the opening price of the following trading day due to the information accumulated overnight. The overnight return between t and $t+1$ is computed over the last reported quote on day t and the first of the day after. The cumulated close-to-close log-return and the intraday realized quadratic variation are plotted in Figure 3.2 while Table 3.1 reports some summary statistics. We observe that the realized variation is subject to extreme shocks during the sub-prime crisis. The highest level of intraday price variability is measured on 10^{th} of October 2008, few days before the Leheman bankruptcy, where the quadratic variation is 47 times larger than its sample mean and roughly 3.5 times larger that the other peaks observed in correspondence of the flash crash on 2010-5-6 and during the Euro sovereign crisis. The quadratic variation measured on 2008-10-8, 2010-5-6 and 2011-8-9 correspond to an annualized volatility of 122%, 64.2% and 63.1% respectively. The largest close-to-close market decline of -11.2% is observed on 8^{th} of August 2011 during the Euro Sovereign crisis.

We decompose the total close to close variation into three distinct components: the intraday continuous price variation IV_t also referred as integrated volatility, a jump component RV_t^J and the overnight variation ON_t equal to the squared overnight return. The total intraday variation RV_t is the sum of IV_t and RV_t^J which can be disentangled using high the frequency methods largely discussed in Chapters 1 and 2. Particularly we apply the sequential version of the test proposed by Corsi, Pirino, and Renò (2010) that is detailed in Section 1.3. This method allows not only to detect days with at least one jump but also to identify its exact intraday time and its size[3]. We can therefore decompose the close to close return as $y_t = y_t^c + J_t$ where the overnight return is assumed for convenience to be included in the continuous component. Possible measures of the

[2] This practice is very common in high frequency literature: the first 5-10 minutes of trading are often excluded due to the erratic price behavior of prices related to the opening procedures.

[3] Note that the test of Lee and Mykland (2008) and its variants are designed exactly to this purpose but they require a smooth evolution of volatility to be consistently applied, thus they are not well suited in presence of volatility jumps. This is why we prefer to use the sequential CPR method.

relative contribution of the three components are the following:

$$\frac{\sum_t IV_t}{\sum_t RV_t + ON_t} = 63.8\% \qquad \frac{\sum_t ON_t}{\sum_t RV_t + ON_t} = 34.4\% \qquad \frac{\sum_t RV_t^J}{\sum_t RV_t + ON_t} = 1.8\%$$

Figure 3.3 shows the evolution of the overnight and the jump contribution across time, computing the ratios above on a two months rolling window. The relative weight of the overnight returns seems non constant in time and ranges from 20% to almost 70%. Interestingly, the overnight contribution starts to increase at the beginning of the Euro sovereign crisis in late 2009, suggesting that the rise is possibly induced by volatility spillovers from the European markets which start operating before the NYSE opening time.

The role of price jumps

The overall contribution of jumps is very small compared to the remaining components. Nevertheless, in the rare cases when jumps occur (approximately once every 8 days) their contribution to the intraday realized quadratic variation ranges between 6.3% and 90.5% with an average value of 23.5%. In Chapter 1 we stressed that the role of jumps appears less relevant during periods of market turmoil when the continuous return variability seems to dominate. In principle one can argue that the decreasing relative contribution of price jumps is determined by a decline of the power of the test due to the high level of volatility and to its extreme variability in time. To provide a more robust evidence, Figure 1.7 reported in Chapter 1 shows the trend of alternative measures for the relative jump contribution in line with Huang and Tauchen (2005):

$$RJ_t = \frac{RV_t - \hat{IV}_t}{RV_t} \tag{3.20}$$

where $\hat{IV}_t$ is a jump robust realized volatility estimator (results are reported for the corrected threshold bipower, the $MinRV$ and the $MedRV$). It is known that the finite sample bias generated by a rapidly changing volatility leads to underestimate IV_t and therefore to overestimate RJ_t. In spite of this, the bottom panel in Figure 1.7 shows a clear decreasing value of RJ_t during the subprime crisis. Concerning instead the inverse dependence between the volatility level and the jump activity that has been observed and largely discussed in Chapter 1, it is more difficult to draw

sound conclusions. As a simple example, consider the case of i.i.d. jumps: the power of the tests is naturally diminishing when the volatility is higher, because disentangling jumps from continuous diffusion becomes more difficult since the size of a jump becomes on average closer to the size of a 5 minutes continuous return. In this simple example we may observe a pattern of RJ_t compatible or at least qualitatively similar to the empirical one.

Andersen et al. (2007a) and Andersen et al. (2010a) point out that daily returns properly standardized by the realized volatility measured at high frequency, may still significantly deviate from normality. Two main determinants of such deviations are identified: the presence of price jumps and the contemporaneous leverage effect. From Table 3.1 we see that the skewness of close-to-close returns reduces marginally when jumps are removed. Analogously, the simultaneous correlation coefficients between close-to-close returns y_t and the change in the total realized variation $RV_t - RV_{t-1}$ moves from -16% to -15% removing jumps. This suggests that the contemporaneous negative dependence between price and volatility is the major determinant of distributional asymmetries characterizing daily returns. Importantly this result is confirmed also under different jump identification methods.

Summary Statistics						
	mean	standard deviation	max	min	skewness	kurtosis
RV_t	$1.23 \cdot 10^{-4}$	$2.95 \cdot 10^{-4}$	$5.78 \cdot 10^{-3}$	$1.25 \cdot 10^{-5}$	8.56	111.0
IV_t	$1.20 \cdot 10^{-4}$	$2.89 \cdot 10^{-4}$	$5.78 \cdot 10^{-3}$	$1.25 \cdot 10^{-5}$	8.85	118.7
RJ_t	2.3%	8.5%	90.5%	0.0%	4.63	25.8
y_t	0.022%	1.49%	10.7%	-11.7%	-0.44	9.7
y_t^c	0.027%	1.46%	10.7%	-11.7%	-0.38	10.2
J_t	-0.04%	0.6%	1.6%	-2.6%	-1.25	4.08
N_t	0.12	0.39	5	0	4.35	27.4

Table 3.1: Summary statistics of log returns and realized variance. N_t is the number of jumps per day.

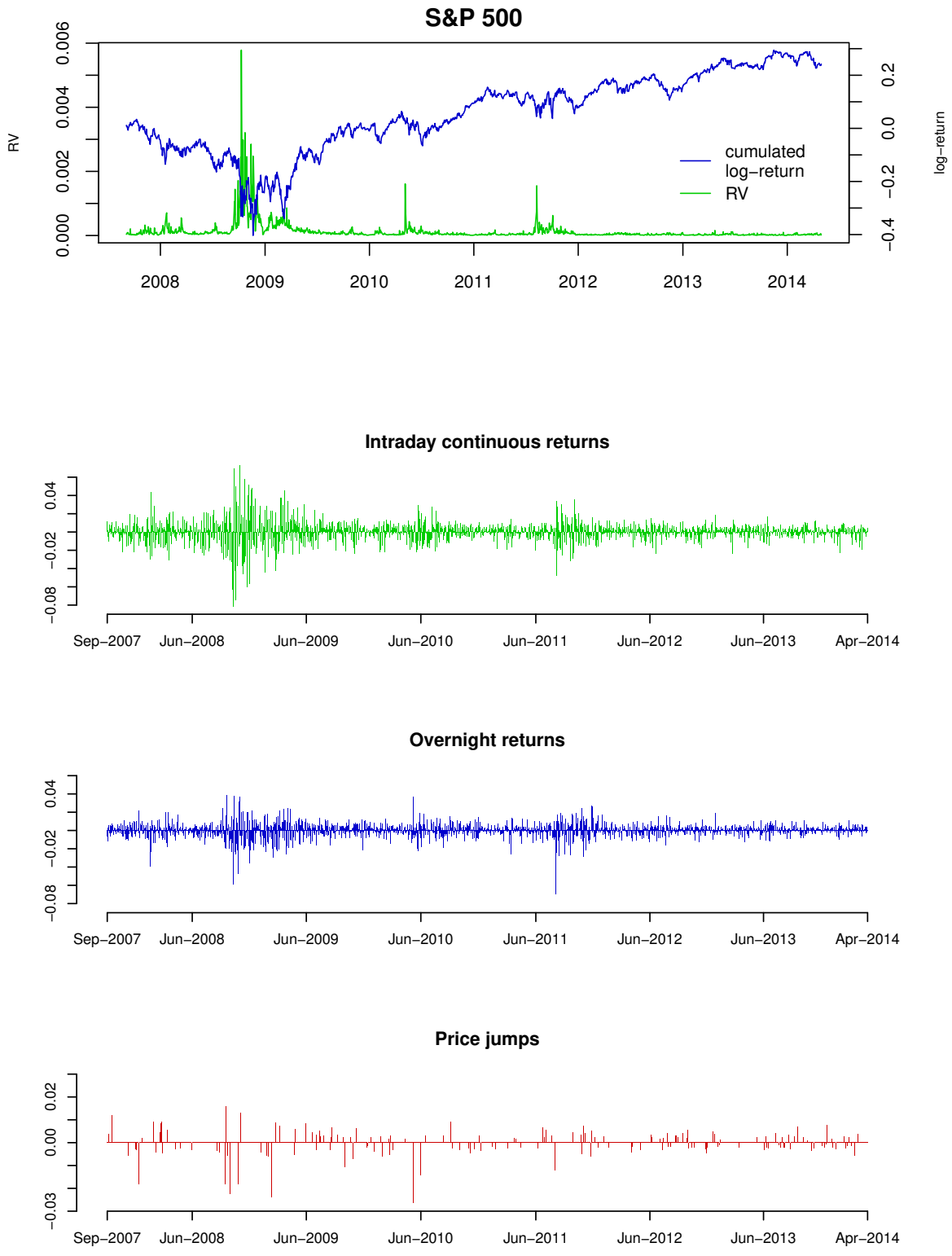


FIGURE3.2: Cumulated close-to-close log-return and the intraday realized variance measured from high frequency data.

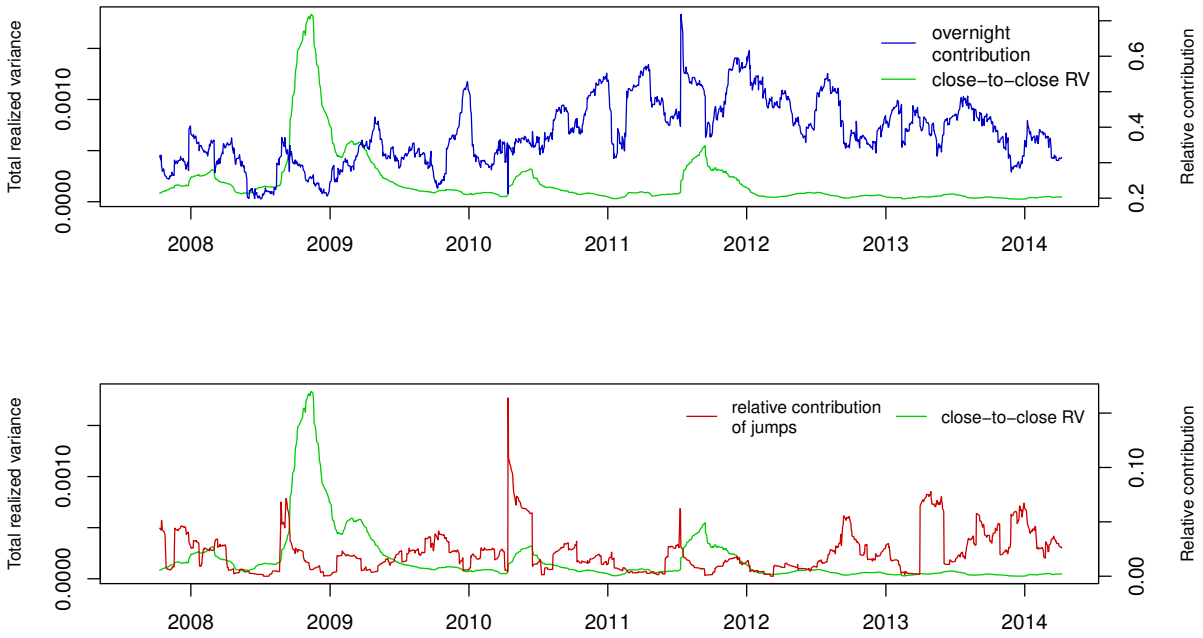


FIGURE3.3: In the top panel the relative contribution of overnight returns calculated as the simple moving average of overnight returns over a two-months rolling window centered on the reference date, divided by the moving average of the close to close realized variance. The contribution of jumps in the bottom panel is computed analogously.

3.4.2 Modeling the Return Dynamics

Our objective is to introduce an affine model for close-to-close returns that is able to account for the stylized facts presented in the previous section. Paralleling many other contributions in the literature, we assume the independence between the continuous and the discontinuous price fluctuations. This simplifying assumption neglects the dependence of the jump size from the volatility level found for instance by Andersen et al. (2011) as well as the dependence between volatility shocks and price jump. Particularly two distinct effects have been documented: the effect of price jumps on future volatility (see Corsi et al. 2010 and Corsi and Renò 2012) and the simultaneous occurrence of price and volatility jumps (Duffie et al. 2000, Todorov and Tauchen 2011, Jacod and Todorov 2010, Corsi and Renò 2012, Bandi and Renò 2016). Concerning the first we have no clear evidence from our data, as already observed in the supplementary analysis presented in Appendix 1.5. For the second, we experimented some alternative specifications where jumps in the price affect

the intensity of volatility jump but according to our analysis (not reported here) such effects are statistically insignificant. This is probably related to the fact that significant volatility shocks occur during the crisis where we find a small number of price jumps.

A model for price jump

The summary statistics in Table 3.1 is not compatible with a constant intensity Poisson jump process: in this case the probability to observe 5 or more jumps in a single trading day is of order 10^{-7}. The empirical analysis of Chapter 1 provides evidence that the jump intensity is not constant in time and is positively affected by the occurrence of jumps (self-excitation). Anyway this effect has a very short persistence, with an half-life time typically shorter than one hour. At the daily time scale, the probability to detect multiple jumps is higher then in a Poisson process with constant intensity. We therefore propose the following specification:

$$J_t = \sum_{i=1}^{N_t} K_i$$

where K_i is the random jump size and N_t is the random number of jumps occurring on day t. If jumps are i.i.d., the affine nature of the model is preserved when N_t is a negative binomial random variable:

$$N_t \sim \begin{cases} nb\left(r\left(\lambda_J\right), q_J\right) & \text{if } q_J > 0 \\ P\left(\lambda_J\right) & \text{if } q_J = 0 \end{cases} \tag{3.21}$$

where λ_J is assumed constant. The negative binomial distribution allows to capture the over-dispersion that is potentially generated by the rapidly decaying self-excitation observed in Chapter 1. Note that, even if λ_J is constant, the jump process can be interpreted as a compound Poisson with a random intensity. In the literature, evidence in favor of time varying intensity is provided by Chan and Maheu (2002), Maheu and McCurdy (2004), Christoffersen et al. (2012), Chen and Poon (2013) and Aït-Sahalia et al. (2015) among others. Importantly, most of this evidence comes from parametric models estimated on daily data. At that time scale, price jumps are hardly distinguishable from sudden volatility rises[4], thus it is not surprising that the empirical features of

[4] According to Christensen et al. (2014) and Bajgrowicz et al. (2016) even at 5 minutes volatility bursts are often miscalssified as jumps.

the jump process discovered by this literature are different from our results. Christoffersen et al. (2015) instead extract jumps from high frequency observations using a method that substantially differs from our approach measuring the realized jump variation as: $RV_t^J = \max(RV_t - BPV_t, 0)$ where BPV is the bipower variation. Unfortunately, in finite samples this quantity can be very often larger than zero due to the statistical error affecting the realized measures. Since the size of this error depends on the realized quarticity, it is not surprising that they find a strong evidence of persistence which is instead not confirmed by our empirical results of Chapter 1. Our modeling choice looks a reasonable and parsimonious description of the jump dynamics according to our empirical findings.

Intraday continuous returns, overnight returns and leverage

To model the continuous return component we are facing two problems: i) estimating the overnight realized variance which is not directly observable, ii) taking into account the deviation from normality which still arises after standardizing log-returns due to the presence of the contemporaneous leverage effects.

Let us start addressing the first problem: a simple and very common approach to take into account the overnight volatility consists in the application of a scaling factor $c > 1$ to the intraday realized volatility:

$$V_t = IV_t + ON_t = c\, IV_t$$

two estimators for c are often considered:

$$\hat{c} = \frac{\sum_t IV_t + ON_t}{\sum_t IV_t}$$

and

$$\hat{c} = \frac{\sum_t (y_t^c - \bar{y}_t^c)^2}{\sum_t IV_t}$$

where $\bar{y}_c = \frac{1}{T}\sum_{t=1}^{T} y_t^c$. For an exhaustive review of this topic see for instance Ahoniemi and Lanne (2013). We propose a slightly more sophisticated approach based on the assumption that the overnight realized variance is a linear combination of the intraday volatility realized on the previous

and the next trading day:

$$V_t = a\ (IV_t + b\, IV_{t-1}) \tag{3.22}$$

Empirical evidence suggests that this method improves substantially the accuracy over a simple scaling. Anyway we must recognize that equation 3.22 implies a relative contribution of overnight returns that is constant in time, in contrast with Figure 3.3. The proposed model leaves therefore room for further improvements based on a more accurate dynamic specification for overnight returns that is left for future developments. Good examples in this direction are Tsiakas (2008), proposing a stochastic volatility model that takes specifically into account overnight returns distinguishing between different non-trading periods (weeknights, weekends, holidays and long weekends) and Andersen et al. (2011) using a GARCH model.

Concerning the second point, we observe that discrete time models are generally able to capture the lagged leverage effect, i.e. the increase in volatility generated at $t+h$ by a negative return occurred at t. The contemporaneous leverage is instead easily featured in continuous time models when price and volatility innovations are assumed to be negatively correlated. To feature this important effect in discrete time, we propose the following specification for the continuous component of close-to-close returns:

$$y_t^c = \mu_t - \frac{V_t}{2} + l\left(\mathbb{E}\left[V_t \middle| \mathcal{F}_{t-1}\right] - V_t\right) + \sqrt{V_t}\epsilon_t \tag{3.23}$$

where $\mathcal{F}_t$ is the information set at the market closing time on day t, μ_t is the deterministic drift that will be specified below, $-V_t/2$ is the standard convexity adjustment and the parameter l (if positive) generates the intraday leverage effect by introducing a negative drift whenever the realized variance on day t is higher than its expected value according to the information available at $t-1$. This simple specification produces a negative skewness in the conditional return distribution at the daily time scale even in absence of price jumps, consistently with the stylized facts presented in

Section 3.4.1:

$$\begin{aligned}\mathbb{E}\left[\left(y_t^c - \mathbb{E}\left(y_t^c\right)\right)^3 \middle| \mathcal{F}_{t-1}\right] &= \mathbb{E}\left\{\left[\left(l+\frac{1}{2}\right)\left(\mathbb{E}\left[V_t \middle| \mathcal{F}_{t-1}\right] - V_t\right) + \sqrt{V_t}\epsilon_t\right]^3 \middle| \mathcal{F}_{t-1}\right\} \\ &= -\left(l+\frac{1}{2}\right)^3 \left(V_t - \mathbb{E}\left[V_t \middle| \mathcal{F}_{t-1}\right]\right)^3 \end{aligned} \tag{3.24}$$

The asymmetry of log-returns at the daily time scale in this model is closely related to the asymmetry of the volatility distribution. A stochastic volatility model that allows the occurrence of large volatility shocks, generates also strong asymmetries in the conditional distribution of daily log-returns. The cost of including the contemporaneous leverage is to introduce an explicit dependence between the return and the volatility dynamics (still unspecified at this stage) through the conditional expectation $\mathbb{E}\left[V_t \middle| \mathcal{F}_{t-1}\right]$ whose functional form depends on the volatility model. This cost is mostly paid in computational terms because the parameters of 3.23 can no longer be estimated separately from other volatility parameters like in Majewski et al. (2015) and Alitab et al. (2016). Notably, if the dynamics of IV_t is affine, then also V_t follows an affine process which in turn implies that $\mathbb{E}\left[V_t \middle| \mathcal{F}_{t-1}\right]$ depends linearly on lags of IV_t, guarantying that the full model is the end affine.

The drift is specified as follows:

$$\mu_t = R_t + \zeta_J \lambda_J + \zeta_V \mathbb{E}\left[V_t \middle| \mathcal{F}_{t-1}\right] \tag{3.25}$$

where R_t is the risk free rate[5], ζ_J is the jump risk premium and ζ_V is the volatility risk premium. Differently from Corsi et al. (2013), Majewski et al. (2015) and Alitab et al. (2016), we assume that the variance risk premium depends only on the information available at the end of the previous trading day. This modeling choice is consistent with the economic theory because the risk premium required by investors depends on their expectations that are incorporated in the closing price.

3.4.3 Stochastic Volatility

To complete the model we need to establish a dynamics for the open-to-close realized volatility IV_t. We consider two novel specifications based on the *ARGG* process introduced in this chapter.

[5] As a proxy we consider the Federal Funds Rate publicly available on the FED web site.

Both of them belong to the class of multivariate Car processes and encompass the LHARG model of Majewski et al. (2015). Importantly, the affine nature of these models guarantees the full tractability of option pricing formulas within the general framework of Majewski et al. where risk neutralization is obtained specifying a stochastic discount from the exponential affine family. In this case, the moment generating function under the risk neutral measure can be derived in semi-closed form according to Gourieroux and Monfort (2007). As a benchmark we consider the LHARG model of Majewski et al. assuming that the conditional distribution of the realized variance is a non-central gamma:

$$IV_{t+1} | \mathcal{F}_t \sim \bar{\gamma}(\delta, \lambda_{t+1}, \theta) \tag{3.26}$$

where

$$\lambda_{t+1} = \left(d + \beta_d IV_t + \beta_w {IV_t}^{(w)} + \beta_m {IV_t}^{(m)} + \alpha_d \ell_t + \alpha_w \ell_t^{(w)} + \alpha_m \ell_t^{(m)}\right)/\theta \tag{3.27}$$

being $IV_t^{(w)}$ and $IV_t^{(m)}$ respectively the weekly and the monthly volatility components that characterize the class of Heterogeneous Autoregressive models introduced by Corsi (2009) to mimic long memory effects:

$${IV_t}^{(m)} = \frac{1}{4}\sum_{i=2}^{5} IV_{t-i} \qquad {IV_t}^{(w)} = \frac{1}{17}\sum_{i=6}^{22} IV_{t-i}$$

the term $\ell_t = \left(\epsilon_t - \gamma_d \sqrt{V_t}\right)^2$ produces the lagged leverage effect in conjunction with $\ell_t^{(w)}$ and $\ell^{(m)}$ representing respectively the weekly and monthly components:

$$\ell_t^{(w)} = \frac{1}{4}\sum_{i=2}^{5}\left(\epsilon_t - \gamma_w \sqrt{V_t}\right)^2 \qquad \ell_t^{(m)} = \frac{1}{17}\sum_{i=6}^{22}\left(\epsilon_t - \gamma_m \sqrt{V_t}\right)^2$$

The Heterogeneous Autoregressive Generalized Gamma Model with Leverage (LHARGG)

We simply extend of the LHARG model using the generalized non-central gamma as a conditional distribution. The model specification differs from the LHARG only for the choice of the conditional density

$$IV_{t+1} | \mathcal{F}_t \sim \psi(\delta, \theta, \lambda_{t+1}, q) \tag{3.28}$$

Importantly in the limit $p \to 0$, this process converges weakly to the LHARG. Interestingly, the generalized model can be interpreted as an LHARG model with additional idiosyncratic heteroskedastic

noise affecting the non-centrality parameter

$$IV_{t+1} | \mathcal{F}_t \sim \bar{\gamma}(\delta, \lambda_{t+1} + \eta_{t+1}, \theta)$$

where η_t is distributed as a translated gamma with density

$$f(\eta | \lambda) = \gamma\left(\eta + \lambda\frac{1-q}{q}; \lambda\frac{1-q}{q}, \frac{q}{1-q}\right)$$

Including volatility jumps

The empirical results that will be presented in the next section show that the LHARGG, despite its superior flexibility with respect to the LHARG model, is still unable to feature the observed volatility dynamics. As already mentioned, a large body of literature suggests that multiple factors are needed to provide a realistic description of volatility. We therefore consider a model designed to capture two distinct scales of volatility innovations that characterize different time periods. At each time the conditional volatility is determined by the contribution of of two components: $IV_t = v_{1,t} + v_{2,t}$ where $v_{1,t}$ accounts for volatility fluctuations under tranquil market conditions, while $v_{2,t}$ produces large volatility shocks characterizing periods of market turmoil. We assume that these components are conditionally independent and distributed as a generalized non-central gamma and gamma-zero respectively:

$$v_{1,t} \sim \psi(\delta, \theta_1, \lambda_{1,t}, q_1) \tag{3.29}$$

$$v_{2,t} \sim \psi(0, \theta_2, \lambda_{2,t}, q_2) \tag{3.30}$$

Note that the latter distribution is simply a compound negative binomial with a finite probability mass at zero. The component $v_{2,t}$ is therefore well designed to feature volatility jumps. The conditional density of IV is

$$IV_{t+1} | \mathcal{F}_t \sim \psi_2(\delta, \theta_1, \lambda_{1,t+1}, q_1, \theta_2, \lambda_{2,t+1}, q_2) \tag{3.31}$$

where ψ_2 is the convolution of the two densities

$$\psi_2(x;\,\delta,\theta_1,\lambda_1,q_1,\theta_2,\lambda_2,q_2) = \sum_{n,m=0}^{\infty} nb(n;r_1,q_1)\; nb(m;r_2,q_2) \int_0^x \gamma(u;\,\delta+n,\theta_1)\,\gamma(x-u;\,m,\theta_2)\,du \tag{3.32}$$

and $r_1 = \dfrac{1-q_1}{q_1}\lambda_1,\; r_2 = \dfrac{1-q_2}{q_2}\lambda_2$. The heterogeneous autoregressive generalized gamma model with leverage and volatility jumps (LHARGG-VJ) is specified as:

$$\lambda_{1,t+1} = \frac{1}{\theta_1}\Big(d_1 + \beta_{1,d}IV_t + \beta_{1,w}{IV_t}^{(w)} + \beta_{1,m}{IV_t}^{(m)} + \alpha_{1,d}\ell_{1,t}^{(d)} + \alpha_{1,w}\ell_{1,t}^{(w)} + \alpha_{1,m}\ell_{1,t}^{(m)}\Big) \tag{3.33}$$

$$\lambda_{2,t+1} = \frac{d_2 + \beta_{2,d}IV_t + \alpha_{2,d}\ell_{2,t}}{\theta_2} \tag{3.34}$$

where $\ell_{i,t} = \left(\epsilon_t - \gamma_{i,d}\sqrt{V_t}\right)^2$ for $i = 1, 2$. Even if the conditional density 3.32 is rather complicated, the corresponding moment generating function has a simple closed form expression:

$$\hat{\psi}_2(z;\,\delta,\theta_1,\lambda_1,q_1,\theta_2,\lambda_2,q_2) = (1-\theta_1 z)^{-\delta}\left[\frac{1-q_1}{1-q_1/(1-\theta_1 z)}\right]^{r_1} \left[\frac{1-q_2}{1-q_2/(1-\theta_2 z)}\right]^{r_2} \quad z \in \mathbb{C} \tag{3.35}$$

the conditional mean and variance of this distribution are:

$$\begin{aligned} \mathbb{E}(x) &= \theta_1(\delta+\lambda_1) + \theta_2\lambda_2 \\ Var(x) &= \theta_1^2\left(\delta + \lambda_1\frac{2-q_1}{1-q_1}\right) + \theta_2^2\left(\lambda_2\frac{2-q_2}{1-q_2}\right) \end{aligned}$$

The numerical computation of the conditional density is discussed in Appendix C.3.2. The interpretation of expression 3.31 is relatively simple: when $m = 0$ the distribution of $v_{2,t}$ degenerates

into a Dirac delta with the whole probability mass concentrated at zero. This scenario corresponds to the absence of volatility jumps and occurs with probability $(1-q_2)^{r_{2,t}}$. Volatility jumps are i.i.d. and non-negative, following a gamma distribution with shape parameter equal to 1 and scale θ_2. The non-centrality parameter λ_2 can be interpreted as the average conditional jump intensity. The effective jump intensity is distributed as a gamma with shape parameter $r_{2,t}$ and scale parameter $q_2/(1-q_2)$. The probability of having n volatility jumps at time t is a negative binomial with parameters q_2 and $r_{2,t}$. The persistence of the jump intensity is determined by the HAR structure of equation 3.33 controlling the component v_1. For this reason it is unnecessary to introduce additional lags into equation 3.34.

3.4.4 Estimation Strategy and Results

Price jump component Under the simplifying assumption that the jump component is independent from the rest of the process and that jumps are exactly observed, the corresponding parameters can be estimates separately. Results on Table 3.2 show that q_J is significantly larger than zero, rejecting the hypothesis of a Poisson process with constant jump intensity. Some experiments have been also conducted using alternative dynamic specifications to include an autoregressive component in the jump activity but it does not produce statistically significant effects according to our results (not reported here).

Continuous return component In analogy with Corsi et al. (2013), Majewski et al. (2015) and Alitab et al. (2016) we assume that the realized volatility IV_t and therefore the close-to-close volatility V_t are observed without error. This assumption remarkably simplifies the statistical inference due to the absence of latent variables. We must anyway recognize that, even if the close to close volatility is exactly determined by equation 3.22, the intaday realized volatility is always measured with an error. To quantify its size note that it has a variance equal to $2\,IQ_t/(M-N_t)$, where N_t is number of intervals containing jumps on day t and IQ_t is the intraday realized quarticity. From empirical data we find that the standard deviation of the relative error on $\hat{IV}_t$ ranges from 16% to 57% with an average value of 19%. Bormetti et al. (2016) study the LHARG model under the realistic assumption that the realized volatility is observed with an error and rely on MCMC techniques to perform the statistical inference, finding an improvement of the pricing performances.

Remarkably, our LHARGG can be regarded as an LHARG model with noisy non-centrality. This allows to design approximated filtering algorithms of the kind proposed in Section 3.3.2 which are applicable in case of not perfectly observed state variables. This development is left for future research.

Under full information, in the specific case $l = 0$ (no contemporaneous leverage), the parameters of the stochastic volatility process can be in principle estimated separately from the parameters of equations 3.23 and 3.25 based on two-stage maximum likelihood. In the general framework instead, the whole bivariate time series of volatility and returns must be jointly considered. The introduction of the simultaneous leverage effect increases the computational burden related to the numerical optimization of the likelihood function. It is worthwhile to mention that two parameters out of five introduced in equations 3.22 and 3.23 can be determined in closed form by solving explicitly the first order conditions after applying appropriate transformations.

Our estimates for the different model specifications are reported on Table 3.2. To reduce the number of free parameters, some constraints have been imposed after some experimentation. Namely, since the parameters d_1 and d_2 are zero at the optimum we assume $d_1 = d_2 = 0$; the parameters $\gamma_{1,d}$, $\gamma_{1,w}$, $\gamma_{1,m}$ and $\gamma_{1,d}$ are constrained to take the same value γ because heterogeneous leverage provides poor improvements to the likelihood. The overdispersion parameters q_1 and q_2 of the LHARGG-VJ model take very similar values at the optimum and their difference is statistically insignificant, thus we impose $q_1 = q_2 = q$. $\beta_{1,d}$, $\beta_{1,w}$, $\beta_{1,m}$ and $\beta_{2,d}$ are very close to the zero lower bound at the maximum and are found to be statistically insignificant. This is not surprising for at least a couple of reasons. First each leverage terms ℓ_t contain a volatility component equal to $\gamma^2 V_t = \gamma^2 \left(a\, RV_t + a\, b\, RV_{t-1}\right)$ and the constraints imposed to prevent the non-centrality parameter from going negative imply that a higher leverage leads also to more volatility persistence[6]. Second, when the volatility is high, the autoregressive term dominates: $\ell_t \simeq \gamma^2 V_t$. The importance of leverage to describe our data is easily arguable looking at Figure 3.2 where the volatility rise in 2008 is accompanied by a large decline of the index. Thus it is not surprising that the size of γ necessary to capture the leverage effect is also sufficiently large to account for the volatility persistence. In order to facilitate the likelihood

[6] Majewski et al. (2015) propose a specification of the LHARG named ZM-LHARG where the non-negativity constraint can be violated.

optimization and its numerical differentiation, we impose the constraints $\beta_{1,d} = \beta_{1,w} = \beta_{1,m} = 0$. Our estimates show a high volatility and leverage persistence reflected by the significance of the monthly coefficient α_m across all model specifications. Concerning the return dynamics, we note that the simultaneous leverage is extremely relevant: the parameter l is always positive, consistently with the negative skewness observed on daily returns (see Table 3.1). The estimates of the market price of variance risk suggest that the positive relation between lagged volatility and the future returns is not very strong in our period of analysis.

The coefficients a and b of equation 3.22 are both positive and significant, estimates are very similar under the different volatility specifications and imply a contribution of around 38% of the overnight to the total realized variation, very close to the non-parametric estimates reported in Section 3.4.1. The overnight volatility is approximated by a combination of the intraday volatility realized on the previous and on the following trading day with weights around 60% and 40% respectively. Anyway we recognize that our description of overnight volatility can be probably improved taking into account a time varying relative contribution to the total realized variance that could be for instance obtained with a supplementary GARCH model.

Moving from the LHARG to the LHARGG, the log-likelihood increases remarkably thanks to the inclusion of a single additional parameter p controlling the over-dispersion of the conditional volatility distribution. This feature seems extremely relevant to describe the volatility dynamics. The crucial role of volatility jumps is primarily reflected by the high significance level of the parameter $\alpha_{2,d}$ for the LHARGG-VJ model. Note that the effect of volatility jump is highly persistent due to HAR structure of the component v_1. Any attempt to introduce additional autoregressive terms in v_2 does not produce relevant improvements of the fit. The over-dispersion parameter p remains significantly positive, causing a strong rejection of the non-central gamma as conditional distribution of the two volatility components.

Figures 3.4 and 3.5 compare the expected and the realized volatility for each model specification. Remarkably the realized volatility falls very often outside the 99.9% confidence bands for the LHARG and the LHARGG which are unable to capture the volatility peaks observed during the sub-prime crisis. The LHARGG-VJ model instead admits large volatility shocks generated by volatility jumps, whose estimated size is about 35 times larger than the scale of ordinary volatility fluctuations. All

the models fail to capture the sudden volatility drops observed within the same period. This problem is probably due to the long persistence generated by the HAR structure[7] Figure 3.6 shows the average contribution of the two volatility components: as expected the role of volatility jumps is particularly relevant during the sub-prime crisis.

To check our assumptions on the dynamic specification of returns, the distribution of residuals is compared on Figure 3.8 with a standard normal. The deviations from normality are not very strong, as confirmed also by the outcomes of the normality tests reported on Table 3.3. Importantly, the normality of residuals is not rejected at the 1% confidence level.

To test the alternative dynamic specifications for stochastic volatility, we construct the series $U_t = F\left(IV_t|\,\mathcal{F}_t\right)$ where $F\left(IV_t|\,\mathcal{F}_t\right)$ is the conditional cdf of the continuous intraday volatility. Under the assumption that the conditional density is correctly specified, the transformed values are uniformly distributed. This property can be used to develop a formal specification test, for example using the Kolmogorv-Smirnov statistics. The empirical distributions of U_t under the alternative specifications are shown on Figure 3.7, where the deviation from the uniform is quite clear for the LHARG model and confirmed also by a p-value of the test which is very close to zero. The LHARGG model seems to produce a better fit but still not compatible with the empirical data according to our test which rejects the null at the 0.1% confidence level. In the LHARGG-VJ model the inclusion of volatility jumps component together with the over-dispersion associated to the generalized non-central gamma distribution, seem crucial to capture the empirical features of the realized volatility: our test returns a p-value of 5%. Figure 3.9 compares a Monte Carlo simulation of three volatility processes with the empirical data. The LHARGG-VJ specification produces a pattern that appears much more similar to the observed dynamics. This results confirm the necessity to include at least two stochastic factors for an acceptable description of the volatility evolution. Nevertheless, we note that volatility jumps seem to occur more often in the simulated process then in reality and at the same time the peaks in volatility are smaller compared to the levels reached in 2008. From Figure 3.10, where the simulated unconditional density of continuous realized volatility and close to close continuous returns are compared with the empirical counterparts, we observe remarkable improvements of the

[7] Perhaps it could be overcome assuming the full independence between the two components. In this case anyway the inference becomes more complicated due to the presence of latent state variables.

fit moving from the LHARG to the LHARGG an then to the LHARGG-VJ specification. The gain is particularly relevant in the tails of the distributions. In this regard, the transition from the *ARG* process to its generalized version clearly plays a crucial role. Nevertheless, non of the models analyzed in this chapter is able to fully capture the empirical features of volatility during the period covered by our analysis. Providing an accurate statistical description of the huge volatility shocks observed during the sub-prime crisis through an affine model remains a challenging problem. While the main limitation of the LHARG and the LHARGG is the impossibility to account for sudden volatility shocks, the LHARGG-VJ model is unable to capture the small values of realized volatility observed during the sub-prime crisis very close in time to high volatility peaks. The problem is that in this model jumps in volatility affect also the ordinary component v_1 which takes a long time to revert to low levels, due to the persistence generated by the HAR structure. A possible solutions to this is to abandon the assumption of full information and to introduce some latent state variables, either assuming the full independence between the two volatility components or taking into account the measurement that affects the realized volatility.

Price jumps	
parameter	estimate
λ_J	$1.17 \cdot 10^{-1***}$
	$(9.23 \cdot 10^{-3})$
q_J	$1.97 \cdot 10^{-1***}$
	$(4.17 \cdot 10^{-2})$
$\log L$	-625.1

$^{***}p < 0.001$ $^{**}p < 0.01$ $^{*}p < 0.05$

Stochastic volatility			
parameter	LHARG	LHARGG	LHARGG-VJ
θ_1	$2.15 \cdot 10^{-5***}$	$3.23 \cdot 10^{-6***}$	$2.22 \cdot 10^{-6***}$
	$(3.93 \cdot 10^{-7})$	$(2.61 \cdot 10^{-7})$	$(2.27 \cdot 10^{-7})$
δ	1.03^{***}	5.61^{***}	5.82^{***}
	$(8.70 \cdot 10^{-2})$	$(3.65 \cdot 10^{-1})$	$(4.64 \cdot 10^{-1})$
$\alpha_{1,d}$	$1.07 \cdot 10^{-6***}$	$9.75 \cdot 10^{-7***}$	$6.55 \cdot 10^{-7***}$
	$(2.02 \cdot 10^{-7})$	$(2.34 \cdot 10^{-7})$	$(1.43 \cdot 10^{-7})$
$\alpha_{1,w}$	$2.24 \cdot 10^{-7***}$	$2.37 \cdot 10^{-7***}$	$2.06 \cdot 10^{-7***}$
	$(5.02 \cdot 10^{-8})$	$(6.97 \cdot 10^{-8})$	$(5.41 \cdot 10^{-8})$
$\alpha_{1,m}$	$9.31 \cdot 10^{-8***}$	$1.32 \cdot 10^{-7***}$	$2.24 \cdot 10^{-7***}$
	$(2.29 \cdot 10^{-8})$	$(3.55 \cdot 10^{-8})$	$(4.51 \cdot 10^{-8})$
γ	$-5.83 \cdot 10^{2***}$	$-6.10 \cdot 10^{2***}$	$-6.02 \cdot 10^{2***}$
	$\left(5.58 \cdot 10^{1}\right)$	$\left(7.63 \cdot 10^{1}\right)$	$\left(6.67 \cdot 10^{1}\right)$
a	1.35^{***}	1.37^{***}	1.39^{***}
	$(5.75 \cdot 10^{-2})$	$(6.62 \cdot 10^{-2})$	$(6.81 \cdot 10^{-2})$
b	$2.59 \cdot 10^{-1***}$	$2.45 \cdot 10^{-1***}$	$2.24 \cdot 10^{-1***}$
	$\left(3.21 \cdot 10^{-2}\right)$	$\left(4.20 \cdot 10^{-2}\right)$	$\left(4.19 \cdot 10^{-2}\right)$
ζ_J	$-4.52 \cdot 10^{-2}$	$6.55 \cdot 10^{-4}$	$4.45 \cdot 10^{-3*}$
	$\left(2.74 \cdot 10^{-3}\right)$	$\left(2.27 \cdot 10^{-3}\right)$	$\left(1.93 \cdot 10^{-3}\right)$
ζ_V	4.25	1.31	$-0.97 \cdot 10^{-1}$
	(2.45)	(2.45)	(3.43)
l	$5.01 \cdot 10^{1***}$	$4.98 \cdot 10^{1***}$	$4.87 \cdot 10^{1***}$
	(4.69)	(4.79)	(4.80)
q	–	$9.52 \cdot 10^{-1***}$	$8.94 \cdot 10^{-1***}$
	–	$\left(4.30 \cdot 10^{-3}\right)$	$\left(1.36 \cdot 10^{-3}\right)$
θ_2/θ_1	–	–	$3.48 \cdot 10^{1***}$
	–	–	(7.76)
$\alpha_{2,d}$	–	–	$9.88 \cdot 10^{-8***}$
	–	–	$(1.36 \cdot 10^{-8})$
$\log L$	19,643.23	19,979.67	20,180.19

Table 3.2: Estimated parameters for the different model specifications.

Normality test on ϵ_t	LHARG	LHARGG	LHARGVJ
Shapiro-Wilk	0.03	0.03	0.03
Kolmogorov-Smirnov	0.13	0.16	0.08

	Ljung-Box test ϵ_t			**Ljung-Box test** ϵ_t^2		
lags	LHARG	LHARGG	LHARGG-VJ	LHARG	LHARGG	LHARGG-VJ
5	0.201	0.112	0.224	0.097	0.100	0.088
10	0.449	0.311	0.130	0.366	0.350	0.304
15	0.264	0.182	0.152	0.071	0.060	0.049
20	0.420	0.324	0.155	0.051	0.044	0.038

Table 3.3: Tests on standardized close-to-close returns.

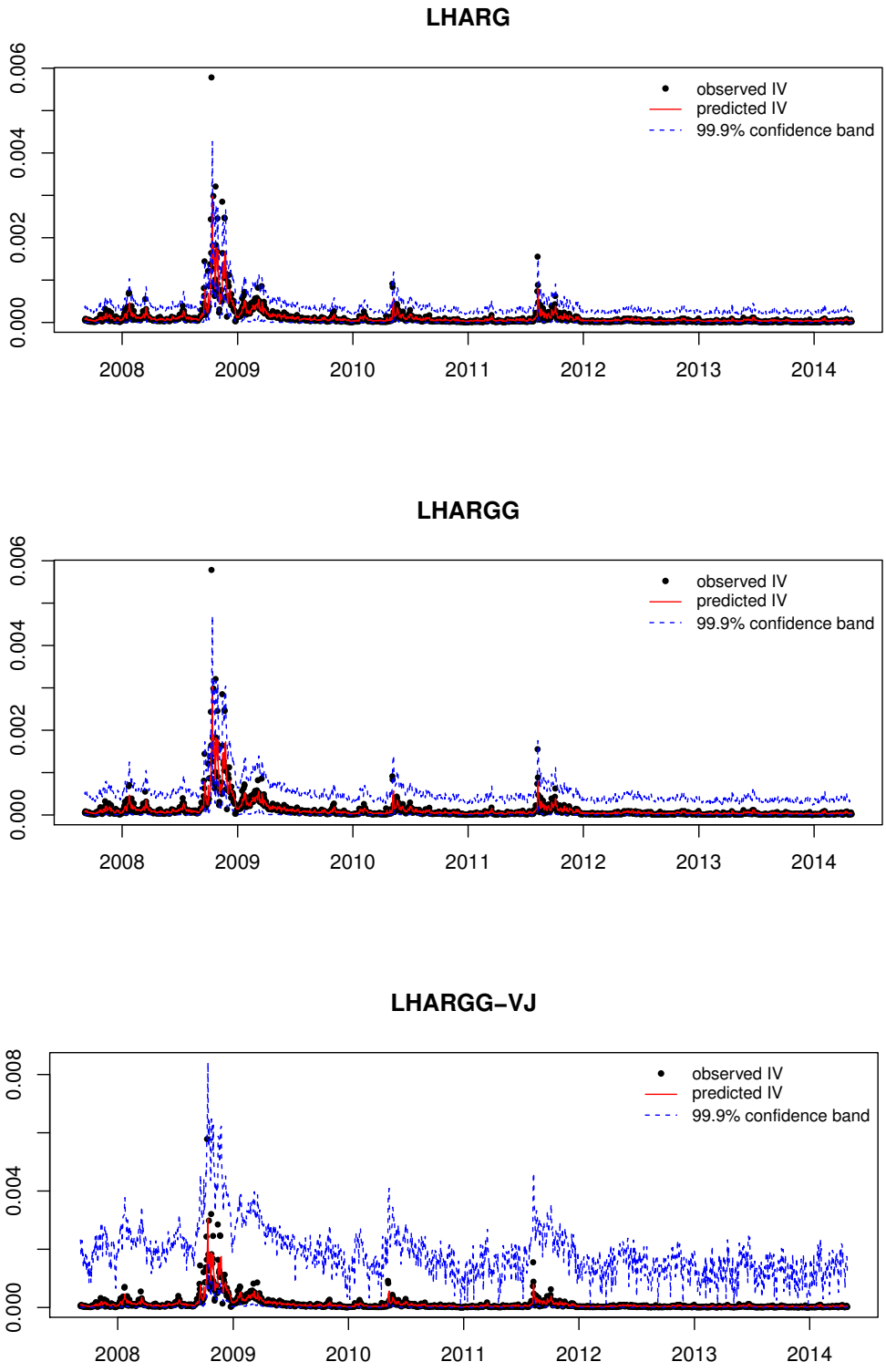


FIGURE3.4: The figures compare the variance predicted by the each model with the realized quadratic variation. The 99.9% confidence bands at each time t are calculated based to the conditional distribution of RV_t with respect to the information set at $t-1$.

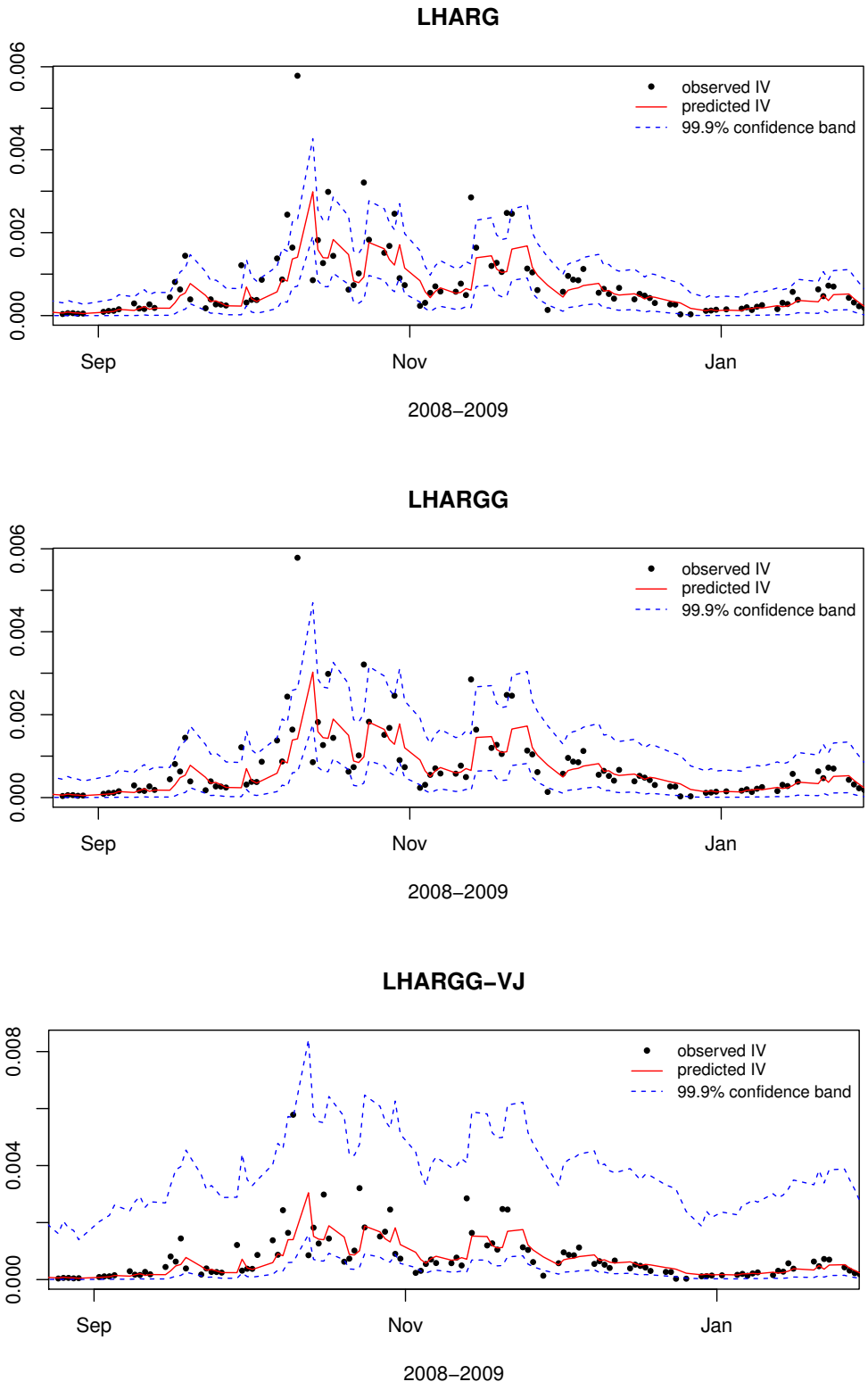


FIGURE3.5: Variance predicted by each model compared with the realized quadratic variation during the sub-prime crisis. The 99.9% confidence bands at each time t are calculated based on the conditional distribution of RV_t w.r.t. the information set at $t-1$.

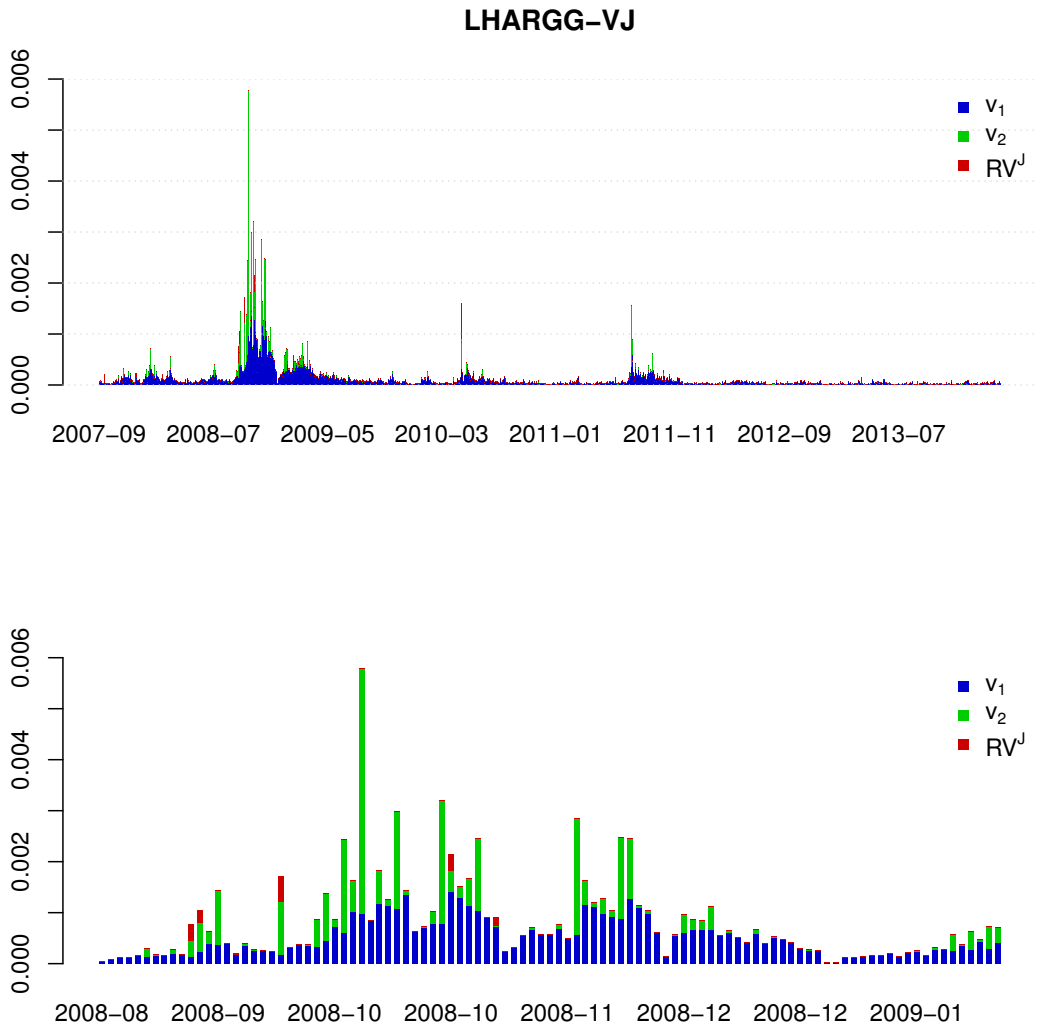


FIGURE3.6: Contribution from the different volatility components.

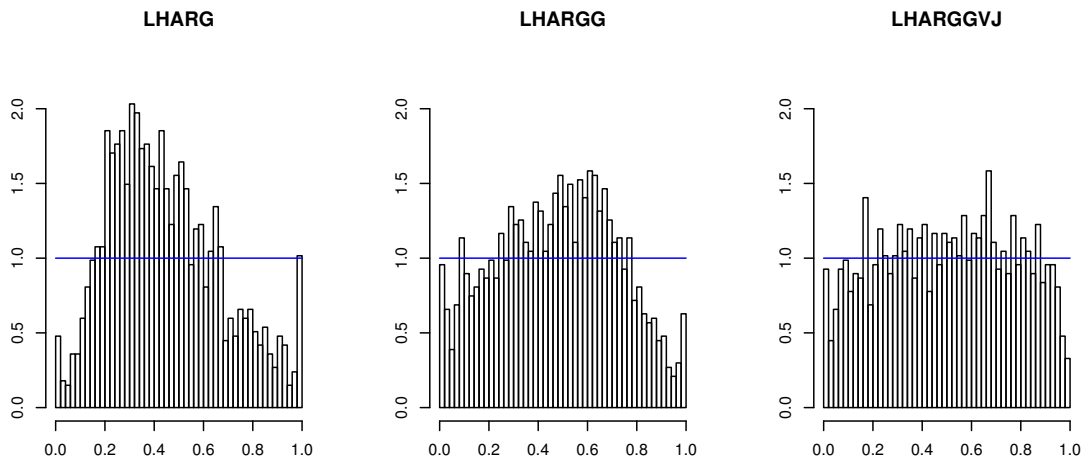


FIGURE3.7: The figure shows the empirical distribution of U_t for the alternative specifications.

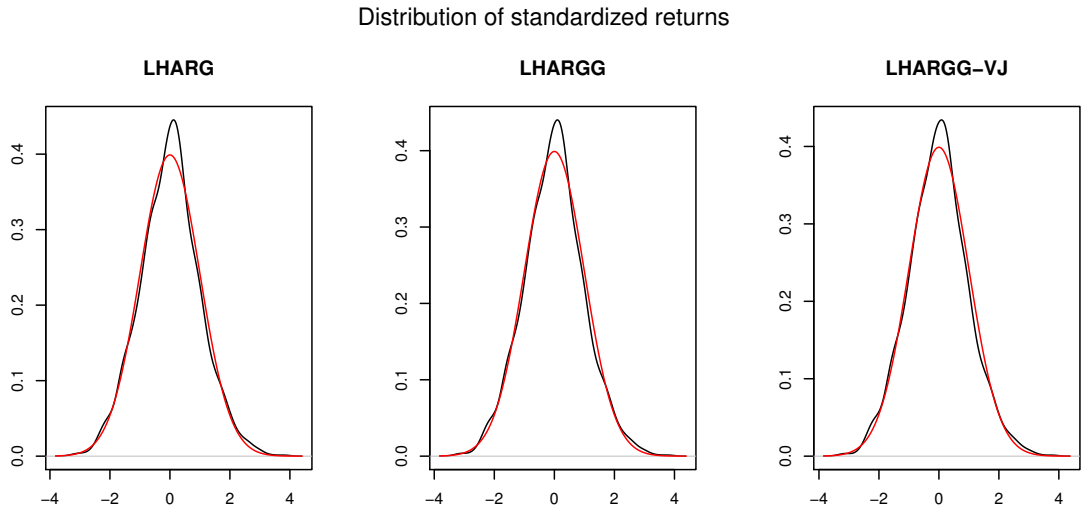


FIGURE3.8: Distribution of residuals ϵ_t computed from standardized close-to-close log returns. The difference between alternative specifications is due only to the conditional expectation $\mathbb{E}[V_t|\mathcal{F}_{t-1}]$ which determines the drift of daily log returns in equation 3.23.

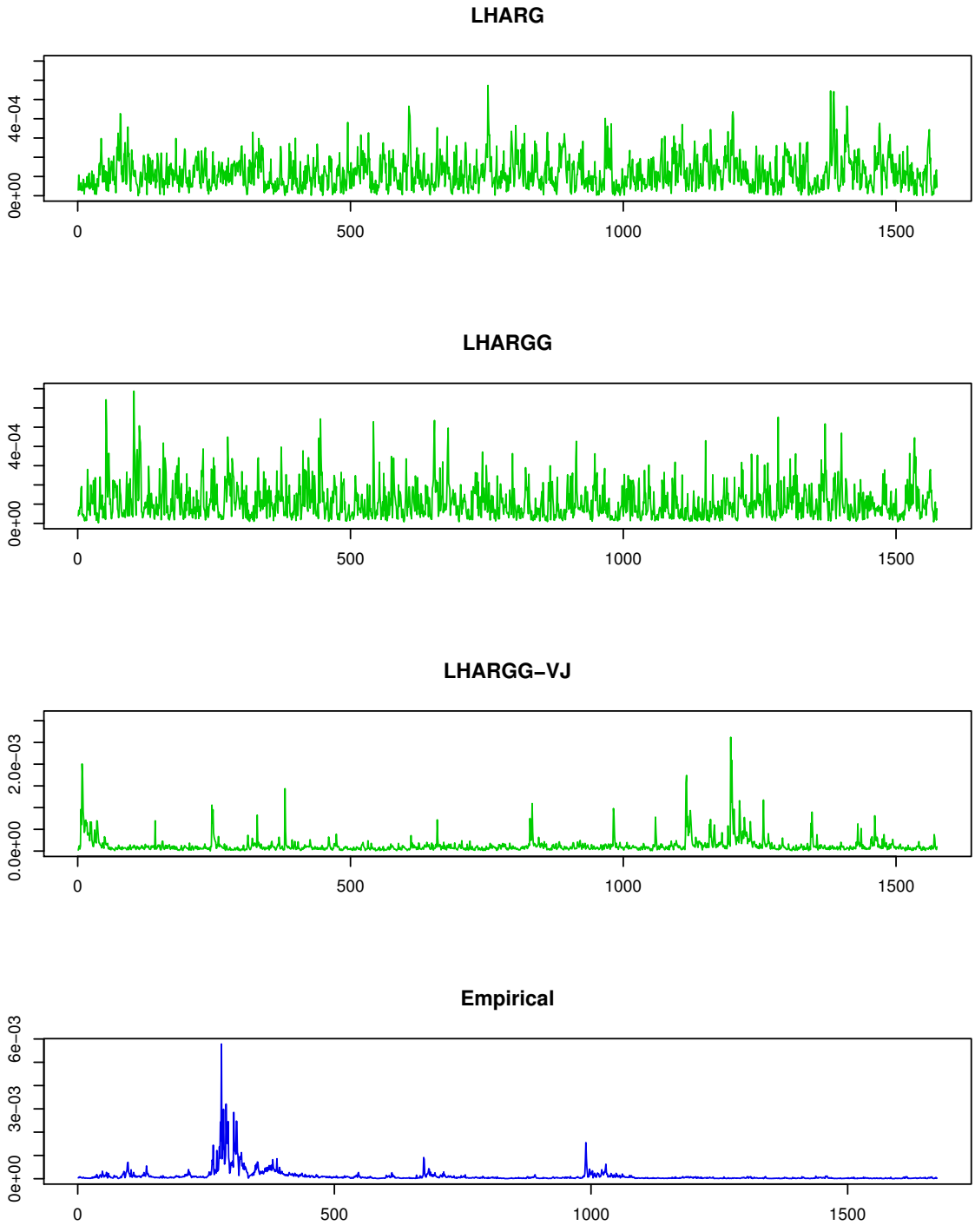


FIGURE3.9: Monte Carlo simulation for the different processes driving the continuous volatility compared with the empirical data.

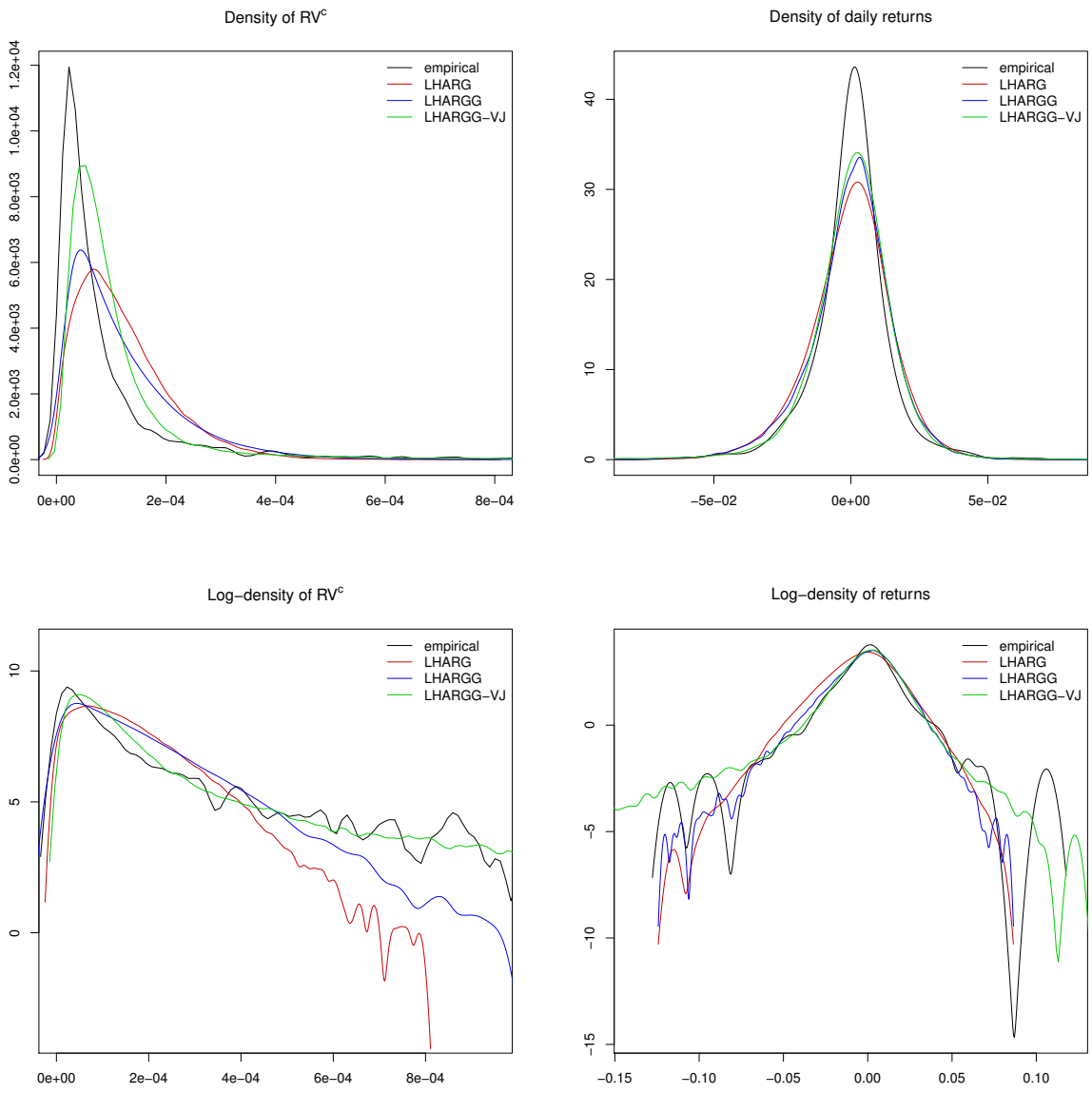


FIGURE3.10: Unconditional distribution of the continuous intraday realized volatility (left) and of close-to-close log returns (right).

3.5 Concluding Remarks

We introduced a new stochastic process generalizing the autoregressive gamma model of Gourieroux and Jasiak (2006) and preserving the same affine properties. The empirical application to stochastic volatility provides clear evidence on its superior flexibility that allows to better fit the dispersion of the data. Consistently with many previous studies, we find confirmation on the need of multiple volatility factors to provide an acceptable description of the empirical data. According to our model

specification, one component produces ordinary volatility fluctuations characterizing tranquil market conditions while the second component is responsible for large volatility shocks with a persistent effect characterizing periods of market turmoil. At the same time we find that price jumps provide on average a small contribution to the total price variation becoming more relevant during periods of low volatility. Moreover, empirical evidence suggests that price jumps are unable to explain the distributional asymmetry of returns observed at the daily time scale, which is instead adequately captured by a contemporaneous leverage effect that we introduce in our affine specifications through a stochastic drift component.

The results of this chapter suggest several directions for future research. Applications of the *ARGG* to option pricing are surely of primary interest: the over-dispersion introduced in conditional volatility and the contemporaneous leverage have important consequences on the tail decay and on the asymmetry of the return distribution. The implications can be extremely relevant for short term option prices that are highly sensitive to short-scale distributional asymmetries, sudden volatility shocks and price jumps. In this regard we remark that our empirical findings induce to reconsider the role of price jumps as a determinant of price variability while a prominent role during periods of stress seems to be played jumps in volatility, consistently also with the empirical evidence reported in Chapter 1.

Finally, in this chapter we have mentioned several enhancements to the proposed models to further improve the fit under the physical measure. The main limitation to the LHARGG-VJ specification is probably related to the requirement of using only fully observed state variables. This imposes severe restrictions to the architecture of the model, probably compromising the flexibility of the autocorrelation structure of the two volatility components. Considering the possibility that some state variables are not observed may allows to improve significantly the performances of this. A second aspect that it is worthwhile to mention is the importance of overnight returns which provide a relevant contribution to the total price variation that is apparently non constant in time. This calls for a more sophisticated description, based for instance on GARCH models like in Andersen et al. (2011), or on stochastic volatility models like in Tsiakas (2008).

Appendix C

Technical Details of Chapter 3

C.1 Properties of the Generalized Non-Central Gamma Distribution

C.1.1 Proof of Proposition 10

Consider the Fourier inversion integral:

$$\psi(x) = \frac{1}{2\pi\imath}\int_{z_0-\imath\infty}^{z_0+\imath\infty} \hat{\psi}(z)\, e^{-zx} dz \tag{C.1}$$

where $Re(z_0) < (1-q)/\theta$. For simplicity we assume $\theta = 1$ from now on. For x large the integral can be approximated using the saddle point method. Let

$$g(z) = \log\hat{\psi}(z) = (r-\delta)\log(1-z) - r\log(1-q-z) + r\log(1-q)$$

the saddle point z^* is the solution of the first order condition $g'(z^*) = x$ with $g''(z^*) > 0$. We have

$$\begin{aligned} g'(z) &= \frac{1}{1-z}\left[\delta + \lambda\frac{1-q}{1-z-q}\right] \\ g''(z) &= \frac{1}{(1-z)^2}\left[\delta + \lambda(1-q)\frac{2(1-q)-q}{(1-z-q)^2}\right] \end{aligned}$$

note that the second order derivative of g is always positive for $z < 1-q$, thus

$$z^*(x) = 1 - \frac{1}{2}\left[q + \frac{\delta}{x} + \sqrt{\left(q-\frac{\delta}{x}\right)^2 + 4\lambda\frac{1-q}{x}}\right] < 1-q \quad \forall x > 0$$

choosing $z_0 = z^*$ in C.1 we obtain

$$\begin{aligned}\psi(x) &= \frac{e^{g(z^*)-z^*x}}{2\pi}\int_{-\infty}^{\infty} e^{-g''(z^*)s^2/2} + h(x,s)\,ds \\ &= \frac{e^{g(z^*)-z^*x}}{\sqrt{2\pi g''(z^*)}}\left\{1+\int_{-\infty}^{\infty} h\left(x, u\sqrt{2/g''(z^*)}\right)du\right\} \qquad \text{(C.2)}\end{aligned}$$

where $h(s,x) = \exp\left[g(z^*+\imath s) - g(z^*) - \imath s x\right] - \exp\left[-g''(z^*)\,s^2/2\right]$. The first term is the saddle point approximation, now we prove that for large values of x, the integral in C.2 converges to a constant. Consider $q > 0$, for x large we have

$$\begin{aligned} z^* &\simeq 1 - q - r/x \\ g(z^*) &\simeq r\,\log x \\ g''(z^*) &\simeq x^2/r \end{aligned}$$

where $r = \lambda(1-q)/q$

$$h\left(x, u\sqrt{2/g''(z^*)}\right) = \left(1-\imath u\sqrt{2/r}\right)^{-r} e^{-\imath u\sqrt{2r}} - e^{-u^2} + O\left(x^{-1}\right)$$

as $x \to \infty$, all the dependence on x is contained in the in the pre-factor of C.2 thus we conclude $\psi(x) \propto x^{r-1}e^{-x(1-q)}$. In the general case $\theta \neq 1$, $\psi(x) \propto (x/\theta)^{r-1}\,e^{-x(1-q)/\theta}$ which proves the result.

Sum of generalized non-central gamma random variables It is well known that the sum of two negative binomial random variables with intensities r_1 and r_2 and the same probability of success p is again a negative binomial with shape $r_1 + r_2$. This property is naturally inherited also by the generalized non-central gamma.

Proposition 13. The sum of the the random variables $X_1 \sim \psi(\delta_1, \theta, \lambda_1, q)$ and $X_2 \sim \psi(\delta_2, \theta, \lambda_2, q)$ with $\theta > 0$ and $\lambda_1, \lambda_2, q, \delta_1, \delta_2 \geqslant 0$ is distributed as a generalized non-central gamma with parameters $\delta = \delta_1 + \delta_2$, θ, $\lambda = \lambda_1 + \lambda_2$, q.

Proof: The m.g.f. of $Y = X_1 + X_2$ is

$$\mathbb{E}\left[e^{z(X_1+X_2)}\right] = (1-\theta z)^{-(\delta_1+\delta_2)} \left(1 - \frac{q}{1-q}\frac{\theta z}{1-\theta z}\right)^{-r(\lambda_1+\lambda_2,q)}$$

Proposition 14. The generalized non-central gamma distribution is infinitely divisible.

Proof: For any $n \in \mathbb{N}$ such that $n > 0$, a generalized non-central gamma random variable Y with parameters $\delta, \theta, \lambda, q$ can be obtained as the sum of n i.i.d. random variables extracted from a distribution of the same type with parameters $\delta, \theta, \lambda/n, q$.

Changing the scale of a gamma random variable The following corollary states that the sum of a gamma random variable and a generalized non-central gamma with appropriate parameters is distributed as a gamma with a larger scale. Since when $\delta = r(\lambda, q)$ the generalized non-central gamma reduces to the gamma distribution with shape parameter $r(\lambda, q)$ and scale parameter $\theta/(1-q)$, this result can be regarded as a corollary of Proposition 13.

Corollary 15. The variable $Y = X_1 + X_2$ such that $X_1 \sim \gamma(\delta_1, \theta_1)$ and $X_2 \sim \psi(\delta_2, \theta_1, \lambda, q)$ with $\lambda = \frac{q}{1-q}(\delta_1 + \delta_2)$, is distributed as gamma with parameters $\delta = \delta_1 + \delta_2$ and $\theta = \frac{\theta_1}{1-q}$.

Proof: The moment generating function of Y is:

$$\mathbb{E}\left[e^{zY}\right] = (1-\theta_1 z)^{-(\delta_1+\delta_2)} \left[\frac{1-q}{1-q/(1-\theta_1 z)}\right]^{\delta_1+\delta_2} = \left(1 - \frac{\theta_1}{1-q}z\right)^{-(\delta_1+\delta_2)}$$

Mixing generalized non-central gamma random variables The next property states that taking an appropriate negative binomial mix of generalized non-central gamma distributions, all characterized by a certain overdispersion parameter $q \in (0,1)$, it is possible to obtain another distribution of the same family with an associate parameter q' such that $0 < q \leqslant q' < 1$.

Proposition 16. Consider the negative binomial random variable N with parameters r, q_{mix} and a non negative random variable X such that its distribution conditionally on N is a gener-

alized non-central gamma with parameters δ, $\lambda_N = (r+N)\,q/(1-q)$, q, θ. Then the unconditional distribution of X is still a generalized non-central gamma with parameters δ, λ, q', θ where $q' = \dfrac{q}{p+(1-q)(1-q_{mix})}$

Proof: Assume $\theta = 1$ without loss of generality. The unconditional moment generating function of X is

$$\begin{aligned}
\mathbb{E}\left[\mathbb{E}\left[e^{z\,X}\middle|\,N\right]\right] &= \mathbb{E}\left[(1-z)^{-\delta}\left(1-\frac{q}{1-q}\,\frac{z}{1-z}\right)^{-r-N}\right] \\
&= (1-z)^{-\delta}\left(1-\frac{q}{1-q}\,\frac{z}{1-z}\right)^{-r}\left[\frac{1-q_{mix}}{1-q_{mix}\left(1-\frac{q}{1-q}\,\frac{z}{1-z}\right)^{-1}}\right]^{r} \\
&= (1-z)^{-\delta}\left[1-\frac{q}{(1-q)(1-q_{mix})}\,\frac{z}{1-z}\right]^{-r} \\
&= (1-z)^{-\delta}\left(1-\frac{q'}{1-q'}\,\frac{z}{1-z}\right)^{-r}
\end{aligned}$$

which coincides with the moment generating function of a generalized non-central gamma with over-dispersion parameter $q' = \dfrac{q}{q+(1-q)(1-q_{mix})}$.

C.2 Properties of the Autoregressive Generalized Gamma Process

C.2.1 Proof of Proposition 11

Consider $q \in (0,1)$ and $z \leqslant 0$, note that

$$\begin{aligned}
a_i'(z_i) &= \theta\beta_i\frac{1-q}{q}\left(\frac{1}{1-q-\theta z_i}-\frac{1}{1-\theta z_i}\right) \leqslant 0 \\
a_i''(z_i) &= \theta\beta_i\frac{1-q}{q}\left[(1-q-\theta z_i)^{-2}-(1-\theta z_i)^{-2}\right] \geqslant 0
\end{aligned}$$

For $z_i \leqslant 0$ we have $a_i(0)\,z_i + z_i < A_i(\mathbf{z}) \leqslant 0$ for $i = 1,\ldots,q-1$ and $A_p(\mathbf{z}) \leqslant -a_p'(0)\,z_p$ where

$a_i'(0) = \theta\beta_i$. Now it is sufficient to prove that $\lim_{n\to\infty} \Phi^n = 0$ where

$$\Phi = \begin{pmatrix} \theta\beta_1 & \theta\beta_2 & \dots & \theta\beta_{p-1} & \theta\beta_p \\ 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 \end{pmatrix}'$$

the condition is satisfied iff $|\Phi| < 1$ which in turn requires the polynomial $1 - \theta\beta_1 x - \ldots - \theta\beta_p x^p$ to have all the roots located within the unit circle. This is guaranteed when the condition $\theta \sum_{i=1}^{q} \beta_i < 1$ holds. Note that $\Phi = \frac{\partial A}{\partial \mathbf{z}'}\Big|_{\mathbf{z}=\mathbf{0}}$, while a_i and b (equations 3.14, 3.15) satisfy conditions $A.1$ and $A.2$ in Darolles et al. (2006), thus under the same condition the process is also geometrically ergodic.

C.2.2 Proof of Proposition 12

Given that the *ARGG* process belongs to the Car class, the mean and the variance conditionally on any lagged realization of Y are linear functions of the state variable

$$\mathbb{E}(Y_t|Y_{t-h}) = \mu_h^0 + \mu_h^1 Y_{t-h} \tag{C.3}$$

$$Var(Y_t|Y_{t-h}) = \upsilon_h^0 + \upsilon_h^1 Y_{t-h} \tag{C.4}$$

where

$$\begin{aligned} \mu_1^0 &= \theta(\delta + d) \\ \mu_1^1 &= \rho \\ \upsilon_1^0 &= \theta^2\left(\delta + \frac{2-q}{1-q}d\right) \\ \upsilon_1^1 &= \theta\rho\frac{2-q}{1-q} \end{aligned}$$

the relations:

$$\begin{aligned} \mathbb{E}(Y_t|Y_{t-h}) &= \mathbb{E}\left[\mathbb{E}(Y_t|Y_{t-h+1})|Y_{t-h}\right] \\ Var(Y_t|Y_{t-h}) &= \mathbb{E}\left[Var(Y_t|Y_{t-h+1})|Y_{t-h}\right] + Var\left[\mathbb{E}(Y_t|Y_{t-h+1})|Y_{t-h}\right] \end{aligned}$$

lead to the following recursive formulas

$$\begin{aligned}
\mu_h^0 &= \mu_{h-1}^0 + \mu_{h-1}^1 \mu_1^0 \\
\mu_h^1 &= \mu_{h-1}^1 \mu_1^1 \\
\upsilon_h^0 &= \upsilon_{h-1}^0 + \mu_1^0 \upsilon_{h-1}^1 + \left(\mu_{h-1}^1\right)^2 \upsilon_1^0 \\
\upsilon_h^1 &= \upsilon_{h-1}^1 \mu_1^1 + \left(\mu_{h-1}^1\right)^2 \upsilon_1^1
\end{aligned}$$

The solutions are

$$\begin{aligned}
\mu_h^0 &= \mu_1^0 \frac{1-\left(\mu_1^1\right)^h}{1-\mu_1^1} \\
\mu_h^1 &= \left(\mu_1^1\right)^h \\
\upsilon_h^0 &= \frac{\mu_0^1 \upsilon_1^1}{\mu_1^1\left(1-\mu_1^1\right)}\left[\frac{1-\left(\mu_1^1\right)^h}{1-\mu_1^1} - \frac{1-\left(\mu_1^1\right)^{2h}}{1-\left(\mu_1^1\right)^2}\right] + \upsilon_1^0 \frac{1-\left(\mu_1^1\right)^{2h}}{1-\left(\mu_1^1\right)^2} \\
\upsilon_h^1 &= \upsilon_1^1 \left(\mu_1^1\right)^{h-1} \frac{1-\left(\mu_1^1\right)^h}{1-\mu_1^1}
\end{aligned}$$

substituting we find expressions 3.18 and 3.19. The unconditional mean and variance are obtained taking the limit $h \to \infty$.

C.3 Computation of the Generalized Non-central Gamma Density Function

In this appendix we describe different computational methods for the calculation of the non-central generalized gamma density with an adequate level of accuracy. We start with the simplest case of a single component, then we move to the more general problem of two independent components according to definitions 3.32 and 3.35.

C.3.1 Mixture of Gamma Distribution

A simple numerical approach to compute the generalized non-central chi squared density consists in the explicit summation of the series 3.6. Two problems must be addressed: i) determining the appropriate truncation to achieve the desired level of accuracy, ii) managing the case $p \to 0$ when the negative binomial converges to the Poisson distribution. The following proposition allows to

address the first problem:

Proposition 17. Let $\{A_n\}_n$ such that $A_n \in \mathbb{R}$, $A_n > 0\ \forall n \in \mathbb{N}$ and for $n > N$, $A_{n+1}/A_n \leqslant C/n$. Then

$$\sum_{n=N+1}^{\infty} A_n \leqslant \frac{A_N e^C C}{N+1}$$

Proof From the assumptions, the following inequality holds:

$$\sum_{n=N+1}^{\infty} A_n \leqslant A_N N! C^{-N} \sum_{n=N+1}^{\infty} \frac{C^n}{n!}$$

the series on the RHS is the reminder of the Taylor series expansion for the function e^C around the origin truncated at the N-th term. Thus from the Lagrange's reminder formula, for some $x \in [0, C]$ we have

$$\sum_{n=N+1}^{\infty} \frac{C^n}{n!} = e^x \frac{x^{N+1}}{(N+1)!} \leqslant e^C \frac{C^{N+1}}{(N+1)!}$$

which leads to the result.

Now assume for simplicity $\theta = 1$. For $k \geqslant 0$ we have

$$\begin{aligned} \frac{nb(k+1; r, q)}{nb(k; r, p)} &= q \frac{k + r(\lambda, q)}{k+1} \\ &= q \frac{k}{k+1} + \frac{\lambda(1-q)}{k+1} \end{aligned} \tag{C.5}$$

where we used the relation $r(\lambda) = \lambda(1-q)/q$. Note that expression C.5 is well behaved for small values of q. The first term in the summation, for small values of q, can be approximated as follows:

$$\begin{aligned} nb(k=0; r, q) &= (1-q)^{\lambda(1-q)/q} \\ &= \exp\left\{\lambda \frac{1-q}{q} \log(1-q)\right\} \\ &= \exp\left\{-\lambda(1-q)\left(1 - \frac{q}{2} + \frac{q^2}{3}\right) + O\left(q^4\right)\right\} \end{aligned}$$

To determine the appropriate truncation of the sum, let A_k denote the k-th term in the summation:

$$\begin{aligned} A_k &= nb\left(k; r, q\right) \frac{e^{-x} x^{\delta+k-1}}{\Gamma\left(\delta+k\right)} \\ &= \frac{\left(1-q\right)^r}{\Gamma\left(\delta\right)} e^{-x} x^{\delta-1} \frac{\left(xq\right)^k}{k!} \frac{r\left(r+1\right)\ldots\left(r+k-1\right)}{\delta\left(\delta+1\right)\ldots\left(\delta+k-1\right)} \end{aligned} \tag{C.6}$$

we have

$$\frac{A_{k+1}}{A_k} = \left[q\,\frac{k}{k+1} + \frac{\lambda\left(1-q\right)}{2\left(k+1\right)}\right] \frac{x}{\delta+k} \leqslant \frac{c_N\, x}{\left[\delta\right]+k} \tag{C.7}$$

where $\left[\delta\right]$ is the integer part of δ and

$$c_N = \left[q + \frac{\lambda\left(1-q\right)}{N+1}\right]$$

Then 17 applies and we have

$$\sum_{k=N+1}^{\infty} A_k \leqslant \frac{A_N\, e^{x\, c_N}\, x\, c_N}{\left[\delta\right]+N+1}$$

note that for $N \to \infty$, $c_N \to q$, thus the convergence is much faster for the non-central gamma characterized by $q = 0$. This explicit summation is efficient only for small values of x. In other cases the Fourier inversion can be more efficient (see Appendix C.3.2).

C.3.2 Computation of the Generalized Non-central Gamma Density with Two Components

Here we address the general problem of finding fast and accurate computational methods to calculate the non-central generalized gamma density in the mots difficult case when the distribution includes two independent components with a different scale. We consider three alternative methods: in Appendix C.3.2 we derive a gamma mixture representation that allows an expansion similar to Appendix C.3.1, in Appendix C.3.2 we elaborate a different expansion starting from equation 3.32 and in Appendix C.3.2 we propose an algorithm for Fourier inversion.

C.3. COMPUTATION OF THE GENERALIZED NON-CENTRAL GAMMA DENSITY FUNCTION

Mixture of gamma distributions

Let $r = \lambda (1-q)/q$ and $u = (1-z)^{-1}$, we have

$$\left[\frac{1-q}{1-q/(1-\theta z)}\right]^r = (1-q)^r \left[\frac{1+u(1-\theta)/\theta}{1+u(1-q-\theta)/\theta}\right]^r$$

From 3.35, assuming for simplicity $\theta_1 = 1$, the moment generating function of the generalized non-central gamma with two components can be written as:

$$\begin{aligned}\hat{\psi}_2(z;\, \delta, \theta_1, \lambda_1, q_1, \theta_2, \lambda_2, q_2) &= (1-z)^{-\delta}(1-q_1)^{r_1}(1-q_1)^{r_1}\left(1-\frac{q_1}{1-z}\right)^{-r_1} \\ &\quad \left[1-\frac{\theta_2-1}{\theta_2(1-z)}\right]^{r_2}\left[1-\frac{\theta_2-1+q_2}{\theta_2(1-z)}\right]^{-r_2} \quad z \in \mathbb{C} \qquad \text{(C.8)}\end{aligned}$$

For $1-z > \max\{q_1, (\theta_2 - 1 + p_2)/\theta_2\}$ this expression can be expanded as

$$\hat{\psi}_2(z;\, \delta, \theta_1, \lambda_1, q_1, \theta_2, \lambda_2, q_2) = \sum_{k=0}^{\infty} a_k (1-z)^{-\delta-k}$$

thus the resulting density can be expanded as mixture of gamma distributions with unitary scale

$$\psi_2(x;\, \delta, \theta_1, \lambda_1, q_1, \theta_2, \lambda_2, q_2) = \sum_{k=0}^{\infty} a_k \gamma(x; \delta + k, 1) \qquad \text{(C.9)}$$

the coefficients can be computed numerically by inversion of the power transform

$$Q(u) = (1-q_1)^{r_1}(1-q_2)^{r_2}(1-q_1 u)^{-r_1}\left(1-\frac{\theta_2-1}{\theta_2}u\right)^{r_2}\left(1-\frac{\theta_2-1+q_2}{\theta_2}u\right)^{-r_2}$$

the first three terms of the expansion are

$$\begin{aligned} a_0 &= (1-q_1)^{r_1}(1-q_2)^{r_2} \\ a_1 &= a_0\left[\lambda_1(1-q_1)+\lambda_2\frac{1-q_2}{\theta_2}\right] \\ a_2 &= \frac{a_0}{2}\Bigg\{\lambda_1(1-q_1)\left[\lambda_1(1-q_1)+q_1\right]+ \\ &\qquad 2\lambda_2\frac{1-q_2}{\theta_2}\left[\lambda_1(1-q_1)+1\right]+\lambda_2(1-q_2)\frac{\lambda_2(1-q_2)+q_2-2}{\theta_2^2}\Bigg\} \end{aligned}$$

Alternative Expansion

For simplicity we consider $\theta_1 = 1$. We start from the representation 3.32:

$$\begin{aligned} \psi_2(x;\,\delta,\theta_1,\lambda_1,q_1,\theta_2,\lambda_2,q_2) &= (1-q_2)^{r_2}\,\psi(x;\,\delta,\theta_1,\lambda_1,q_1)+ \\ &\quad \sum_{n=0}^{\infty}\sum_{m=1}^{\infty} nb(n;r_1,q_1)\; nb(m;r_2,q_2) \\ &\quad \int_0^x \gamma(u;\,\delta+n,\theta_1)\,\gamma(x-u;\,m,\theta_1) \end{aligned} \tag{C.10}$$

The integral inside the summation for $m > 0$ is:

$$\begin{aligned} \int_0^x \gamma(u;\,\delta+n,\theta_1)\,\gamma(x-u;\,m,\theta_2) &= \frac{e^{-x/\theta_J}}{\Gamma(\delta+n)\,\Gamma(m)\,\theta_2^m}\int_0^x e^{u(1/\theta_2-1)}u^{\delta+n-1}(x-u)^{m-1}\,du \\ &= \frac{e^{-x/\theta_2}x^{\delta+n+m-1}}{\Gamma(\delta+n)\,\Gamma(m)!\theta_2^m}\int_0^1 e^{\eta x t}t^{\delta+n-1}(1-t)^{m-1}\,dt \\ &= \frac{e^{-x/\theta_2}x^{\delta+n+m-1}}{\theta_2^m\Gamma(\delta+n+m)}\hat{\beta}(\eta x;\delta+n,m) \end{aligned}$$

where $\eta = 1/\theta_2 - 1$, $B(x,y) = \Gamma(x)\,\Gamma(y)/\Gamma(x+y)$ and $\hat{\beta}(u;a,b)$ is the moment generating function of the beta distribution with parameters a, b:

$$\hat{\beta}(u;a,b) = \frac{\int_0^1 e^{ut}t^{a-1}(1-t)^{b-1}\,dt}{B(a,b)} = 1+\sum_{k=1}^{\infty}\frac{u^k}{k!}\prod_{l=0}^{k-1}\frac{a+l}{a+b+l} \tag{C.11}$$

Thus

$$\begin{aligned}\psi_2\left(x;\,\delta,\theta_1,\lambda_1,q_1,\theta_2,\lambda_2,q_2\right) &= (1-q_2)^{r_2}\,\psi\left(x;\,\delta,\theta_1,\lambda_1,q_1\right)+\\ &\quad\sum_{n=0}^{\infty}\sum_{m=1}^{\infty} nb\left(n;r_1,q_1\right)\,nb\left(m;r_2,q_2\right)\cdot\\ &\quad\theta_2^n\gamma\left(x;\,\delta+n+m,\theta_J\right)\tilde{\beta}\left(\eta x;\delta+n,m\right)\end{aligned}\tag{C.12}$$

Note that for $a,b>0$ and $u\geqslant 0$ the following inequalities hold:

$$1\leqslant\tilde{\beta}\left(u;a,b\right)\leqslant e^u \text{ for } u\geqslant 0\qquad e^{-u}\leqslant\tilde{\beta}\left(u;a,b\right)\leqslant 1\quad\text{for } u<0\tag{C.13}$$

moreover it can be easily shown that $\tilde{\beta}$ enjoys the following properties

$$\tilde{\beta}\left(u;a,b\right)=e^u\tilde{\beta}\left(-u;b,a\right)\tag{C.14}$$

$$\tilde{\beta}\left(u;a=0,b\right)=1\quad\text{and}\quad\tilde{\beta}\left(u;a,b=1\right)=e^u\tag{C.15}$$

$$\tilde{\beta}\left(u;a+1,b+1\right)=\frac{a+b+1}{u}\left[\tilde{\beta}\left(u;a+1,b\right)-\tilde{\beta}\left(u;a,b+1\right)\right]\tag{C.16}$$

$\tilde{\beta}$ can be computed numerically. To truncate the summation properly note that for $u\geqslant 0$ we have:

$$\sum_{k=n+1}^{\infty}\frac{u^k}{k!}\prod_{l=0}^{k-1}\frac{a+l}{a+b+l}\leqslant\sum_{k=n+1}^{\infty}\frac{u^k}{k!}\leqslant\frac{e^u u^{n+1}}{(n+1)!}$$

for $u<0$ the we can use relation C.14. From the representation C.11 we see that for $u>0$, $\tilde{\beta}\left(u;a,b\right)$ is strictly decreasing in b. Using property C.14 we conclude that for $u<0$, $\tilde{\beta}\left(u;a,b\right)$ is strictly decreasing in a.

Adopting the definition C.11 and thanks to properties C.15 the density can be represented in the compact form:

$$\psi_2\left(x;\,\delta,\theta_1,\lambda_1,q_1,\theta_2,\lambda_2,q_2\right)=e^{-x/\theta_2}x^{\delta-1}\sum_{m=0}^{\infty}nb\left(m;r_2,q_2\right)\left(\frac{x}{\theta_2}\right)^m\sum_{n=0}^{\infty}\frac{nb\left(n;r_1,q_1\right)}{\Gamma\left(\delta+n+m\right)}x^n\tilde{\beta}\left(x\eta;\delta+n,m\right)\tag{C.17}$$

The summation has the form $\sum_{m=0}^{\infty}B_m\sum_{n=0}^{\infty}A_{m,n}$. Under the assumption $\theta_2\geqslant 1$ (i.e. $\eta<0$) we

have:

$$\frac{A_{m,n+1}}{A_{m,n}} = \left[q_1 \frac{n}{n+1} + \frac{\lambda_1 (1-q_1)}{n+1}\right] \frac{x\tilde{\beta}(x\eta; \delta+n+1, m)}{(\delta+n+m)\,\tilde{\beta}(x\eta; \delta+n, m)} \leqslant \frac{c_N^{\nu} x}{\lfloor\delta\rfloor + m + n}$$

where

$$c_N^1 = q_1 + \frac{\lambda_1 (1-q_1)}{n+1}$$

therefore proposition 17 of Appendix C.3.1 applies:

$$\sum_{k=N+1}^{\infty} A_{m,k} \leqslant \frac{A_{m,N}\, e^x x}{\lfloor\delta\rfloor + m + N + 1} \quad \text{for } N \geqslant \max(\lambda_1, x)$$

Let us now concentrate on the summation over m. Let $u > 0$, the following relations hold:

$$\begin{aligned}
\tilde{\beta}(-u; a, b) - \frac{\tilde{\beta}(-u; a, b+1)}{a+b} &= e^{-u}\left[\tilde{\beta}(u; b, a) - \frac{\tilde{\beta}(u; b+1, a)}{a+b}\right] \\
&= e^{-u}\left[1 - \frac{1}{a+b} + \sum_{k=1}^{\infty} \frac{u^k}{k!}\left(\frac{b}{a+b} - \frac{1}{a+b}\frac{b+k}{a+b+k}\right)\prod_{l=1}^{k-1}\frac{b+l}{a+b+l}\right] \\
&\geqslant e^{-u}\left[1 - \frac{1}{a+b} + \sum_{k=1}^{\infty} \frac{u^k}{k!}\left(\frac{b-1}{a+b}\right)\prod_{l=1}^{k-1}\frac{b+l}{a+b+l}\right] \\
&= e^{-u}\left[1 - \frac{1}{a+b} + \frac{b-1}{b}\sum_{k=1}^{\infty} \frac{u^k}{k!}\prod_{l=0}^{k-1}\frac{b+l}{a+b+l}\right] \\
&= e^{-u}\left\{1 - \frac{1}{a+b} + \left(1 - \frac{1}{b}\right)\left[\tilde{\beta}(u; b, a) - 1\right]\right\} \\
&= \frac{e^{-u} a}{b(a+b)} + \left(1 - \frac{1}{b}\right)\tilde{\beta}(-u; a, b)
\end{aligned}$$

thus

$$\frac{\tilde{\beta}(-u; a, b+1)}{a+b} \leqslant \frac{1}{b}\left[\tilde{\beta}(-u; a, b) - \frac{e^{-u} a}{a+b}\right]$$

now consider $a = \delta + n$, $b = m$ and $u = -\eta x$ we have:

$$\sum_{n=0}^{\infty} A_{n,m+1} \leqslant \frac{1}{m}\sum_{n=0}^{\infty} A_{n,m}$$

follows that in the series C.17 the ratio between two consecutive terms w.r.t. the index $m > 0$ is bounded as follows:

$$0 < \frac{B_{m+1} \sum_{n=0}^{\infty} A_{n,m+1}}{B_m \sum_{n=0}^{\infty} A_{n,m}} \leqslant \left[q_2 \frac{m}{m+1} + \frac{\lambda_2 (1-q_2)}{m+1} \right] \left(\frac{x}{m \theta_J} \right) \qquad \forall x > 0$$

Again proposition 17 applies:

$$\sum_{m=M+1}^{\infty} B_m \sum_{n=0}^{\infty} A_{n,m} \leqslant \frac{e^{C_M x/\theta_2} C_M x/\theta_2}{M+1} B_M \sum_{n=0}^{\infty} A_{n,M} \text{ for } M > 0$$

where $C_M = \left[q_2 + \frac{\lambda_J (1-q_2)}{M+1} \right]$

Fourier inversion

The task consists in the numerical evaluation of the contour integral

$$\psi_2 (x) = \frac{1}{2\pi\imath} \int_{z_0 - \imath\infty}^{z_0 + \imath\infty} \hat{\psi}_2 (z) \, dz$$

where $\hat{\varphi}$ is defined in 3.35. For simplicity we assume $\theta_1 = 1$. We consider the equivalent problem of evaluating the Bromwich integral:

$$\psi_2 (x) = \frac{2e^{-z_0}}{\pi} \int_0^{+\infty} Re \left\{ \hat{\psi}_2 (z_0 + \imath t) \right\} \cos (tx) \, dt \tag{C.18}$$

the integration can be performed using the trapezoidal rule with an incremental step $h > 0$

$$\psi_2 (x) \simeq \hat{\psi}_{2,h} (x) \equiv h \frac{e^{-z_0 x}}{\pi} Re \left\{ \hat{\psi}_2 (z_0) \right\} + h \frac{e^{-z_0 x}}{\pi} \sum_{k=1}^{\infty} Re \left\{ \hat{\psi}_2 (z_0 + \imath t) \right\} \cos (khx) \tag{C.19}$$

We need to control three types of error: the discretization (or aliasing) error, the truncation error and the round off error. Concerning the truncation error we use the Euler summation method to accelerate the convergence as suggested by Abate et al. (2000), the main reference for this technical appendix. In the reminder of this appendix will focus on the control of the aliasing and the round off errors. Importantly, the trapezoidal rule reveals particularly efficient when applied to Fourier inversion problems since the aliasing error, usually of order $O\left(h^2\right)$, decreases at much higher rate

in this case. More precisely, it can be proved that the aliasing error is given by

$$\varepsilon_h(x) = \sum_{k=1}^{\infty} \psi_2\left(x + \frac{2\pi k}{h}\right) e^{z_0 2\pi k/h} \text{ for } h < 2\pi/x \tag{C.20}$$

Following Abate et al., we conveniently consider $h = \pi/xL$ with L positive integer. If an upper bound $C_L(x)$ can be found such that $\phi(x') \leqslant C_L(x)$ for $x' \geqslant x(1+2L)$, then for any $z_0 < 0$ we have

$$\varepsilon_h(x) \leqslant C_L(x) \sum_{k=1}^{\infty} e^{k \cdot 2 z_0 L x} = C_L(x) \frac{e^{2z_0 L x}}{1 - e^{2z_0 L x}} \tag{C.21}$$

It is relatively easy to find an upper bound $C_L(x)$ for a given $x > 0$ and to make the aliasing error arbitrarily small by choosing a negative an sufficiently large value of z_0. Unfortunately, round off errors can explode due to the pre-factor $e^{-z_0 x}$ in C.19, especially when x is in the tail of the distribution. Abate et al. suggest how to optimally set z_0 in order to minimize both type of errors but their approach allows to reach a level of accuracy that is of the same order of the machine precision. Our purpose is to compute the density accurately also in the tail, thus we want to control the relative rather then absolute errors. Note that for $x \to \infty$, $\psi_2(x) \propto e^{-\xi x}$ where $\xi = \min\{1 - q_1, (1-q_2)/\theta_2\}$ as we can see from the singularities of the moment generating function. Thus considering a constant upper bound for φ is too conservative, especially for x large. The idea is to fully exploit equation C.20 taking into account also the tail decay of the density. In this case z_0 can be allowed to take also positive values permitting to reduce round off errors. More precisely, we have

$$\varepsilon_h(x) = \sum_{k=1}^{\infty} \psi_2(x + 2kLx) e^{2z_0 kLx}$$

assuming that Lx is sufficiently large and $z_0 < \xi$, this expression is approximated by

$$\begin{aligned} \varepsilon_h(x) &\simeq \psi_2(x + 2Lx) e^{2z_0 Lx} \sum_{k=0}^{\infty} e^{2(z_0 - \xi)kLx} \\ &= \frac{\psi_2(x + 2Lx) e^{2z_0 Lx}}{1 - e^{2(z_0 - \xi)kLx}} \end{aligned} \tag{C.22}$$

Concerning the round off error, assume that the machine precision is 10^{-q} where $q = 16$ if we work

with double precision. To estimate the order of magnitude for this error, consider the sum in C.19: the first term is $\hat{\psi}_2(z_0)$ computed with a relative error of order 10^{-q}. The round off error is therefore of order

$$\epsilon_R(x) \simeq \frac{\hat{\psi}_2(z_0)}{Lx} e^{-z_0 x - q\log 10} \tag{C.23}$$

Note that the summation involves terms with alternating sign and the relative round off error can underestimated by expression C.23 when the rate of decay of those terms is very slow (which can happen for large negative value of z_0), but the exponential factor remains the main driver. The round off errors are minimized at the saddle point z^* solving the equation $g'(z^*) - x = 0$ with $g(z) = \log \hat{\psi}_2(z)$ and $g''(z^*) \geqslant 0$, $z \in \mathbb{R}$. Our purpose is to minimize the sum of the two errors. For $z < z^*$the round off error is decreasing in z while the aliasing error increases, their sum is approximately minimized when they are of the same order: $\epsilon_R(x) \simeq \varepsilon_h(x)$ which leads to the equation

$$z_0 = \frac{g(z_0) - \log(Lx) - q\log 10 - \log\left[1 - e^{2(z_0-\xi)kLx}\right] - \log\psi_2(x + 2Lx)}{(2L+1)x} \tag{C.24}$$

let us first consider this equation in the limit $x \to \infty$, $\log\psi_2(x + 2Lx) \simeq A - \xi x(1 + 2L) + (r^* - 1)\log x$ where

$$\begin{array}{lllll} \text{for } 1-q_1 & < & (1-q_2)/\theta_2 & r^* = r(\lambda_1) & \xi = 1 - q_1 \\ \text{for } 1-q_1 & > & (1-q_2)/\theta_2 & r^* = r(\lambda_2) & \xi = (1-q_2)/\theta_2 \\ \text{for } 1-q_1 & = & (1-q_2)/\theta_2 & r^* = r(\lambda_1) + r(\lambda_2) & \xi = 1 - q_1 \end{array}$$

an approximated solution to C.24 for x large is

$$z_0(x) \simeq \xi - \frac{\log(Lx)}{(2L+1)x} \tag{C.25}$$

this can be checked by direct substitution taking into account that as $z \to \xi$, $g(z) \simeq -r^* \log(\xi - z)$ and dropping all terms going to zero faster than $\log x/x$. Substituting C.25 into C.23 we find

$$\varepsilon_h(x) \simeq \epsilon_R(x) \propto x^{r^* - \frac{2L}{2L+1}} e^{-\xi x - q\log 10}$$

Note also that the saddle point for $x \to \infty$ is approximately $z^*(x) = \xi - r^*/x$, thus for x sufficiently large $z^* < z_0$ and we conclude that choosing z_0 as a solution of equation C.24 is suboptimal in this case because both errors can be reduced by decreasing z_0. The best choice in this case is $z_0 = z^*$and we have

$$\varepsilon_h(x) < \epsilon_R(x) \propto \frac{1}{L}\left(\frac{e}{r^*}\right)^{r^*} x^{r^*-1} e^{-\xi x - q\log 10}$$

The round off error dominates for x large while the relative error is constant an proportional to the machine precision. Choosing a large value of L has a marginal impact on the error for x large, thus it can be convenient to set $L = 1$ to save computational time. Note that equation C.24 contains the distribution ψ_2 that we want to evaluate, hence we need to approximate the density in order to use this equation in practice. To this purpose we use the saddle point approximation which works well for values of x that are not too small (see Section C.1.1 for further details). In practice we rely on this proxy for $x(1+L)$ larger than the mean which implies $L > (\mathbb{E}[x]/x - 1)/2$.

Equation C.24 can be solved numerically, for example by iteration. Anyway solutions are not guaranteed to exist in general: it is in fact possible that the round off error dominates for all admissible values of z_0, the optimal choice in this case is $z_0 = z^*$.

Bibliography

Abate, J., Choudhury, G. L., and Whitt, W. (2000), *Computational Probability*, chap. An Introduction to Numerical Transform Inversion and Its Application to Probability Models, pp. 257–323, Springer US, Boston, MA.

Adrian, T. and Rosenberg, J. (2008), "Stock Returns and Volatility: Pricing the Short-Run and Long-Run Components of Market Risk," *The Journal of Finance*, 63, 2997–3030.

Ahoniemi, K. and Lanne, M. (2013), "Overnight stock returns and realized volatility," *International Journal of Forecasting*, 29, 592 – 604.

Ait-Sahalia, Y. (2004), "Disentangling diffusion from jumps," *Journal of Financial Economics*, 74, 487–528.

Ait-Sahalia, Y. and Hurd, T. R. (2015), "Portfolio Choice in Markets with Contagion," *Journal of Financial Econometrics*, 14, 1–28.

Aït-Sahalia, Y. and Jacod, J. (2009), "Testing for jumps in a discretely observed process," *Annals of Statistics*, 37, 184–222.

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

Alfonsi, A. (2010), *Cox-Ingersoll-Ross (CIR) Model*, John Wiley & Sons, Ltd.

Alitab, D., Majewski, A. A., Bormetti, G., and Corsi, F. (2016), "A Jump and Smile Ride: Continuous and Jump Variance Risk Premia in Option Pricing," Available at ssrn.

Andersen, T., Bollerslev, T., and Huang, X. (2011), "A reduced form framework for modeling

volatility of speculative prices based on realized variation measures," *Journal of Econometrics*, 160, 176–189.

Andersen, T. G. and Bollerslev, T. (1997), "Intraday periodicity and volatility persistence in financial markets," *Journal of Empirical Finance*, 4, 115–158.

Andersen, T. G., Benzoni, L., and Lund, J. (2002), "An Empirical Investigation of Continuous-Time Equity Return Models," *Journal of Finance*, 57, 1239–1284.

Andersen, T. G., Bollerslev, T., and Dobrev, D. (2007a), "No-arbitrage semi-martingale restrictions for continuous-time volatility models subject to leverage effects, jumps and i.i.d. noise: Theory and testable distributional implications," *Journal of Econometrics*, 138, 125 – 180, 50th Anniversary Econometric Institute.

Andersen, T. G., Bollerslev, T., and Diebold, F. X. (2007b), "Roughing It Up: Including Jump Components in the Measurement, Modeling, and Forecasting of Return Volatility," *The Review of Economics and Statistics*, 89, 701–720.

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

Andersen, T. G., Bollerslev, T., and Diebold, F. X. (2010b), "Parametric and Nonparametric Volatility Measurement," in *Handbook of Financial Econometrics*, eds. Y. Ait-Sahalia and L. P. Hansen, Elsevier.

Andersen, T. G., Dobrev, D., and Schaumburg, E. (2012), "Jump-robust volatility estimation using nearest neighbor truncation," *Journal of Econometrics*, 169, 75–93.

Bajgrowicz, P., Scaillet, O., and Treccani, A. (2016), "Jumps in High-Frequency Data: Spurious Detections, Dynamics, and News," *Management Science*, 62, 2198–2217.

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

Barndorff-Nielsen, O. E. and Shephard, N. (2004), "Power and Bipower Variation with Stochastic Volatility and Jumps," *Journal of Financial Econometrics*, 2, 1–37.

Barndorff-Nielsen, O. E. and Shephard, N. (2006), "Econometrics of Testing for Jumps in Financial Economics Using Bipower Variation," *Journal of Financial Econometrics*, 4, 1–30.

Barndorff-Nielsen, O. E., Shephard, N., and Winkel, M. (2006), "Limit theorems for multipower variation in the presence of jumps," *Stochastic Processes and their Applications*, 116, 796–806.

Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A., and Shephard, N. (2009), "Realized kernels in practice: trades and quotes," *Econometrics Journal*, 12, C1–C32.

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

Bates, D. S. (2006), "Maximum Likelihood Estimation of Latent Affine Processes," *Review of Financial Studies*, 19, 909–965.

Bates, D. S. (2008), "The market for crash risk," *Journal of Economic Dynamics and Control*, 32, 2291–2321.

Black, F. (1976), "Studies of stock price volatility changes," in *Proceedings of the 1976 Meetings of the American Statistical Association, Business and Economics Statistics Section*, pp. 177–181.

Bollerslev, T. and Todorov, V. (2011), "Tails, Fears, and Risk Premia," *The Journal of Finance*, 66, 2165–2211.

Bollerslev, T. and Zhou, H. (2002), "Estimating stochastic volatility diffusion using conditional moments of integrated volatility," *Journal of Econometrics*, 109, 33 – 65.

Bollerslev, T., Litvinova, J., and Tauchen, G. (2006), "Leverage and Volatility Feedback Effects in High-Frequency Data," *Journal of Financial Econometrics*, 4, 353–384.

Bollerslev, T., Todorov, V., and Li, S. Z. (2013), "Jump tails, extreme dependencies, and the distribution of stock returns," *Journal of Econometrics*, 172, 307–324.

Bormetti, G., Calcagnile, L. M., Treccani, M., Corsi, F., Marmi, S., and Lillo, F. (2015), "Modelling systemic price cojumps with Hawkes factor models," *Quantitative Finance*, 15, 1137–1156.

Bormetti, G., Casarin, R., Corsi, F., and Livieri, G. (2016), "Smile at Errors: A Discrete-Time Stochastic Volatility Framework for Pricing Options with Realized Measures," Available at SSRN.

Boudt, K., Croux, C., and Laurent, S. (2011), "Robust estimation of intraweek periodicity in volatility and jump detection," *Journal of Empirical Finance*, 18, 353 – 367.

Bowsher, C. G. (2007), "Modelling security market events in continuous time: Intensity based, multivariate point process models," *Journal of Econometrics*, 141, 876 – 912.

Calcagnile, L. M., Bormetti, G., Treccani, M., Marmi, S., and Lillo, F. (2015), "Collective synchronization and high frequency systemic instabilities in financial markets," Working Paper.

Caporin, M., Kolokolov, A., and Renó, R. (2014), "Multi-jumps," "Marco Fanno" Working Papers 0185, Dipartimento di Scienze Economiche "Marco Fanno".

Chan, W. and Maheu, J. (2002), "Conditional Jump Dynamics in Stock Market Returns," *Journal of Business & Economic Statistics*, 20, 377–89.

Chen, K. and Poon, S.-H. (2013), "Variance swap premium under stochastyic volatility and self-exciting jumps," Available at SSRN.

Chen, X., Ghysels, E., and Wang, F. (2011), "HYBRID GARCH Models and Intra-Daily Return Periodicity," *Journal of Time Series Econometrics*, 3, 1–28.

Chernov, M., Ronald Gallant, A., Ghysels, E., and Tauchen, G. (2003), "Alternative models for stock price dynamics," *Journal of Econometrics*, 116, 225–257.

Christensen, K., Oomen, R., and Podolskij, M. (2010), "Realised quantile-based estimation of the integrated variance," *Journal of Econometrics*, 159, 74 – 98.

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

Christensen, K., Oomen, R., and Renò, R. (2016), "The Drift Burst Hypothesis," Available at ssrn.

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

Christoffersen, P., Jacobs, K., and Ornthanalai, C. (2012), "Dynamic jump intensities and risk premiums: Evidence from S&P500 returns and options," *Journal of Financial Economics*, 106, 447–472.

Christoffersen, P., Feunou, B., Jacobs, K., and Meddahi, N. (2014), "The Economic Value of Realized Volatility: Using High-Frequency Returns for Option Valuation," *Journal of Financial and Quantitative Analysis*, 49, 663–697.

Christoffersen, P., Feunou, B., and Jeon, Y. (2015), "Option valuation with observable volatility and jump dynamics," *Journal of Banking & Finance*, 61, Supplement 2, S101 – S120, Recent Developments in Financial Econometrics and Applications.

Corradi, V., Distaso, W., and Fernandes, M. (2012), "International market links and volatility transmission," *Journal of Econometrics*, 170, 117 – 141.

Corsi, F. (2009), "A Simple Approximate Long-Memory Model of Realized Volatility," *Journal of Financial Econometrics*, 7, 174–196.

Corsi, F. and Renò, R. (2012), "Discrete-Time Volatility Forecasting With Persistent Leverage Effect and the Link With Continuous-Time Volatility Modeling," *Journal of Business & Economic Statistics*, 30, 368–380.

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

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

Cox, J. C., , Ingersoll, J. E., and Ross, S. A. (1985), "A Theory of the Term Structure of Interest Rates," *Econometrica*, 53, 385–407.

Dai, Q. and Singleton, K. J. (2000), "Specification Analysis of Affine Term Structure Models," *The Journal of Finance*, 55, 1943–1978.

Darolles, S., Gourieroux, C., and Jasiak, J. (2006), "Structural Laplace Transform and Compound Autoregressive Models," *Journal of Time Series Analysis*, 27, 477–503.

Duffie, D. and Kan, R. (1996), "A Yield-Factor Model Of Interest Rates," *Mathematical Finance*, 6, 379–406.

Duffie, D. and Pan, J. (1997), "An Overview of Value at Risk," *The Journal of Derivatives*, 4, 7–49.

Duffie, D., Pan, J., and Singleton, K. (2000), "Transform Analysis and Asset Pricing for Affine Jump-Diffusions," *Econometrica*, 68, 1343–1376.

Duffie, D., Filipovic, D., and Schachermayer, W. (2003), "Affine processes and applications in finance," *Annals of Applied Probability*, 13, 984–1053.

Dumitru, A.-M. and Urga, G. (2011), "Identifying Jumps in Financial Assets: A Comparison Between Nonparametric Jump Tests," *Journal of Business & Economic Statistics*, 30, 242–255.

Engle, R. (2002), "Dynamic Conditional Correlation: A Simple Class of Multivariate Generalized Autoregressive Conditional Heteroskedasticity Models," *Journal of Business & Economic Statistics*, 20, 339–50.

Engle, R. and Lee, G. (1999), "A permanent and transitory component model of stock return volatility," in *Cointegration, Causality, and Forecasting: A Festschrift in Honor of Clive W. J. Granger*, R. Engle and H. White.

Eraker, B. (2004), "Do Stock Prices and Volatility Jump? Reconciling Evidence from Spot and Option Prices," *The Journal of Finance*, 59, 1367–1404.

Eraker, B., Johannes, M., and Polson, N. (2003), "The Impact of Jumps in Volatility and Returns," *The Journal of Finance*, 58, pp. 1269–1300.

Filipović, D., Mayerhofer, E., and Schneider, P. (2013), "Density approximations for multivariate affine jump-diffusion processes," *Journal of Econometrics*, pp. –.

Gilder, D., Shackleton, M. B., and Taylor, S. J. (2014), "Cojumps in stock prices: Empirical evidence," *Journal of Banking & Finance*, 40, 443 – 459.

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

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

Granelli, A. and Veraart, A. (2016), "Modelling the variance risk premium of equity indices: the role ofdependence and contagion," *SIAM Journal on Financial Mathematics*, 7, 382–417.

Hansen, P. R. and Lunde, A. (2006), "Realized Variance and Market Microstructure Noise," *Journal of Business & Economic Statistics*, 24, 127–161.

Hansen, P. R., Huang, Z., and Shek, H. H. (2012), "Realized GARCH: a joint model for returns and realized measures of volatility," *Journal of Applied Econometrics*, 27, 877–906.

Harris, L. (1986), "A transaction data study of weekly and intradaily patterns in stock returns," *Journal of Financial Economics*, 16, 99–117.

Hasbrouck, J. (1999), "The Dynamics of Discrete Bid and Ask Quotes," *The Journal of Finance*, 54, pp. 2109–2142.

Hausman, J. A. (1978), "Specification Tests in Econometrics," *Econometrica*, 46, 1251–1271.

Hawkes, A. (1971a), "Point Spectra of Some Mutually Exciting Point Processes," *Journal of the Royal Statistical Society. Series B (Methodological)*, 33, 438–443.

Hawkes, A. (1971b), "Spectra of Some Self-Exciting and Mutually Exciting Point Processes," *Biometrika*, 58, 83–90.

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, 327–343.

Heston, S. L. and Nandi, S. (2000), "A Closed-Form GARCH Option Valuation Model," *Review of Financial Studies*, 13, 585–625.

Hilbe, J. (2011), *Negative Binomial Regression*, Cambridge University Press.

Huang, X. and Tauchen, G. (2005), "The Relative Contribution of Jumps to Total Price Variance," *Journal of Financial Econometrics*, 3, 456–499.

Jacod, J. and Todorov, V. (2010), "Do price and volatility jump together?" *Ann. Appl. Probab.*, 20, 1425–1469.

Jiang, G. J. and Oomen, R. C. (2008), "Testing for jumps when asset prices are observed with noise-a "swap variance" approach," *Journal of Econometrics*, 144, 352–370.

Kolokolov, A. and Renó, R. (2016), "Efficient multipower," Available at SSRN.

Lee, S. S. and Mykland, P. A. (2008), "Jumps in Financial Markets: A New Nonparametric Test and Jump Dynamics," *Review of Financial Studies*, 21, 2535–2563.

Liu, J., Longstaff, F. A., and Pan, J. (2003), "Dynamic Asset Allocation with Event Risk," *The Journal of Finance*, 58, 231–259.

Maheu, J. M. and McCurdy, T. H. (2004), "News Arrival, Jump Dynamics, and Volatility Components for Individual Stock Returns," *The Journal of Finance*, 59, 755–793.

Majewski, A. A., Bormetti, G., and Corsi, F. (2015), "Smile from the past: A general option pricing framework with multiple volatility and leverage components," *Journal of Econometrics*, 187, 521 – 531, Econometric Analysis of Financial Derivatives.

Mancini, C. (2009), "Non-parametric Threshold Estimation for Models with Stochastic Diffusion Coefficient and Jumps," *Scandinavian Journal of Statistics*, 36, 270–296.

Merton, R. C. (1976), "Option pricing when underlying stock returns are discontinuous," *Journal of Financial Economics*, 3, 125 – 144.

Monfort, A. and Pegoraro, F. (2007), "Switching VARMA Term Structure Models," *Journal of Financial Econometrics*, 5, 105–153.

Monfort, A., Pegoraro, F., Renne, J.-P., and Roussellet, G. (2015), "Staying at Zero with Affine Processes: An Application to Term Structure Modelling," Working papers, Banque de France.

Pan, J. (2002), "The jump-risk premia implicit in options: evidence from an integrated time-series study," *Journal of Financial Economics*, 63, 3 – 50.

Pan, J. and Duffie, D. (2001), "Analytical value-at-risk with jumps and credit risk," *Finance and Stochastics*, 5, 155–180.

Patton, A. J. and Sheppard, K. (2009), "Optimal combinations of realised volatility estimators," *International Journal of Forecasting*, 25, 218–238.

Podolskij, M. and Ziggel, D. (2010), "New tests for jumps in semimartingale models," *Statistical Inference for Stochastic Processes*, 13, 15–41.

Rognlie, M. (2010), "Spurious Jump Detection and Intraday Changes in Volatility," Duke university economics honor thesis, Duke University.

Schwert, M. (2010), "Hop, Skip and Jump - What are Modern 'Jump' Tests Finding in Stock Returns?" Available at SSRN.

Shephard, N. and Sheppard, K. (2010), "Realising the future: forecasting with high-frequency-based volatility (HEAVY) models," *Journal of Applied Econometrics*, 25, 197–231.

Stentoft, L. (2008), "Option Pricing using Realized Volatility," Creates research papers, Department of Economics and Business Economics, Aarhus University.

Taylor, S. J. and Xu, X. (1997), "The incremental volatility information in one million foreign exchange quotations," *Journal of Empirical Finance*, 4, 317–340.

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

Tsiakas, I. (2008), "Overnight information and stochastic volatility: A study of European and {US} stock exchanges," *Journal of Banking & Finance*, 32, 251 – 268.

Vasicek, O. (1977), "An equilibrium characterization of the term structure," *Journal of Financial Economics*, 5, 177–188.

Wei, S. (2012), "Simultaneous Occurrence of Price Jumps and Changes in Diffusive price Volatility," Duke university economics honor thesis, Duke University.

Wood, R. A., McInish, T. H., and Ord, J. K. (1985), "An Investigation of Transactions Data for NYSE Stocks," *The Journal of Finance*, 40, pp. 723–739.

Wright, J. H. and Zhou, H. (2009), "Bond risk premia and realized jump risk," *Journal of Banking & Finance*, 33, 2333 – 2345.