Grasping Economic Jumps by Sparse Sampling Using Intradaily Highs and Lows

Stefan Klößner*
Saarland University

February 15, 2011

Preliminary version, comments welcome

Abstract

Economic shocks like unexpected changes in federal funds rate typically cause jumps in economic time series such as asset prices, indices or exchange rates. However, existing high-frequency methods to ex post measure and detect jumps fail to capture these economic jumps. In this paper, we exploit the information embodied in intradaily highs and lows to develop new methods for disentangling volatility contributions originating from diffusive price behaviour and economic jumps. Treating positive and negative jumps separately, we provide estimators and tests for diffusive volatility, positive and negative jumps. Empirically, we find more economic jumps than previously reported.

JEL-classifications: C13, C14, G10
Keywords: Economic Jumps, High-Frequency Data, High-Low-Prices, Integrated Volatility, Sum of Squared Positive (Negative) Jumps

*email: S.Kloessner@mx.uni-saarland.de

1 Introduction

In recent years, measuring and detecting jumps in high-frequency data has been quite a prominent topic in econometrics, testified by the following, albeit incomplete list of publications that all in some form deal with measuring and detecting jumps: Aït-Sahalia & Jacod (2009b,a), Andersen et al. (2007), Andersen et al. (2008), Barndorff-Nielsen et al. (2008a, 2009), Barndorff-Nielsen & Shephard (2004, 2006), Barndorff-Nielsen et al. (2006), Bollerslev et al. (2009a), Huang & Tauchen (2005), Jiang & Oomen (2008), Lee & Mykland (2008), Lee & Ploberger (2009), etc. Numerous papers have used the estimators and tests proposed in this literature, for a multitude of applications, including among others estimating and modeling volatility, detecting jumps in prices, estimating and modeling jump intensities and jump sizes, calibrating option pricing models, and, recently, for computing variance risk premia[1].

In this literature, jumps are understood as discontinuities in the price path, typically assumed to represent market participants' reaction to some important information becoming known. However, there is significant evidence that news are not processed instantaneously, but gradually by what is called 'partial price adjustment':

- 'Clearly, the price does not jump from the old equilibrium level to the new. Instead, trades occur at almost all of the possible nonequilibrium prices along the way.' (Ederington & Lee (1995))
- Adams et al. (2004) also state: 'stocks trade at several interim prices on their way to a new equilibrium price that fully incorporates the news.'

Therefore, the notion of a jump, interpreted from an econometric point of view, does not necessarily coincide with that of a jump, as determined by the data. In this paper, we will therefore distinguish between 'economic' jumps, i.e. movements of prices constituting jumps from an economic perspective, and 'mathematical' jumps. The difference between these may also be stated in terms of reaction speed: a mathematical jump is an inifinitely fast reaction to news, while economic jumps are allowed to take several minutes in order to fully reflect the new information.

There are prominent examples for economic jumps: Figures 1 and 2 display intraday charts for Dow Jones Industrial Average and General Electrics on January 3, 2001. On this day, federal funds rate was cut from 6.5% to 6%,

[1] cf. Carr & Wu (2009), Bollerslev et al. (2009b), Todorov (2010).

and markets reacted to this surprising news by huge upward jumps. However, in compliance with the literature on partial price adjustment, these economic jumps are not mathematical jumps in the data: prices adjusted more or less gradually within several minutes following the news, taking a lot of intermediate values before reaching their new equilibrium. This kind of price movement has been called 'gradual jump' by Barndorff-Nielsen et al. (2009), and it causes severe problems when applying standard estimators and tests. For instance, prevalent estimators of the jumps' contribution to volatility fail to capture the aforementioned economic jumps on January 3, 2001[2].

Until now, literature has neither provided procedures to cope with gradual jumps nor analyzed the reasons for the standard procedures' failure.
In this paper, we will fill both these gaps. The first, quite counter-intuitive lesson to learn here is that increasing frequency does not solve the problem, but worsens it dramatically. Empirical, theoretical as well as Monte Carlo results indicate that, as a rule of thumb, sampling frequency should not be chosen smaller than 5 or 10 minutes, as this is the amount of time markets usually need to fully incorporate surprising new information. The second lesson is that the most widely used approaches to measure and detect jumps are heavily affected by gradual jumps, while some newer test procedures do a better, yet still unsatisfactory job.
We therefore develop new estimators and tests build on intradaily highs and lows of moderate frequency. This type of data is readily available, and the information content embedded in highs and lows is surprisingly useful for disentangling volatility and jumps. Additionally, OHLC data can be visualized by candlestick charts, from which one can draw conclusions about diffusive volatility and jumps: candlesticks with huge body and small wicks indicate jumps, while those with rather small body and large wicks correspond to diffusive price behaviour. Building on this intuition, we develop estimators to separate diffusive volatility from positive and negative jumps. As these can be computed individually for each intradaily subperiod, a bunch of applications becomes available: e.g., locating the timing of jumps within days, studying diurnal patterns in diffusive volatility and jump intensities, and measuring the volatility of volatility.
In line with the literature on partial price adjustment, applying the new procedures to empirical data reveals the importance of economic jumps in empirical data, in particular for stock market indexes, which exhibit much more (economic) jumps than previously reported.

[2]For more details, see the following section.

The rest of the paper is organized as follows: in Section 2, we consider gradual jumps and show empirically how existing estimators react to their presence, in particular when sampling frequency is chosen too high. Section 3 discusses the modeling of intraday prices, with special emphasis on the notions of economic and mathematical jumps as well as volatility. In Section 4, we provide some theory and Monte Carlo results on how different estimators and tests are affected by gradual jumps, underlining the need to sample sparsely. Section 5 demonstrates the merits of intradaily OHLC data of moderate frequency, especially with respect to disentangling volatility and jumps, which in turn is the subject of Section 6, where we construct new estimators for diffusive volatility and the contributions of (upward, downward) jumps to volatility. After presenting new tests to detect economic jumps, we apply the new procedures both to simulated and empirical data in Section 7, while Section 8 concludes.

2 Gradual Jumps: Why it is Necessary to Sample Sparsely

Detecting jumps in asset prices and measuring their contribution to volatility are very important issues in financial econometrics, testified by numerous publications in recent years. Their importance stems from the fact that understanding the dynamics of jump times and jump sizes is indispensable for applications such as derivatives pricing or estimating value at risk. However, modeling these dynamics builds on time series of jump times and jump sizes, which makes it crucial to have procedures at hand to ex post detect and measure jumps. Typically, the contribution of jumps to volatility is measured via an indirect approach: first, an estimate is computed for 'overall volatility'[3], which is comprised of diffusive volatility and the sum of the jumps' squares, and then we subtract from this quantity an estimate of diffusive volatility. For 'overall volatility', the most prominent estimator is the widely used realized variance:

$$\mathrm{RV} := \sum_{i=1}^{N} r_{i,N}^2, \tag{1}$$

[3]'Overall volatility', 'diffusive volatility' and similar notions wille be made more precise in Subsection 3.2.

where $r_{i,N} := p_{\frac{i}{N}} - p_{\frac{i-1}{N}}$ $(i = 1, \ldots, N)$ denotes the returns of some log-price process p_t over the time interval[4] $[\frac{i-1}{N}, \frac{i}{N}]$. To measure diffusive volatility, the dominant approach is to use multipower variation, e.g. bipower or tripower variation, developed by Barndorff-Nielsen & Shephard (2004, 2006):

$$\mathrm{BPV} := \frac{\pi}{2} \frac{N}{N-1} \sum_{i=1}^{N-1} |r_{i,N}|\,|r_{i+1,N}|, \tag{2}$$

$$\mathrm{TPV} := \frac{\pi^{\frac{3}{2}}}{2\,\Gamma(\frac{5}{6})^3} \frac{N}{N-2} \sum_{i=1}^{N-2} |r_{i,N}|^{\frac{2}{3}}\,|r_{i+1,N}|^{\frac{2}{3}}\,|r_{i+2,N}|^{\frac{2}{3}}. \tag{3}$$

Alternatively, one can use the so-called minimum and median realized variance estimators of Andersen et al. (2008):

$$\mathrm{MinRV} := \frac{\pi}{\pi-2} \frac{N}{N-1} \sum_{i=1}^{N-1} \min(|r_{i,N}|, |r_{i+1,N}|)^2, \tag{4}$$

$$\mathrm{MedRV} := \frac{\pi}{\pi - 4\sqrt{3} + 6} \frac{N}{N-2} \sum_{i=1}^{N-2} \mathrm{med}(|r_{i,N}|, |r_{i+1,N}|, |r_{i+2,N}|)^2. \tag{5}$$

On January 3, 2001, federal funds rate was cut to 6%. This very surprising news became known shortly after 1 p.m., and stock markets reacted to it by a huge upward jump, as can be seen from Figures 1 and 2, which show the intraday prices on that day for Dow Jones Industrial Average market index and General Electrics. For DJIA, this more than 3% jump was the largest intraday jump ever in its history. Correspondingly, the sum of squared jumps on that day was about $10\%^2$, while it was even slightly higher for GE.

We computed estimates of the squared jumps' sum for both assets using the aforementioned estimators[5] for frequencies ranging from 1 minute to 30 minutes ($N = 390$ to $N = 13$), including all subsampled versions for frequencies larger than 1 minute, e.g. for 5 minutes frequency, one subsampled version makes use of the intervals 09:30-09:35, 09:35-09:40, etc., the next one uses 09:30-09:31, 09:31-09:36, etc., …, and the last one builds on 09:30-09:34, 09:34-09:39, etc. The first two columns of Figures 3 and 4 show the results (measured in $\%^2$): most of the estimates fall well below the blue line which

[4] For ease of notation, we take the trading interval to be $[0, 1]$ and refer to it as a 'day', although other time spans are perfectly feasible. We also suppress a further index denoting the day under consideration.

[5] We also tried other estimators and different days with changes in federal funds rate, with essentially the same results, which are therefore not presented here.

indicates the squared jump estimated by visual inspection of Figures 1 and 2, and the average of the subsampled estimators, given by the red line, not only fails to reach the blue target line, but even becomes negative for the highest frequencies, although the sum of squared jumps clearly is a non-negative quantity. As a consequence, tests on the presence of jumps, based on these estimators, will therefore typically fail to detect a jump on this particular day.

Naturally, in view of this surprising and very unpleasant result, the following question arises: why do the most prominent estimators fail to capture these jumps and for high frequencies even tend to produce negative values for the sum of squared jumps, a quantity that is known to be nonnegative? The answer becomes evident upon a closer look at Figures 1 and 2: the jumps happen gradually over several minutes, i.e. prices do not adjust to the new value instantaneously, but by a movement that might be seen either as a very steep, more or less straight line, or as a series of smaller jumps.[6] This phenomenon, i.e. the 'gradual adjustment of prices to the new information' (Smidt (1968, p. 252)), has been called 'gradual jump' by Barndorff-Nielsen et al. (2009), and until now, there is no estimator known to be able to accomodate gradual jumps: Barndorff-Nielsen et al. (2009) suggest to simply remove them from the data, which obviously presumes the (non-trivial) ability to detect those days when prices exhibit a gradual jump. In the remainder of this paper, we will fill this gap and develop both methods to detect gradual jumps and to measure them appropriately. However, before pursuing this, we want to direct the reader's attention to another phenomenon identifiable by Figures 3 and 4: for increasing frequencies, the estimates seem to converge to 0. A possible explanation comes from the gradual nature of the jump: it does not consist of one large jump, but rather is composed of several still large, but smaller upward movements. Due to the convexity of the square function, the sum of the squared parts of the gradual jump is much smaller than the squared jump itself, thereby entailing an ever smaller value for the squared jumps' sum.

As a consequence, too high frequencies should be avoided, as choosing too small subintervals leads to the jump being divided into its gradual parts. Instead we should sample with moderate frequencies of say 5 or 10 minutes, because we can safely expect markets to react to news within a period of 10 minutes. At this point, we also want to stress that too high frequencies do not only cause problems regarding the estimation of the jumps' contribution

[6]The nature of gradual jumps will be discussed in Subsection 3.1, while we will investigate the estimators' response to gradual jumps more thoroughly in Subsection 4.1.

to volatility, but also typically produce too small estimates of 'overall volatility', for exactly the same reasons as explained above[7].

A further advantage of sampling sparsely is that a possible bias due to microstructure noise becomes vastly irrelevant, thus relieving us from the burden to investigate estimators with respect to their robustness to the presence of microstructure noise. In the end, for all the reasons given above, we deem it best to sample at moderate frequencies, i.e. frequencies not exceeding 5 minutes.

We will now proceed with some general contemplations concerning the notions of volatility, diffusions, and jumps.

3 On the Modeling of Intraday Asset Prices: Jumps, Drift & Volatility

In high-frequency finance, the prototypical model for intraday prices is given by the 'Brownian semimartingale plus (rare) jumps' model

$$p_t = p_0 + \int_0^t \mu_u du + \int_0^t \sigma_u dW_u + J_t^{(p)}, \tag{6}$$

where

- p_t denotes the logarithmic prices of some asset at time t,
- the drift process μ_t governs the instantaneous mean return at time t,
- the time-varying spot volatility[8] σ_t accounts for the variability of the returns,
- W denotes a standard Brownian motion,
- $J^{(p)}$ is a pure-jump process describing discontinuous changes in log-prices, often assumed to exhibit only finitely many jumps in any time interval[9].

[7] See also Subsection 4.1.

[8] As is common practice in econometrics, we will call σ as well as σ^2 volatility.

[9] The stochastic processes appearing in (6), i.e. μ, σ, and $J^{(p)}$, are subject to some merely technical conditions, for which we refer the reader to the respective literature.

This model, in slightly less general form, dates back at least to Merton (1976): according to Merton, the continuous price components, i.e. the drift part $\int_0^t \mu_u du$ and the diffusive part $\int_0^t \sigma_u dW_u$, model 'normal' vibrations in price, taking place due to new information that causes only marginal changes in the stock's value, while the jump part models 'abnormal' vibrations in price, happening due to the arrival of important new information about the stock that has more than a marginal effect on price.

Although the 'Brownian semimartingale plus (rare) jumps' model is by far the most widely entertained one in high-frequency econometrics, it should be noticed that several authors prefer other models in favor of (6), for reasons such as better description of available data or better explaining of option prices. These models include replacing the rare jumps process and/or the Brownian motion of (6) by more general Lévy processess or fractional Brownian motion, an approach esp. popular in financial mathematics. We will, however, stick to the model (6), with the understanding that (6) describes the unobservable 'true' prices, while empirical data are seen as a noisy version of (6). Within this framework, we will now turn our attention to the notion of jumps.

3.1 Mathematical vs. Economic Jumps

When given a discrete time series, it is typically quite hard to decide whether this time series exhibits a jump. The difficulty does not stem from the fact that we might not have enough data to decide the question at hand, it rather is caused by a more fundamental problem: due to the absence of an exact mathematical definition, we have to resort to more vague concepts like

> 'jumps [...] represent large, infrequent moves over a short time. For example, with a very lengthy historical data set of, say daily prices on stock, you might want to define a jump as any move greater than 10% or, perhaps any move greater than 5 times the historical daily standard deviation.'[10]

With continuous-time models, this ambiguity seems to disappear, as in that case there is an obvious mathematical definition of a jump:

> 'With the underlying theoretical models, it's clear: jumps are discontinuous moves.'[10]

[10]Quoted from www.optioncity.net/faqs.htm

In fact, this mathematical definition is predominant in the literature: a continuous-time log-price process $(p_t)_{t\in[0,1]}$ is said to jump at some time t_J if there is a difference between the price p_{t_J} at time t_J and the price immediately before time t_J, $p_{t_J-} := \lim_{t\uparrow t_J} p_t$.[11] The size of the jump at time t_J, Δp_{t_J}, is then defined as

$$\Delta p_{t_J} := p_{t_J} - p_{T_J-}. \tag{7}$$

However, the existence of a sound mathematical definition of jumps doesn't disburden us from our duties as econometricians, namely putting theoretical (mathematical) concepts into economic contexts.[12] Phrased differently, we have to ask ourselves about what constitutes an 'economic' jump. In this regard, there are at least two important lines of thought:

- Taking the mathematical notion of a jump for granted, an economic jump certainly has to be economically relevant. We do not want to give a precise definition of what should be understood by 'economically relevant', but the idea here is to rule out theoretical (mathematical) jumps that are so tiny, say a hundredth of a tick, that their economic significance might be doubted[13]. We also emphasize that 'economic significance' does not constitute an absolute concept but must be interpreted within a given context: tiny (mathematical) jumps might for some applications be more or less irrelevant, while they may play a crucial role for others, for instance in derivatives pricing.

- Taking a step back, we might also ask whether an economic jump necessarily has to be a mathematical jump in observed prices: a discontinuity in prices is represented by a vertical line in a plot, i.e. by a line of infinite slope. Now, for instance, if prices suddenly start to rise by a very steep line and stop doing so after a few minutes, then we might interpret that movement as a jump, although it is mathematically not necessarily discontinuous[14]. As an example, we recall the up to now largest intraday jump in the history of Dow Jones Industrial Average on January 3, 2001 (see Figure 1). This gradual jump most clearly has to be considered as an economic jump, even if it may not be a

[11]Prices are typically assumed to be right-continuous. This means that for every time t_0, prices p_{t_0} at time t_0 and $p_{t_0+} := \lim_{t\downarrow t_0} p_t$ immediately after time t_0 coincide: $p_{t_0+} = p_{t_0}$.

[12]I'm indebted to Christian Pierdzioch for confronting me with the question 'What is the meaning of a jump in a continuous-time setting?'.

[13]Particularly small jumps are very hard to detect anyway, for details the reader is referred to Aït-Sahalia & Jacod (2009a,b), Lee & Ploberger (2009).

[14]Rasmussen (2009) models jumps by short periods of large drift, i.e. as continuous movements.

mathematical jump in the data, but possibly a more or less continuous movement from the old equilibrium price to the new one.

Building on the ideas used for discrete time series, we might therefore think of an economic jump as an abrupt, rather large price movement, typically taking place during a very short time period after some economic news is released.

Given the above, our understanding of economic jumps can be summarized as follows: economic jumps are jumps in the 'true', but unobservable price process that are relevant to the economic application or caused by economic news, while observed prices do not necessarily have to show a price discontinuity: in the data, the economic jump might appear as comprised of a series of smaller jumps, or as a short period of quite large drift[15].

Although it might also be possible to model gradual jumps by postulating specific properties of the jump process $J^{(p)}$, namely clustering of jumps in time as well as dependence of jump sizes, we will not pursue this approach for the following reasons: first of all, we want to catch the jump in its entirety, not divided into several parts: in the DJIA example, we think there was a more than 3% upward jump, and not several, say 10, upward jumps of 0.3% each, in particular, as it was a reaction to exactly one economic news, the cutting of the interest rate. A second reason not to follow this approach is the fact that we don't even know whether the only discretely observed data are better described by a series of several smaller jumps or by a short period of large drift.

On theoretical grounds, assume that at time t_J an important economic news becomes known which causes log-prices to jump by an amount of Δ. If observed log-prices $\widetilde{p}$ adjust to the new level instantaneously, we have a mathematical jump:

$$\widetilde{p}_{t_J} = \widetilde{p}_{t_J -} + \Delta. \tag{8}$$

If, however, the reaction to the news takes some time Δ_t, then the new log-price $\widetilde{p}_{t_J -} + \Delta$ will result only at time $t_J + \Delta_t$,

$$\widetilde{p}_{t_J + \Delta_t} = \widetilde{p}_{t_J -} + \Delta, \tag{9}$$

while between times t_J and $t_J + \Delta_t$, observed prices somehow adjust to their new level. From this point of view, mathematical jumps are infinitely fast

[15]From a more statistical point of view, it might be interesting to develop methods to be able to decide on how exactly the economic jump comes about in the data. From the econometric point of view taken in this paper, however, this question is of only minor importance.

reactions to news, while economic jumps are reactions to news at finite speed. There are in fact several papers in the finance literature that address the way and speed with which prices adjust to new information: e.g., inter alia, Ederington & Lee (1995), Adams et al. (2004), Chen & Rhee (2010). The common finding of all these papers is what is called 'partial price adjustment': prices do not adjust immediately to new information, but gradually over several minutes, the speed of price adjustment depending on lots of variables like asset class, type of news, market organization, firm size, short sales being allowed or not, and so on.

In particular, we learn from the above that economic jumps might appear in the data disguised as short periods of large drift and small spot volatility, while mathematical jumps could be interpreted as the limit for reaction speed approaching infinity. We will therefore in Section 4 investigate existing estimators and tests with respect to their behaviour to short intraday periods of significant drift and small volatility. In the next subsection, however, we first define more precisely the quantities that we are interested in estimating.

3.2 Volatility: A Single Word to Denote a Plethora of Related, Yet Different Concepts

'Arguably, no concept in financial mathematics is as loosely interpreted and as widely discussed as 'volatility'. [...] 'volatility' has many definitions, and is used to denote various measures of changeability.' [16]

'Volatility is a measure of price variability over some period of time. [...] Volatility can be defined and interpreted in five different ways.' [17]

'Yet, the concept of volatility is somewhat elusive, as many ways exist to measure it and hence to model it.' [18]

Although the above quotations make it clear that 'volatility' is a rather vague concept, volatility estimation has been a very prominent topic in financial econometrics for decades, since many financial applications require knowledge or at least some estimate of price volatility, e.g. asset and derivatives pricing, risk management or portfolio selection. As both measuring and testing for jumps are inextricably linked to volatility measurement, this paper is also going to contribute to the literature concerned with estimating volatility in

[16] Shiryaev (1999, p. 345)
[17] Taylor (2005, p. 189)
[18] Engle & Gallo (2006, p. 4)

a continuous-time setting. Yet even in this setting and letting aside implied volatility, volatility may refer to different concepts:

- we could be interested in what is called both 'notional volatility' (Andersen et al. (2010)) and 'ex-post variation' (Barndorff-Nielsen et al. (2008a)): the quadratic variation of the log-price process (over some day)

$$\mathrm{QV} := \lim_{N\to\infty} \sum_{i=1}^{N} r_{i,N}^2. \tag{10}$$

  Given log-prices on an equidistant grid $t_{i,N} := \frac{i}{N}$, $i = 0, \ldots, N$, the most natural estimator for QV is realized volatility as given by (1). Indeed, there are a lot of papers trying to model and forecast the time series of daily realized volatilities, e.g. inter alia Andersen et al. (2007), Bollerslev et al. (2009a), Corsi et al. (2008).

- overall volatility QV coincides with

$$\mathrm{IV} := \int_0^1 \sigma_u^2 du, \tag{11}$$

  when log-prices follow a Brownian semimartingale, i.e.

$$p_t = p_0 + \int_0^t \mu_u du + \int_0^t \sigma_u dW_u. \tag{12}$$

  IV is called integrated variance or integrated volatility[19]. In the more general case, i.e. in the presence of jumps, QV can be decomposed into IV and SSJ,

$$\mathrm{QV} = \mathrm{IV} + \mathrm{SSJ}, \tag{13}$$

  with the sum of the squared jumps

$$\mathrm{SSJ} := \sum_{t\in[0,1]} (\Delta p_t)^2. \tag{14}$$

  As IV contains information about the contribution of the continuous part of the log-prices to 'volatility', whereas SSJ can tell us something about the variation of returns due to jumps in prices, one often speaks of IV as diffusive volatility or diffusive risk and calls SSJ

[19] As is common practice in econometrics, we will use 'variance' and 'volatility' exchangeably.

jump risk or volatility due to jumps. Starting with the seminal paper Barndorff-Nielsen & Shephard (2004), there are a lot of papers addressing the issue of disentangling QV into IV and SSJ, e.g. inter alia Huang & Tauchen (2005), Barndorff-Nielsen & Shephard (2006), Barndorff-Nielsen et al. (2006), Veraart (2010). In addition, several authors have come up with models for daily IV, often modeling bi- or multivariate series containing at least (transformations of) estimators for (QV, IV), e.g. inter alia Andersen et al. (2007), Bollerslev et al. (2009a), Corsi et al. (2008).

- In many contexts, for instance when considering the value at risk of some financial position, it is not only important to know the risk due to jumps, but to have asymmetric measures of both risks of upward and downward jumps. Theoretically, this amounts to decomposing SSJ into

$$\mathrm{SSJ} = \mathrm{SSpJ} + \mathrm{SSnJ}, \tag{15}$$

with the sum of squared positive jumps

$$\mathrm{SSpJ} := \sum_{t \in [0,1]} 1_{\Delta p_t > 0} \, (\Delta p_t)^2 \tag{16}$$

and the sum of squared negative jumps

$$\mathrm{SSnJ} := \sum_{t \in [0,1]} 1_{\Delta p_t < 0} \, (\Delta p_t)^2, \tag{17}$$

a topic to which the literature has turned only recently, see Barndorff-Nielsen et al. (2008b), Klößner (2008), Patton & Sheppard (2009). In particular, SSpJ and SSnJ are found to have different impacts on future volatility, even for quite long horizons, see Patton & Sheppard (2009).

- Another decomposition of volatility, more specifically quadratic variation QV, is proposed by Barndorff-Nielsen et al. (2008b) and Patton & Sheppard (2009): 'good' volatility emanating from positive returns,

$$\frac{1}{2}\,\mathrm{IV} + \mathrm{SSpJ}\,, \text{ estimated by } \mathrm{RS}^+ := \sum_{i=1}^{N} 1_{r_{i,N} > 0} \, r_{i,N}^2, \tag{18}$$

and 'bad' volatility emanating from negative returns,

$$\frac{1}{2}\,\mathrm{IV} + \mathrm{SSnJ}\,, \text{ estimated by } \mathrm{RS}^- := \sum_{i=1}^{N} 1_{r_{i,N} < 0} \, r_{i,N}^2. \tag{19}$$

Again, good and bad volatility are found to have different impact on future volatility.

As all the mentioned notions of volatility are functions of IV, SSpJ, and SSnJ, the aim of this paper is to provide methods to measure these quantities. However, in view of the above contemplations concerning economic jumps, we will put special emphasis on adequately measuring these quantities when jumps manifest themselves as economic jumps in the data.

4 The Effect of Gradual Jumps on Existing Estimators and Tests

In order to study the effects of economic jumps, we assume that the unobserved true log-prices p exhibit a jump at time t_J, when prices jump by Δ from p_{t_J-} to $p_{t_J-} + \Delta$. Observed prices $\widetilde{p}$, however, are characterized by a finite reaction speed, thus it takes a time span of Δ_t until observed prices have fully adjusted to the new level $p_{t_J-} + \Delta$ at time $t_J + \Delta_t$. Although a more general treatment would be possible, for the sake of convenience we assume the adjustment to be linear[20]:

$$\widetilde{p}_t = p_{t_J-} + \frac{\Delta}{\Delta_t}(t - t_J) \text{ for } t_J \leqslant t \leqslant t_J + \Delta_t. \tag{20}$$

We further assume that we are given a rather high, but fixed frequency such that

$$\frac{1}{N} \ll \Delta_t. \tag{21}$$

In this setting, the jump 'begins' at $t_J \in [\frac{i_1-1}{N}, \frac{i_1}{N}]$, with $i_1 = \lceil t_J N \rceil$, and 'ends' at $t_J + \Delta_t \in [\frac{i_2-1}{N}, \frac{i_2}{N}]$, with $i_2 = \lceil t_J N + \Delta_t N \rceil$. Additionally, we assume that there are no other jumps on that day and no other frictions, entailing $\widetilde{p} = p$ outside of $[t_J, t_J + \Delta_t]$ and $\widetilde{r}_{i,N} := \widetilde{p}_{i,N} - \widetilde{p}_{i-1,N} = r_{i,N}$ for $i < i_1$ and $i > i_2$.

4.1 How Gradual Jumps Affect Existing Estimators

To start with, we have a look at the properties of RV as an estimator for QV. Under (20) and (21), some easy calculations show that for the observed data, realized variance $\widetilde{\mathrm{RV}} := \sum_{i=1}^{N} \widetilde{r}^2_{i,N}$ takes the following form:

$$\widetilde{\mathrm{RV}} = \sum_{i<i_1 \vee i>i_2} r^2_{i,N} + \left((p_{t_J-} - p_{\frac{i_1-1}{N}}) + (\frac{i_1}{N} - t_J)\frac{\Delta}{\Delta_t}) \right)^2 \tag{22}$$

[20]Obviously, $\frac{\Delta}{\Delta_t}$ plays the role of a drift rate during $[t_J, t_J + \Delta_t]$.

$$+\frac{i_2-i_1-1}{\Delta_t^2 N^2}\Delta^2+\left((t_J+\Delta_t-\frac{i_2-1}{N})\frac{\Delta}{\Delta_t}+(p_{\frac{i_2}{N}}-(p_{t_J-}+\Delta))\right)^2.$$

As both $\frac{i_1}{N}-t_J\leqslant\frac{1}{N}$ and $t_J+\Delta_t-\frac{i_2-1}{N}\leqslant\frac{1}{N}$, we can infer from (21) that $(p_{t_J-}-p_{\frac{i_1-1}{N}})+(\frac{i_1}{N}-t_J)\frac{\Delta}{\Delta_t}$ and $(t_J+\Delta_t-\frac{i_2-1}{N})\frac{\Delta}{\Delta_t}+(p_{\frac{i_2}{N}}-(p_{t_J-}+\Delta))$ will be essentially determined by the diffusive parts in $p_{t_J-}-p_{\frac{i_1-1}{N}}$ and $p_{\frac{i_2}{N}}-(p_{t_J-}+\Delta)$. Comparing unobservable 'true' realized volatility $\mathrm{RV}=\sum_{i=1}^{N}r_{i,N}^2$ to empirical realized volatility $\widetilde{\mathrm{RV}}$, using $i_2-i_1-1\approx\Delta_t N$, then shows that $\mathrm{RV}-\widetilde{\mathrm{RV}}$ is essentially given by

$$\widetilde{\mathrm{RV}}\approx-\mathrm{RV}-\sum_{i_1<i<i_2}r_{i,N}^2-\Delta^2\,(1-\frac{1}{\Delta_t N}). \tag{23}$$

From (23), we can learn two things: first, the economic jump leads to integrated volatility being slightly underestimated, as the equivalent of diffusive volatility over $[\frac{i_1}{N},\frac{i_2-1}{N}]$ is missing in $\widetilde{\mathrm{RV}}$. The effect is however more or less negligible since the interval's length is smaller than Δ_t, which comprises only a few minutes.
The second consequence of (23) is much more important: due to (21), $\frac{1}{\Delta_t N}$ is quite small so that $\widetilde{\mathrm{RV}}$ doesn't capture the jump adequately, the problem getting even worse with increasing frequency. Put differently, for increasing frequency, $\widetilde{\mathrm{RV}}$ will more and more fail to grasp the gradual jump's contribution to overall volatility, eventually completely ignoring the gradual jump.

We will now turn our attention to bipower variation BPV as an estimator of diffusive volatility IV. Again, by some easy calculations and similar approximations as above, we find for $\widetilde{\mathrm{BPV}}:=\frac{\pi}{2}\sum_{i=2}^{N}|\widetilde{r}_{i_N}||\widetilde{r}_{i-1,N}|$:

$$\widetilde{\mathrm{BPV}}\approx\frac{\pi}{2}\sum_{i<i_1\,\vee\, i>i_2+1}|r_{i_N}||r_{i-1,N}|+\frac{\pi}{2}\Delta^2\frac{1}{\Delta_t N}, \tag{24}$$

from which we see that $\widetilde{\mathrm{BPV}}$ is distorted as an estimator for IV, the strength of the distortion depending on squared jump size Δ^2 and becoming negligible when frequency is high enough.

Combining (22), (24), we find for $\widetilde{\mathrm{RV}}-\widetilde{\mathrm{BPV}}$ as an estimator of SSJ:

$$\widetilde{\mathrm{RV}}-\widetilde{\mathrm{BPV}}\approx\sum_{i<i_1\,\vee\, i>i_2}r_{i,N}^2-\frac{\pi}{2}\sum_{i<i_1\,\vee\, i>i_2+1}|r_{i_N}||r_{i-1,N}|+(1-\frac{\pi}{2})\frac{\Delta^2}{\Delta_t N}. \tag{25}$$

In (25), the difference of the sums is an estimator for the squared jumps outside of $[\frac{i_1-1}{N}, \frac{i_2}{N}]$, which equals 0, as we have assumed the absence of further jumps. Therefore, at least for large economic jumps Δ, $\widetilde{\mathrm{RV}} - \widetilde{\mathrm{BPV}}$ is dominated by $(1-\frac{\pi}{2})\frac{\Delta^2}{\Delta_t N} < 0$, which explains both the negative estimates for SSJ in Figures 3 and 4 as well as their converging to 0 for increasing frequencies.

With regard to tripower variation TPV, minimum realized variance MinRV, and median realized variance MedRV, similar approximations show that these estimators have the same properties as BPV: integrated volatility IV will be overestimated in the presence of economic jumps, while the economic jumps themselves aren't captured properly, esp. for high frequencies, when there is a tendency to produce negative estimates for the squared jumps' sum SSJ, resulting from the constants $\frac{\pi^{3/2}}{2\Gamma(5/6)^3}$, $\frac{\pi}{\pi-2}$, $\frac{\pi}{\pi-4\sqrt{3}+6}$ in (3), (4), (5) exceeding 1.

Summarizing, we can state that existing estimators of diffusive volatility IV are upward biased, while both overall volatility QV and jump volatility SSJ are severely underestimated in the presence of economic jumps, the underestimation growing stronger with increasing frequency, routinely producing negative estimates of SSJ for very high frequencies.

We will now investigate several tests for detecting jumps with respect to their reaction to gradual jumps.

4.2 How Gradual Jumps Affect Existing Jump Tests

The first procedures to test for jumps have been developed by Barndorff-Nielsen & Shephard (2004, 2006), while Huang & Tauchen (2005) have singled out one specific test of that family, which they found to have the best properties with respect to size. In the sequel, literature has seen several other proposals, including[21] those by Jiang & Oomen (2008), Aït-Sahalia & Jacod (2009b), and Lee & Ploberger (2009). We will now have a look at how these tests respond to the presence of gradual jumps.

The well-known jump test by Barndorff-Nielsen & Shephard (2006) in its

[21]In our comparison, we don't include the tests by Lee & Mykland (2008) and Rasmussen (2009), as these are similar to the Lee/Ploberger test, but much worse in terms of holding their level for moderate frequencies or heavily fluctuating volatility.

version by Huang & Tauchen (2005) looks as follows:

$$z_{TP,rm} := \frac{\frac{\mathrm{RV}-\mathrm{BPV}}{\mathrm{RV}}}{\sqrt{(\frac{\pi^2}{4}+\pi-5)\frac{1}{N}\max(1,\frac{\mathrm{TPQ}}{\mathrm{BPV}^2})}} \dot{\sim} N(0,1), \tag{26}$$

where TPQ denotes an estimator of integrated quarticity $\mathrm{IQ} := \int_0^1 \sigma_t^4 dt$, and the presence of mathematical jumps shifts the test statistic to the right. In view of the results of the previous subsection concerning the properties of $\mathrm{RV}-\mathrm{BPV}$ as an estimator of SSJ, we can infer that for high frequencies, gradual jumps will tend to produce small, albeit negative values for the numerator[22] of (26). As the test is right-sided, we expect it to have only poor power against this alternative, esp. for high frequencies[23]. As there is a tendency for negative values of the test statistic, we even expect the test's power eventually to fall below its nominal size.
A particularly nasty consequence of the above finding is that gradual jumps and mathematical jumps tend to cancel out each other: while the test statistic is diminished by gradual jumps, it increases due to mathematical jumps, which results in gradual jumps reducing the test's ability to detect mathematical jumps, if these occur on the same day.

The jump test of Jiang & Oomen (2008) in its 'ratio'-version[24] works as follows:

$$\mathrm{SwV}_r := \frac{N\,\widehat{\mathrm{IV}}}{\sqrt{\frac{15}{9}\widehat{\mathrm{IS}}}} \frac{\mathrm{SwV}-\mathrm{RV}}{\mathrm{RV}} \dot{\sim} N(0,1), \tag{27}$$

where SwV (called 'swap-variance') is defined as

$$\mathrm{SwV} := 2\sum_{i=1}^{N}(e^{r_{i,N}} - 1 - r_{i,N}), \tag{28}$$

and $\widehat{\mathrm{IV}}$ and $\widehat{\mathrm{IS}}$ are estimators of integrated volatility IV and integrated sixticity $\mathrm{IS} := \int_0^1 \sigma_t^6 dt$. Upward jumps shift SwV_r to the right, while downward

[22]For the test at hand, as well as for some of the following ones, the effect of gradual jumps on the denominator can be more or less safely ignored.

[23]For all tests discussed in the text, we expect significant power against a gradual jump as soon as frequency is chosen low enough, i.e. such that the the jump is contained in only one subperiod $[\frac{i-1}{N}, \frac{i}{N}]$.

[24]There are also versions called 'diff' and 'log', cf. Jiang & Oomen (2008).

jumps shift the test statistic to the left[25]. The null hypothesis of no jumps thus is rejected when SwV_r takes a too large absolute value.

Using the notations and techniques of the previous subsection, we easily find for $\widetilde{\mathrm{SwV}} := 2\sum_{i=1}^{N}(e^{\widetilde{r}_{i,N}} - 1 - \widetilde{r}_{i,N})$ the following approximation for $\widetilde{\mathrm{SwV}} - \widetilde{\mathrm{RV}}$:

$$\widetilde{\mathrm{SwV}} - \widetilde{\mathrm{RV}} \approx \sum_{i<i_1 \vee i>i_2} \left(2(e^{r_{i,N}} - 1 - r_{i,N}) - r_{i,N}^2\right) \tag{29}$$

$$+\frac{1}{\Delta_t N}\left(2(e^{\frac{\Delta}{\Delta_t N}} - 1 - \frac{\Delta}{\Delta_t N}) - (\frac{\Delta}{\Delta_t N})^2\right)$$

In the absence of further jumps, the first sum of (29) will be of order $O(\frac{1}{N^{3/2}})$, while the second term, which is of the same sign as Δ, might be substantial, provided that $\frac{\Delta}{\Delta_t N}$ is not too small. We therefore expect the swap variance test to have non-trivial power against gradual jumps, which decreases with increasing frequency.

The jump test of Aït-Sahalia & Jacod (2009b) builds on power variations of different frequencies. Again, we select a specific test[26]: the in the absence of jumps asymptotically standard normally distributed test statistic

$$\frac{\sqrt{N}(\widehat{S}(4,2) - 2)}{\sqrt{\frac{160}{3}\frac{\widehat{\mathrm{IO}}}{\widehat{\mathrm{IQ}}^2}}} \dot{\sim} N(0,1), \tag{30}$$

with $\widehat{\mathrm{IO}}$ and $\widehat{\mathrm{IQ}}$ estimators for integrated octicity $\mathrm{IO} := \int_0^1 \sigma_t^8 dt$ and integrated quarticity IQ. The ratio

$$\widehat{S}(4,2) := \frac{\sum_{i=1}^{N/2}(p_{\frac{2i}{N}} - p_{\frac{2(i-1)}{N}})^4}{\sum_{i=1}^{N}(p_{\frac{i}{N}} - p_{\frac{i-1}{N}})^4} \tag{31}$$

is shifted towards 1 in the presence of mathematical jumps, so that we reject the null hypothesis of no jumps for strongly negative values of the test statistic. Now, what happens in the presence of a gradual jump? In the

[25]In principle, positive and negative jumps on the same day could therefore cancel out each other.

[26]Results for other members of that test family are similar.

numerator of $\widehat{S}(4,2)$, the term $(2\frac{\Delta}{\Delta_t N})^4$ will appear essentially $\frac{i_2-i_1}{2} \approx \frac{1}{2\Delta_t N}$ times, while in the denominator, essentially $i_2 - i_1 \approx \frac{1}{\Delta_t N}$ instances of $(\frac{\Delta}{\Delta_t N})^4$ will show up. Thus $\widehat{S}(4,2)$ will be shifted towards 8, resulting in largely positive values of the test statistic. For this reason, we expect this test to have particularly low power against gradual jumps, which will be vanishing for high frequencies.
For the same reasons as explained above for the Barndorff-Nielsen/Shephard test, gradual jumps will hinder the test detecting mathematical jumps happening on the same day, as these two types of jumps shift the test statistic to different directions.

Finally, we want to discuss the test of Lee & Ploberger (2009), which attains the optimal detection rate for testing against mathematical jumps. It is based on the test statistic

$$\tau := \max_{i=l+1,\dots,N} \frac{r_{i,N}^2}{\frac{1}{l}\sum_{k=1}^{l} r_{i-k,N}^2}, \tag{32}$$

where l is chosen either as $l = 4\log N$ or $l = 2\log N$, and a certain positive transformation of τ is exponentially distributed under the null hypothesis of no jumps[27]. The effect of gradual jumps on τ are not that clear: on one hand, the gradual jump will enlarge $r_{i,N}^2$ in the numerator approximately by $\frac{\Delta^2}{\Delta_t^2 N^2}$ for $i_1 < i < i_2$, on the other hand, the denominator will also increase for certain values of i (in relation to l, i_1, and i_2). In particular, we expect higher ratios for i in the vicinity of i_1, as the numerator will increase by a larger amount than the denominator, while for i slightly larger than i_2, the ratios will be reduced, as only the denominator becomes affected by the gradual jump. All in all, however, due to the test statistic τ being the maximum over all ratios, we expect the test to have significant power against the alternative of gradual jumps.

4.3 Monte Carlo Study

In order to confront the previous subsections' theoretical considerations with data, we have simulated 30,000 paths of Brownian motion subjected to at least one gradual jump per day: more precisely, we calibrated a high and a low volatility setting, by taking $\sigma^2 = 10^{-4}$ ($\sigma^2 = 9 \cdot 10^{-6}$), matching typical empirical estimates for daily volatility of stocks (indexes). The gradual

[27] cf. Lee & Ploberger (2009)

jump's starting time t_J was drawn uniformly on $[0, 1]$, the jump lengths Δ_t according to a shifted binomial distribution taking values from 1 minute to 13 minutes with a mean jump length of 7 minutes, and jump sizes of either sign were drawn with absolute values from a uniform distribution on[28] $[0.3\%, 0.9\%]$.

Table 1 and the first four rows of Tables 2 and 3 (p. 41, 42) show bias and RMSE for the estimators considered in this section, for frequencies ranging from sampling every minute to the very sparse half-hourly sampling, with bold printing indicating optimal frequency. In general, all estimators' variances decrease with frequency, so that the RMSE-minimizing frequency is always less or equal to the bias-minimizing frequency. The reader should keep in mind that diffusive volatility IV equals $\text{IV} = \sigma^2 = 1\%^2$ ($\text{IV} = \sigma^2 = 0.09\%^2$) in the high (low) voaltility setting, while the uniformly on $[0.3\%, 0.9\%]$ distributed absolute values of the jumps are on average responsible for jump volatility of $\text{SSJ} = 0.39\%^2$. Quadratic variation QV therefore varies between $1.09\%^2$ and $1.81\%^2$ ($0.18\%^2$ and $0.9\%^2$), with an expected value of $1.39\%^2$ ($0.48\%^2$).

	1 min	3 min	5 min	10 min	15 min	30 min
$\sigma^2 = 10^{-4}$	-0.344	-0.245	-0.175	-0.090	-0.057	**-0.026**
	(0.404)	(0.321)	**(0.287)**	(0.319)	(0.378)	(0.526)
$\sigma^2 = 9 \cdot 10^{-6}$	-0.346	-0.247	-0.177	-0.091	-0.060	**-0.028**
	(0.399)	(0.290)	(0.217)	(0.153)	**(0.139)**	(0.141)

Table 1: Bias (and RMSE) of RV (in $\%^2$)

Table 1 shows simulation bias and RMSE of RV, when estimating $\text{QV} = \text{IV} + \text{SSJ}$: as predicted by the theoretical results in Subsection 4.1, the gradual jumps are not captured well by RV, resulting in a negative bias of RV. It is also evident from Table 1 that this effect becomes more and more prominent when frequency increases. Both bias and RMSE are substantial, esp. for high frequencies and the low volatility setting, but also when choosing optimal frequencies.

We will now turn our attention to Table 2 (p. 41), the first four rows of which display the performance of bi- and tripower variation as well as minimum and

[28] Fair (2002) defines 'large price changes' or 'events' by returns of a maximal length of 5 minutes exceeding 0.75% in absolute value.

median realized variance when estimating diffusive volatility IV.
From theory, we know that the bias will disappear when frequency approaches infinity. However, let us explain why both mean and RMSE do not depend monotonuously upon frequency: on one hand, increasing sampling frequency helps reducing the impact of a jump, as all estimators combine in some way or another the jump return with returns of adjacent intervals, which will become smaller with increasing frequency, provided that these intervals do not contain parts of the jump. On the other hand, for high frequencies, the jump becomes more and more 'gradual', i.e. the jump is divided into more and more parts, resulting in an ever higher number of adjacent intervals containing parts of the jump, which will therefore not be damped at all.
In the high volatility setting, sampling at the 1 minute frequency turns out to be optimal among the frequencies considered, while in the low volatility setting, results are varying very much across estimators. In view of the quantity to be estimated, $\text{IV} = \sigma^2 = 1\%^2$ ($\text{IV} = \sigma^2 = 0.09\%^2$), bias and RMSE are substantial, esp. in the low volatility setting, when they are almost always larger than IV, the quantity to be estimated.
All in all, when estimating diffusive volatility IV in the presence of gradual jumps, sampling frequency should be chosen as high as possible, in particular when jump sizes are large in comparison to diffusive volatility, as in the low volatility setting, for which frequencies beyond 1 minute are needed for estimating diffusive volatility IV adequately.

With respect to the estimation of SSJ, Table 3 (p. 42) tells a completely different story: in line with theory, $\text{RV} - \text{BPV}$ etc. are not capable of capturing the gradual jumps, their bias increasing with frequency even beyond the average SSJ of $0.39\%^2$, indicating that at the 1 minute frequency, these estimators perform worse than the very bad estimator $\widehat{\text{SSJ}} \equiv 0$, which would be less biased[29].

Summarizing the effect of gradual gradual jumps on the estimation of the different components of volatility, we find that the presence of gradual jumps causes prevalent estimators to be severely biased. While the bias disappears asymptotically when estimating diffusive volatility IV, increasing the sampling frequency makes things even worse when estimating the jumps' contribution to volatility, SSJ, or, consequently, overall volatility $\text{QV} = \text{IV} + \text{SSJ}$. We therefore need new estimators for these quantities based on moderate sampling frequencies, which we will develop in the next sections.

[29]For all estimators of SSJ, both bias and RMSE could be slightly improved by truncation at 0.

In Figure 5, we report the power of the jump tests discussed in the previous subsection: the graphs show power after size-adjusting with quantiles obtained from simulating Brownian motion (without jumps), as some of the tests exhibit severe size distortions for moderate frequencies[30]. The theoretical results are confirmed by Figure 5: while the tests by Barndorff-Nielsen/Shephard and Aït-Sahalia/Jacod have generally only low power, falling below the non-trivial line for the frequency of 1 minute, the tests by Jiang/Oomen and Lee/Ploberger perfom much better, esp. still showing non-trivial power when using 1 minute data.

All in all, we emphasize the need for sparse sampling: all the tests' power is very poor for frequencies beyond one minute, so that in order to be able to detect gradual jumps, sparse sampling is a must. However, sparse sampling brings about a loss in efficiency, may violate the statistician's principle of using all available data, and comes at the cost of not being able to detect very small jumps, as this would necessitate the use of high frequencies. In the next sections, we will discuss a solution to this dilemma, the use of moderately sampled intradaily highs and lows.

5 A Non-Chartist View on the Informational Content of Candlestick Charts

Intradaily data of moderate frequency, say 5 or 10 minutes, are often available at quite reasonable prices. In many cases, it is possible to acquire intradaily OHLC data, i.e. in addition to log-prices $p_{\frac{i}{N}}$ on some equidistant grid $0, \frac{1}{N}, \ldots, 1$ ($[0,1]$ denoting the daily trading interval), we are also given intradaily highs $(p^*)_{i,N}$ and lows $(p_*)_{i,N}$:

$$(p^*)_{i,N} := \sup_{\frac{i-1}{N} \leqslant t \leqslant \frac{i}{N}} p_t,\ (p_*)_{i,N} := \inf_{\frac{i-1}{N} \leqslant t \leqslant \frac{i}{N}} p_t\,. \tag{33}$$

Then we have four log-prices in every subinterval $[\frac{i-1}{N}, \frac{i}{N}]$:

- opening price $p_{\frac{i-1}{N}}$,
- closing price $p_{\frac{i}{N}}$,
- highest price $(p^*)_{i,N}$,

[30]For details, see Balter & Klößner (2010).

- and lowest price $(p_*)_{i,N}$.

These can be visualized by a method quite popular among chartists, the so-called candlestick charts (see Figures 7 and 8). They are constructed by applying the following procedure to every subinterval:

- plot the candlestick's body, i.e. a rectangular box whose lower (upper) coordinate is given by the minimum (maximum) of the subinterval's opening and closing price,
- colour the rectangle in green (or blue) for upward movements, i.e. when closing price exceeds opening price, and in red for downward movements, i.e. when closing price is smaller than opening price,
- draw the upper wick, i.e. atop of the rectangle, draw a line whose upper coordinate is given by the highest price,
- draw the lower wick, i.e. beneath the rectangle, draw a line whose lower coordinate is given by the lowest price.

We will refrain from discussing the merits ascribed to candlesticks by technical analysts, but instead concentrate on the informational content of candlesticks with respect to diffusive price behaviour and jumps. Figure 7 shows simulated candlesticks for the most simple diffusion process, namely a Brownian motion. It reveals that for a diffusive process the candlesticks can take a lot of different forms: a large body with quite small wicks like the one at 09:55 a.m., a very small body with large wicks like the one at 10:05 a.m., candlesticks with a small and a long wick like the ones at 10:00 a.m. and 10:10 a.m., etc. However, for a continuous price process (with only moderate drift), it is very unlikely to see a candlestick with a huge body and almost no wicks, the form that we expect to observe in case of a (large) jump (cf. the candlestick at 1:15 p.m. in Figure 8). In order to discuss the reason underlying this observation, we will now start to analyze diffusion processes more thoroughly: to that end, assume that log-prices p exhibit no jumps, i.e.

$$p_t = p_0 + \int_0^t \mu_u du + \int_0^t \sigma_u dW_u. \tag{34}$$

For $t \in [\frac{i-1}{N}, \frac{i}{N}]$, we will then have

$$p_t = p_{\frac{i-1}{N}} + \int_{\frac{i-1}{N}}^t \mu_u du + \int_{\frac{i-1}{N}}^t \sigma_u dW_u. \tag{35}$$

Rescaling time as $\tau := \frac{t-\frac{i-1}{N}}{\frac{1}{N}} \in [0,1]$, we get using

$$\widetilde{\mu}_\tau := \mu_{\frac{i-1}{N}+\frac{\tau}{N}},\ \widetilde{\sigma}_\tau := \sigma_{\frac{i-1}{N}+\frac{\tau}{N}} \tag{36}$$

and the Brownian motion

$$\widetilde{W}_\tau := \sqrt{N}\left(W_{\frac{i-1}{N}+\frac{\tau}{N}} - W_{\frac{i-1}{N}}\right)\ (\tau \in [0,1]) \tag{37}$$

the following formula:

$$p_t - p_{\frac{i-1}{N}} = \frac{1}{\sqrt{N}}\left(\frac{1}{\sqrt{N}}\int_0^\tau \widetilde{\mu}_u du + \int_0^\tau \widetilde{\sigma}_u d\widetilde{W}_u.\right) \tag{38}$$

In fact, (38) is at the heart of high-frequency volatility measurement: it shows that asymptotically for ever higher frequencies, i.e. for N approaching infinity, returns are of order $\frac{1}{\sqrt{N}}$ and essentially determined by prices' diffusive part, while the drift's influence vanishes due to the additional factor $\frac{1}{\sqrt{N}}$ in front of the first integral in (38). As a consequence of (38), for most applications one can proceed assuming locally constant volatility as well as absent drift[31]. We will follow the same strategy here, setting $\mu \equiv 0$ and, within subintervals, $\sigma_t \equiv \sigma$, for derivations and motivations. We therefore have the following simplified version of (38):

$$p_t - p_{\frac{i-1}{N}} \approx \frac{\sigma_{\frac{i-1}{N}}}{\sqrt{N}}\widetilde{W}_\tau. \tag{39}$$

Approximately, this gives

$$p_{\frac{i}{N}} \approx p_{\frac{i-1}{N}} + \frac{\sigma_{\frac{i-1}{N}}}{\sqrt{N}}\widetilde{W}_1, \tag{40}$$

$$(p^*)_{i,N} \approx p_{\frac{i-1}{N}} + \frac{\sigma_{\frac{i-1}{N}}}{\sqrt{N}}\widetilde{W}^*, \tag{41}$$

$$(p_*)_{i,N} \approx p_{\frac{i-1}{N}} + \frac{\sigma_{\frac{i-1}{N}}}{\sqrt{N}}\widetilde{W}_*, \tag{42}$$

with

$$\widetilde{W}^* := \sup_{0\leqslant t\leqslant 1} \widetilde{W}_t,\ \widetilde{W}_* := \inf_{0\leqslant t\leqslant 1} \widetilde{W}_t. \tag{43}$$

[31] Of course, this approach requires some mathematics, tackled brilliantly by Mykland & Zhang (2009).

For the candlestick's body's length $b_{i,N}$, the upper and lower wicks' lengths $\mathrm{uw}_{i,N}$ and $\mathrm{lw}_{i,N}$, we get:

$$b_{i,N} = |p_{\frac{i}{N}} - p_{\frac{i-1}{N}}| \approx \frac{\sigma_{\frac{i-1}{N}}}{\sqrt{N}} |\widetilde{W}_1|, \tag{44}$$

$$\mathrm{uw}_{i,N} = (p^*)_{i,N} - \max(p_{\frac{i-1}{N}}, p_{\frac{i}{N}}) \approx \frac{\sigma_{\frac{i-1}{N}}}{\sqrt{N}} \left(\widetilde{W}^* - \max(\widetilde{W}_1, 0)\right), \tag{45}$$

$$\mathrm{lw}_{i,N} = \min(p_{\frac{i-1}{N}}, p_{\frac{i}{N}}) - (p_*)_{i,N} \approx \frac{\sigma_{\frac{i-1}{N}}}{\sqrt{N}} \left(\min(\widetilde{W}_1, 0) - \widetilde{W}_*\right) \tag{46}$$

(44)-(46) in particular show that for not too small frequencies, the sizes of the candlestick's body and wicks are up to a scaling factor determined by certain functions of the triple (W_1, W^*, W_*) of terminal, maximal, and miminal value of a Brownian motion, the distribution of which is given in Borodin & Salminen (2002). It is precisely this distribution which makes occurences of candlesticks with huge body but tiny wicks quite improbable in the absence of jumps.

We will now study how (44)-(46) change when there is an economic jump within $]\frac{i-1}{N}, \frac{i}{N}[$: we assume that between times $t_J > \frac{i-1}{N}$ and $t_J + \Delta_t < \frac{i}{N}$ log-prices 'jump' from p_{t_J-} by Δ to $p_{t_J+\Delta_t} = p_{t_J-} + \Delta$.[32] Introducing the notations

$$T_1 := \frac{t_J - \frac{i-1}{N}}{\frac{1}{N}}, T_2 := \frac{t_J + \Delta_t - \frac{i-1}{N}}{\frac{1}{N}}, \; (0 < T_1 \leqslant T_2 < 1) \tag{47}$$

$$\sigma_l := \sigma_{t_J-}, \sigma_r := \sigma_{t_J+\Delta_t}, \tag{48}$$

and the independent Brownian motions

$$W^l_\tau := \sqrt{\frac{N}{T_1}} \left(W_{\frac{i-1}{N}+\frac{T_1}{N}\tau} - W_{\frac{i-1}{N}}\right) \; (\tau \in [0,1]), \tag{49}$$

$$W^r_\tau := \sqrt{\frac{N}{1-T_2}} \left(W_{t_J+\Delta_t+\frac{1-T_2}{N}\tau} - W_{t_J+\Delta_t}\right) \; (\tau \in [0,1]), \tag{50}$$

we then have

$$p_t \approx \begin{cases} p_{\frac{i-1}{N}} + \sigma_l \sqrt{\frac{T_1}{N}} W^l_{\frac{N}{T_1}(t-\frac{i-1}{N})} & , \quad \frac{i-1}{N} \leqslant t < t_J \\ p_{\frac{i-1}{N}} + \sigma_l \sqrt{\frac{T_1}{N}} W^l_1 + \Delta + \sigma_r \sqrt{\frac{1-T_2}{N}} W^r_{\frac{N}{1-T_2}(t-(t_J+\Delta_t))} & , \; t_J + \Delta_t \leqslant t \leqslant \frac{i}{N} \end{cases} . \tag{51}$$

[32] $\Delta_t = 0$ delivers the case of a mathematical jump, while $\Delta_t > 0$ for a gradual jump.

As long as we do not specify the exact mathematical nature of the economic jump, we do not have a formula for p_t 'during' the jump, i.e. for $t_J \leqslant t < t_J + \Delta_t$. However, at least approximately, we may assume that p_t is between p_{t_J-} and $p_{t_J+\Delta_t}$. If the jump size Δ is not too small in comparison to $\frac{\sigma_l}{\sqrt{N}}, \frac{\sigma_r}{\sqrt{N}}$[33], a positive jump $\Delta > 0$ will induce

$$p_{\frac{i}{N}} \approx p_{\frac{i-1}{N}} + \sigma_l \sqrt{\frac{T_1}{N}} W_1^l + \Delta + \sigma_r \sqrt{\frac{1-T_2}{N}} W_1^r, \tag{52}$$

$$(p^*)_{i,N} \approx p_{\frac{i-1}{N}} + \sigma_l \sqrt{\frac{T_1}{N}} W_1^l + \Delta + \sigma_r \sqrt{\frac{1-T_2}{N}} (W^r)^*, \tag{53}$$

$$(p_*)_{i,N} \approx p_{\frac{i-1}{N}} + \sigma_l \sqrt{\frac{T_1}{N}} (W^l)_*, \tag{54}$$

while for a negative jump $\Delta < 0$:

$$p_{\frac{i}{N}} \approx p_{\frac{i-1}{N}} + \sigma_l \sqrt{\frac{T_1}{N}} W_1^l + \Delta + \sigma_r \sqrt{\frac{1-T_2}{N}} W_1^r, \tag{55}$$

$$(p^*)_{i,N} \approx p_{\frac{i-1}{N}} + \sigma_l \sqrt{\frac{T_1}{N}} (W^l)^*, \tag{56}$$

$$(p_*)_{i,N} \approx p_{\frac{i-1}{N}} + \sigma_l \sqrt{\frac{T_1}{N}} W_1^l + \Delta + \sigma_r \sqrt{\frac{1-T_2}{N}} (W^r)_*. \tag{57}$$

This implies for the lengths of the candlestick's body and wicks:

$$b_{i,N} \approx \left| \sigma_l \sqrt{\frac{T_1}{N}} W_1^l + \Delta + \sigma_r \sqrt{\frac{1-T_2}{N}} W_1^r \right| \approx |\Delta|, \tag{58}$$

$$\mathrm{uw}_{i,N} \approx \begin{cases} \sigma_r \sqrt{\frac{1-T_2}{N}} \left((W^r)^* - W_1^r\right) & , \Delta > 0 \\ \sigma_l \sqrt{\frac{T_1}{N}} (W^l)^* & , \Delta < 0 \end{cases}, \tag{59}$$

$$\mathrm{lw}_{i,N} \approx \begin{cases} \sigma_l \sqrt{\frac{T_1}{N}} \left(-(W^l)_*\right) & , \Delta > 0 \\ \sigma_r \sqrt{\frac{1-T_2}{N}} \left(W_1^r - (W^r)_*\right) & , \Delta < 0 \end{cases}. \tag{60}$$

From (58)-(60), we find that a jump causes the candlestick to have a rather large body of length $|\Delta|$, while both wicks are essentially unaffected by jumps, as a comparison of (45), (46) with (59), (60) shows.

[33]For a more thorough discussion concerning the jump size necessary for being detected, see Lee & Ploberger (2009).

Both for the theoretical reasons and the empirical facts given above it is therefore possible to use intradaily OHLC data and candlestick charts for inference on the presence of economic jumps, because

- large wicks, as compared to the candlestick's body, indicate diffusive behaviour in the absence of jumps,
- a large body, in combination with tiny wicks, indicates a jump.

In the following, we will use these facts to construct estimators that can disentangle variation into diffusive volatility and volatility due to jumps.

6 Disentangling Diffusive Volatility and Economic Jumps

6.1 Estimation in the Presence of Large Jumps

With the insights of the previous section, it is obvious that we should exploit the different behaviour of the candlestick's body and wicks in order to disentangle diffusive movements from jumps. Before developing estimators for SSJ, SSpJ, and SSnJ, we first construct a very jump-robust estimator of diffusive volatility

$$\mathrm{IV}_{i,N} := \int_{\frac{i-1}{N}}^{\frac{i}{N}} \sigma_t^2 dt \tag{61}$$

within subperiod $[\frac{i-1}{N}, \frac{i}{N}]$. To this end, we make use of upper and lower wick lengths $\mathrm{uw}_{i,N}$, $\mathrm{lw}_{i,N}$, while disregarding the body length $b_{i,N}$, as only the wicks' lengths are robust to the presence of jumps. To estimate $\mathrm{IV}_{i,N}$, we consider quadratic functions of the upper and lower wicks' lengths $\mathrm{uw}_{i,N}$ and $\mathrm{lw}_{i,N}$, i.e. $\mathrm{IV}_{i,N}$ will be estimated by a linear combination of $\mathrm{uw}_{i,N}^2$, $\mathrm{lw}_{i,N}^2$, and $\mathrm{uw}_{i,N}\,\mathrm{lw}_{i,N}$.

From Klößner (2009) and (45), (46), we can infer that the estimators[34]

$$4\,\mathrm{uw}_i^2,\ 4\,\mathrm{lw}_i^2,\ \frac{4}{8\log(2)-5}\,\mathrm{lw}_i\,\mathrm{uw}_i \tag{62}$$

are unbiased for IV_i whenever there are no jumps and the approximation in (45), (46) holds exactly[35]. In this case, it can also be shown that the optimal

[34] For notational simplicity, we from now on suppress the index N.

[35] This will especially be true when p is Brownian motion.

linear combination of the aforementioned estimators is given by

$$\widehat{\mathrm{IV}}_i^{(l)} := 1.3227\,\mathrm{uw}_i^2 + 1.3227\,\mathrm{lw}_i^2 + 2.4847\,\mathrm{uw}_i\,\mathrm{lw}_i, \tag{63}$$

where 'optimal' means that under all linear combinations of the estimators (62), consistent for IV_i, $\widehat{\mathrm{IV}}_i^{(l)}$ has smallest variance, namely $0.7244\frac{\sigma^4_{\frac{i-1}{N}}}{N^2}$. Notice that $\widehat{\mathrm{IV}}_i^{(l)}$ is by construction an estimator for $\mathrm{IV}_i = \int_{\frac{i-1}{N}}^{\frac{i-1}{N}} \sigma_t^2 dt$, implying that $N\,\widehat{\mathrm{IV}}_i^{(l)}$ might be used to estimate spot volatility σ^2 in $[\frac{i-1}{N}, \frac{i}{N}]$.

We now turn to the estimation of

$$\mathrm{SSJ}_i := \sum_{t \in [\frac{i-1}{N}, \frac{i}{N}[} \Delta_t^2, \tag{64}$$

the sum of squared jumps during $[\frac{i-1}{N}, \frac{i}{N}[$. From (58), we find that SSJ_i should be estimated by the squared body length b_i^2, if $[\frac{i-1}{N}, \frac{i}{N}[$ contains a large economic jump. However, when looking for an estimator for SSJ_i, we must also make sure that the estimator behaves appropriately in the absence of jumps, when body length b_i is essentially given by (44). Unfortunately, (44) shows that b_i^2 will approximately estimate IV_i when there are no jumps during $[\frac{i-1}{N}, \frac{i}{N}[$. So we are in a dilemma: in the presence of a large jump, we would like to use b_i^2 as an estimate of SSJ_i, but in the absence of jumps, this produces too high values, as it then estimates $\mathrm{IV}_i > 0$, which is especially bad if we sum up estimators over subperiods, which is the natural way to come up with estimates of daily SSJ. The solution is to decrease b_i^2 in such a way that the estimator will have vanishing mean and as small as possible variance in the absence of jumps, while essentially measuring Δ^2 for large economic jumps. This leads to the following estimator:

$$\widehat{\mathrm{SSJ}}_i^{(l)} := b_i^2 - 1.4383\,\mathrm{uw}_i^2 - 1.4383\,\mathrm{lw}_i^2 - 2.0605\,\mathrm{uw}_i\,\mathrm{lw}_i, \tag{65}$$

whose mean and variance are essentially 0 and $3.7474\frac{\sigma^4_{\frac{i-1}{N}}}{N^2}$, resp., in the absence of jumps.

In order to estimate the squared positive jumps' sum

$$\mathrm{SSpJ}_i := \sum_{t \in [\frac{i-1}{N}, \frac{i}{N}[} 1_{\Delta_t > 0} \Delta_t^2, \tag{66}$$

we have to distinguish between positive and negative returns, i.e. between $r_i > 0$ and $r_i < 0$. Again, we can not just use $1_{r_i>0}\, b_i^2$, as this estimator will tend to estimate $\frac{1}{2}\,\mathrm{IV}_i$ in the absence of jumps. Similar considerations as above give

$$\widehat{\mathrm{SSpJ}}_i^{(l)} := \begin{cases} b_i^2 - 3.2047\,\mathrm{uw}_i^2 - 3.2047\,\mathrm{lw}_i^2 - 3.0301\,\mathrm{uw}_i\,\mathrm{lw}_i & , r_i > 0 \\ 1.7663\,\mathrm{uw}_i^2 + 1.7663\,\mathrm{lw}_i^2 + 0.9697\,\mathrm{uw}_i\,\mathrm{lw}_i & , r_i < 0 \end{cases}, \tag{67}$$

with mean and variance essentially 0 and $4.9360\frac{\sigma^4_{\frac{i-1}{N}}}{N^2}$, resp., in the absence of jumps.

In an analogous way,

$$\mathrm{SSnJ}_i := \sum_{t\in[\frac{i-1}{N},\frac{i}{N}[} 1_{\Delta_t<0}\Delta_t^2 \tag{68}$$

is estimated by

$$\widehat{\mathrm{SSnJ}}_i^{(l)} := \begin{cases} 1.7663\,\mathrm{uw}_i^2 + 1.7663\,\mathrm{lw}_i^2 + 0.9697\,\mathrm{uw}_i\,\mathrm{lw}_i & , r_i < 0 \\ b_i^2 - 3.2047\,\mathrm{uw}_i^2 - 3.2047\,\mathrm{lw}_i^2 - 3.0301\,\mathrm{uw}_i\,\mathrm{lw}_i & , r_i > 0 \end{cases}, \tag{69}$$

with mean and variance essentially 0 and $4.9360\frac{\sigma^4_{\frac{i-1}{N}}}{N^2}$, resp., in the absence of jumps.

In order to estimate daily quantities IV, SSJ, SSpJ, and SSnJ, we just sum up the contributions of all subperiods:

$$\widehat{\mathrm{IV}}^{(l)} := \sum_{i=1}^{N} \widehat{\mathrm{IV}}_i^{(l)}, \tag{70}$$

$$\widehat{\mathrm{SSJ}}^{(l)} := \sum_{i=1}^{N} \widehat{\mathrm{SSJ}}_i^{(l)}, \tag{71}$$

$$\widehat{\mathrm{SSpJ}}^{(l)} := \sum_{i=1}^{N} \widehat{\mathrm{SSpJ}}_i^{(l)}, \tag{72}$$

$$\widehat{\mathrm{SSnJ}}^{(l)} := \sum_{i=1}^{N} \widehat{\mathrm{SSnJ}}_i^{(l)}. \tag{73}$$

For these estimators, it is possible to write down a central limit theorem for $N \to \infty$, even in the presence of mathematical jumps. However, the

CLT crucially depends on the jumps being mathematical jumps, i.e. for the reasons discussed earlier, too high frequencies will result in gradual jumps being distributed among several subintervals, thereby rendering jump measurement inappropriate. We therefore refrain from discussing the CLT here and refer the interested reader to Klößner (2009).

6.2 Estimation in the Presence of Moderate Jumps

We now consider again the issue of estimating IV_i: the estimator $\widehat{\mathrm{IV}}_i^{(l)}$ makes only use of the wicks' lengths, for the sake of robustness with respect to jumps. In particular, $\widehat{\mathrm{IV}}_i^{(l)}$ does not include products of body and wicks' lengths, which will be small for moderate jumps as multiplication by a wick's length dampens the jump's influence on the body. Incorporating these products brings about a significant reduction in variance in the absence of jumps: while $\widehat{\mathrm{IV}}_i^{(l)}$ has a variance of approximately $0.7244\frac{\sigma^4_{\frac{i-1}{N}}}{N^2}$, we can come up with a new estimator $\widehat{\mathrm{IV}}_i^{(m)}$, whose approximate variance in the absence of jumps is only $0.2921\frac{\sigma^4_{\frac{i-1}{N}}}{N^2}$:

$$\widehat{\mathrm{IV}}_i^{(m)} := 0.4416\,(\mathrm{uw}_i^2 + \mathrm{lw}_i^2) + 1.3851\,\mathrm{uw}_i\,\mathrm{lw}_i + 1.1809\,(\mathrm{uw}_i\,b_i + \mathrm{lw}_i\,b_i). \quad (74)$$

In a similar way, estimators for SSJ_i, SSpJ_i, and SSnJ_i can be constructed:

$$\widehat{\mathrm{SSJ}}_i^{(m)} := 0.6576\,(\mathrm{uw}_i^2 + \mathrm{lw}_i^2) + 0.5552\,\mathrm{uw}_i\,\mathrm{lw}_i - 2.8089\,(\mathrm{uw}_i\,b_i + \mathrm{lw}_i\,b_i) + b_i^2, \quad (75)$$

with mean and variance given by 0 and $1.3014\frac{\sigma^4_{\frac{i-1}{N}}}{N^2}$, resp., in the absence of jumps, and

$$\widehat{\mathrm{SSpJ}}_i^{(m)} := \begin{cases} b_i^2 + 0.7706\,(\mathrm{uw}_i^2 + \mathrm{lw}_i^2) + 0.7394\,\mathrm{uw}_i\,\mathrm{lw}_i \\ \qquad -3.1847\,(\mathrm{uw}_i\,b_i + \mathrm{lw}_i\,b_i) & ,\ r_i > 0 \\ -0.1130\,(\mathrm{uw}_i^2 + \mathrm{lw}_i^2) - 0.1842\,\mathrm{uw}_i\,\mathrm{lw}_i \\ \qquad +0.3758\,(\mathrm{uw}_i\,b_i + \mathrm{lw}_i\,b_i) & ,\ r_i < 0 \end{cases}, \quad (76)$$

$$\widehat{\mathrm{SSnJ}}_i^{(m)} := \begin{cases} -0.1130\,(\mathrm{uw}_i^2 + \mathrm{lw}_i^2) - 0.1842\,\mathrm{uw}_i\,\mathrm{lw}_i \\ \qquad +0.3758\,(\mathrm{uw}_i\,b_i + \mathrm{lw}_i\,b_i) & ,\ r_i > 0 \\ b_i^2 + 0.7706\,(\mathrm{uw}_i^2 + \mathrm{lw}_i^2) + 0.7394\,\mathrm{uw}_i\,\mathrm{lw}_i \\ \qquad -3.1847\,(\mathrm{uw}_i\,b_i + \mathrm{lw}_i\,b_i) & ,\ r_i < 0 \end{cases}, \quad (77)$$

which both have approximate mean and variance of 0 and $0.7071\frac{\sigma^4_{\frac{i-1}{N}}}{N^2}$, resp., in the absence of jumps.

For these new estimators, it is also possible to construct estimators $\widehat{\mathrm{IV}}^{(m)}$, $\widehat{\mathrm{SSJ}}^{(m)}$, $\widehat{\mathrm{SSpJ}}^{(m)}$, and $\widehat{\mathrm{SSnJ}}^{(m)}$ for daily quantities IV, SSJ, SSpJ, and SSnJ, and to derive a CLT for these estimators. We will refrain from discussing these matters, however, for the same reasons as given above.

To see these estimators at work, let us have a look at Figures 3 and 4 again: these contrast the traditional estimators of SSJ with the new estimators $\widehat{\mathrm{SSJ}}^{(l)}$ and $\widehat{\mathrm{SSJ}}^{(m)}$. While it is evident that the new estimators do a much better job than the old ones, it can also be seen that $\widehat{\mathrm{SSJ}}^{(l)}$ performs slightly better than $\widehat{\mathrm{SSJ}}^{(m)}$. Another interesting fact is the decreasing behaviour of the estimators for (too) high frequencies, stemming from the gradual nature of the jump and necessitating moderate frequencies for the jump to be adequately measured.

We now discuss the new estimators' performance in the Monte Carlo experiment described above, where we have simulated 30,000 paths of Brownian motion subjected to gradual jumps[36]. Table 2 (p. 41) shows that $\widehat{\mathrm{IV}}^{(l)}$, being solely based on wicks' lengths, is unbiased for all frequencies, while $\widehat{\mathrm{IV}}^{(m)}$ has a small positive bias, which becomes negligible for high frequencies. Comparing $\widehat{\mathrm{IV}}^{(l)}$, $\widehat{\mathrm{IV}}^{(m)}$ to standard estimators shows the superior properties of the new estimators: in the high volatility setting, their performance at frequencies of 10 or 15 minutes is comparable to that of the old estimators at the 1 minute frequency, while in the low volatility setting, the new estimators sampled every 30 minutes even outperform the old ones sampled at their optimal frequency[37].
All in all, $\widehat{\mathrm{IV}}^{(l)}$ and $\widehat{\mathrm{IV}}^{(m)}$ are excellent estimators for diffusive volatility IV, both at high and moderate frequencies, due to their very effective using of intradaily highs and lows.

The estimation of volatility due to gradual jumps turns out to be much harder, as Tables 3-5 (p. 42, 43) reveal: $\widehat{\mathrm{SSJ}}^{(l)}$ achieves a significant reduction of bias, esp. for moderate frequencies of 10 or 15 minutes, however, its

[36]See Subsection 4.3.

[37]Despite these results, sampling frequency should not be chosen too small in practice, as intradaily fluctuations of volatility could then no longer be captured.

drawback is the relatively high variance, which makes this estimator unnecessarily volatile in the absence of jumps. $\widehat{\mathrm{SSJ}}^{(m)}$ performs much better in terms of variance, admittedly at the cost of a higher bias than $\widehat{\mathrm{SSJ}}^{(l)}$. In the end, both estimators perform similarly in terms of RMSE, improving estimation of SSJ as compared to prevalent estimators. However, further efficieny gains would be desirable, which may possibly be achieved by combining $\widehat{\mathrm{SSJ}}^{(l)}$ and $\widehat{\mathrm{SSJ}}^{(m)}$. The same assertions hold true with respect to the estimation of SSpJ and SSnJ.

Summing up, the new estimators do an outstanding job in the estimation of diffusive volatility IV, improve the estimation of SSJ, and make it possible to estimate SSpJ and SSnJ. It would be desirable to further improve the estimation of volatility due to gradual jumps, for instance by combining $\widehat{\mathrm{SSJ}}^{(l)}$ and $\widehat{\mathrm{SSJ}}^{(m)}$, or by using subsampled versions of these estimators. This, however, is left for future research.

6.3 Testing for the Presence of (Positive, Negative) Jumps

Both the theoretical results in Subsection 4.2 as well as Figure 5 show that only the relatively new tests by Jiang & Oomen (2008) and Lee & Ploberger (2009) have significant power against gradual jumps, while the tests by Aït-Sahalia & Jacod (2009b) and Barndorff-Nielsen & Shephard (2006) are essentially unable to detect these economic jumps. However, for moderate frequencies, all these tests suffer from quite substantial size distortions when spot volatility σ_t^2 exhibits intradaily variation[38], be it a deterministic diurnal pattern in volatility or stochastically varying volatility. Comparing Figure 5 to Figure 6 shows how power is severly reduced when using critical values obtained from simulating 30,000 paths of a two-factor stochastic volatility model without jumps. Therefore, in a related paper, Balter & Klößner (2010) developed jump tests which, to the author's best knowledge, are the only tests for jumps that are both capable of capturing gradual jumps and robust to heavy fluctuations in volatility. As these are also based on sparse sampling using intradaily highs and lows, we will shortly describe their construction, again referring the reader interested in more details to Balter & Klößner (2010). These tests, which can also separately detect positive and negative jumps, are build on the estimators $\widehat{\mathrm{SSJ}}^{(m)}$ for SSJ, $\widehat{\mathrm{SSpJ}}^{(t)}$ for SSpJ,

[38]For details on the following as well as for testing for jumps under high volatility of volatility, we refer the reader to Balter & Klößner (2010).

and $\widehat{\text{SSnJ}}^{(t)}$ for SSnJ, where $\widehat{\text{SSpJ}}^{(t)}$ and $\widehat{\text{SSnJ}}^{(t)}$ are given by the sums of

$$\widehat{\text{SSpJ}}_i^{(t)} := \begin{cases} 1.3982\,(\text{uw}_i^2 + \text{lw}_i^2) - 3.968\, b_i(\text{uw}_i + \text{lw}_i) & \\ \qquad +2.0902\ \text{uw}_i\,\text{lw}_i + b_i^2 & , \quad r_i > 0 \\ 0 & , \quad r_i < 0 \end{cases} , \tag{78}$$

$$\widehat{\text{SSnJ}}_i^{(t)} := \begin{cases} 0 & , \quad r_i > 0 \\ 1.3982\,(\text{uw}_i^2 + \text{lw}_i^2) - 3.968\, b_i(\text{uw}_i + \text{lw}_i) & \\ \qquad +2.0902\ \text{uw}_i\,\text{lw}_i + b_i^2 & , \quad r_i < 0 \end{cases} \tag{79}$$

over all intradaily subperiods. Under suitable conditions, with $\widehat{\text{IQ}}$ a consistent estimator of $\text{IQ} := \int_0^1 \sigma_t^4 dt$, the test statistics

$$\widetilde{\text{TJ}} := \sqrt{N}\frac{\sum_{i=1}^{N} \widehat{\text{SSJ}}_i^{(m)}}{\sqrt{1.3014\,\widehat{\text{IQ}}}},\ \widetilde{\text{TJ}}_p := \sqrt{N}\frac{\sum_{i=1}^{N} \widehat{\text{SSpJ}}_i^{(t)}}{\sqrt{0.8602\,\widehat{\text{IQ}}}},\ \widetilde{\text{TJ}}_n := \sqrt{N}\frac{\sum_{i=1}^{N} \widehat{\text{SSnJ}}_i^{(t)}}{\sqrt{0.8602\,\widehat{\text{IQ}}}} \tag{80}$$

are asymptotically standard normal in the absence of mathematical (positive, negative) jumps. For these tests to be able to detect gradual jumps, they have to be applied to moderately sampled data, which renders their approximation by a standard normal distribution more or less inappropriate, especially when volatility is changing fast. Balter & Klößner (2010) circumvent this problem by replacing the standard normal distribution in (80) by Pearson distributions, with the approximating distribution's parameters computed on a day by day basis, taking into account the variation of intraday volatility, as measured by the Gini coefficient of intradaily estimates $\widehat{\text{IV}}_i^{(m)}$.[39] Figures 5 and 6 show the superior power of $\widetilde{\text{TJ}}$, esp. when volatility is varying heavily. Comoparing these figures, it also becomes evident that the new test is the only one for which critical values under Brownian motion are essentially the same as under high volatility of volatility.

7 Empirical Application

As we have already documented by Figures 3 and 4, the new estimators $\widehat{\text{SSJ}}^{(l)}$ and $\widehat{\text{SSJ}}^{(m)}$ do a good job in capturing the economic jumps in DJIA and GE on January 3, 2001. Using the estimators $\widehat{\text{IV}}_{i,N}^{(l)}$, $\widehat{\text{SSJ}}_{i,N}^{(l)}$, $\widehat{\text{SSpJ}}_{i,N}^{(l)}$, and $\widehat{\text{SSnJ}}_{i,N}^{(l)}$,

[39] For details see Balter & Klößner (2010).

it is possible to have a look at the behavior of both diffusive volatility and jumps within the day, see Figure 9.[40] It is easy to locate the timing of the jump, shortly after 1 p.m., by inspecting the intradaily estimates $\widehat{\mathrm{SSJ}}_{i,N}^{(l)}$ of the jumps' contribution to volatility or the corresponding estimates $\widehat{\mathrm{SSpJ}}_{i,N}^{(l)}$ for positive jumps, while the contribution of negative jumps, estimated by $\widehat{\mathrm{SSnJ}}_{i,N}^{(l)}$, meanders around 0. Another interesting insight we can infer from Figure 9 is the increase in diffusive volatility $\widehat{\mathrm{IV}}_{i,N}^{(l)}$ following immediately after the jump.
Figure 10 shows the analogon to Figure 9 for the correspondig estimators suited to moderate jumps. Although the jump is rather large, these estimators still work properly, with the additional benefit of less variation in the estimates, as can be seen in particular by looking at the estimates of $\mathrm{SSnJ}_{i,N}$, which are much less volatile than their counterparts in Figure 9, as predicted by theory.

With a sample of several days, it is also possible to compute diurnal patterns for the intradaily estimates of volatility due to diffusion, positive, and negative jumps. For DJIA, this has been done using 5 minutes OHLC data of DJIA ranging from January 2001 to December 2006. The results are displayed in Figures 11 and 12: diffusive volatility as well as jump activity are by far highest at opening time and exhibit a U-shaped pattern after opening with an additional spike at 10 a.m., probably due to occasional news releases at that time.
The U-shaped diurnal pattern in volatility testifies volatility's changing over time. To measure the intraday variation of volatility, Balter & Klößner (2010) compute the Gini coefficients of $\widehat{\mathrm{IV}}_{i,N}$ on a daily basis. Figure 13 displays the density of these Gini coefficients for different data generating processes:

- Brownian motion,
- Brownian motion with a deterministic diurnal volatility pattern approximating empirically observed U-shaped volatility:

$$\sigma_t^2 = \sigma_0^2 \left(\frac{23}{36} + 4\,(t - \frac{7}{12})^2 \right) \; (t \in [0,1]),$$

- Merton jump-diffusions with variance of the jumps equal to 5% and 20% of squared volatility, resp.,

[40]For all the figures presented in this section, similar results for other shares are available from the author upon request.

- Brownian motion with gradual jumps,
- the two-factor stochastic volatility model SV2F used in Huang & Tauchen (2005) and Balter & Klößner (2010),
- DJIA from 2001 to 2006,
- and GE from 2001 to 2008.

Figure 13 tells us that both mathematical and economic jumps induce only a small increase in estimated Gini coefficients, while imposing a diurnal volatility pattern has a slightly stronger effect, albeit it can not explain the empirically observed Gini coefficients of GE and DJIA[41]. For Gini coefficients of intradaily volatility estimates to behave similarly to empirical ones, incorporating high volatility of volatility, as e.g. in the SV2F model, is indispensable. However, as already mentioned earlier, with the exception of the tests by Balter & Klößner (2010), all test procedures for jumps suffer from severe size distortions when volatility fluctuates heavily. In Figures 14 and 15, we therefore not only present the empirical rejection rates for all the tests discussed in this paper, but also the rejection rates resulting from the use of critical values obtained from simulating 30,000 paths of the SV2F model.

With respect to data, we emphasize that in contrast to other researchers, we did not leave out data from the first or last minutes of trading time, which was 9:30 a.m. to 4 p.m., entailing $N = 78$ when working with 5 minutes frequency and $N = 39$ when aggregating to 10 minutes data. This is most probably the reason for the comparatively low empirical jump detection rate of the Lee/Ploberger test, which in contrast to the other tests can by its construction not detect jumps occuring during the first l subperiods, which comprise 45 (90) minutes for 5 minutes data when using $l = 2 \ln N$ ($l = 4 \ln N$) and even 80 (150) minutes when sampling every 10 minutes. This effect is quite severe due to the high jump activity at market opening and at 10 a.m., as revealed by the diurnal patterns in Figures 11 and 12, resulting in this test being the one with lowest empirical jump detection rate.

As we have already pointed out in Section 4.2, gradual and mathematical jumps tend to cancel out each other in the computation of the Barndorff-Nielsen/Shephard and Aït-Sahalia/Jacod test statistics, so that it is no wonder that we find these tests' detection rates to be above that of the Lee/Ploberger test, but well below the one of the Jiang/Oomen test[42].

[41] Other assets induce similar Gini coefficients.

[42] This effect becomes much less prominent when size-adjusting is applied, as the Jiang/Oomen test has the largest size distortion of all tests considered here, cf. Balter & Klößner (2010).

The detection rate of the test by Balter & Klößner (2010) is by far the highest, esp. after size-adjusting the results. The reason presumably is that this test is unaffected by all the above mentioned shortcomings: every subperiod is taken care of, gradual and mathematical jumps both enlarge the test statistic, regardless of their sign. The last point, i.e. robustness to the jumps' sign, probably causes the higher detection rate than that of the Jiang/Oomen test, which, as seen earlier, is shifted to the right by positive jumps, while shifted to the left by negative jumps. Therefore, positive and negative jumps happening on the same day will at least partially cancel out each other, thereby possibly resulting in the Jiang/Oomen test's inability to detect the fact that there have been jumps. This assertion can be backed up using the asymmetric jump tests by Balter & Klößner (2010) to separately detect positive and negative jumps: indeed, the proportion of days with significant[43] jumps of both sign to days with at least one significant jump always exceeds 13%, varying with the asset and frequency under consideration.

For the sake of fairness, we emphasize that the test by Balter & Klößner (2010) is the only one that makes use of and thus requires intradaily OHLC data: it essentially uses a three times higher amount of data, i.e. not only the conventional subperiod's return, but also the subperiod's highest and lowest return, so that comparing this test to other tests is somewhat unfair, as using more data should go hand in hand with improved power.

As a result of our objective to capture not only mathematical jumps, but also economic ones, we find a dramatically higher contribution of jumps to volatility than previously reported. In particular, we find that market indexes exhibit much more jumps than individual stocks[44]. In this regard, we emphasize the fact that the test's ability to detect a jump on a particular day is, apart from the frequency employed, essentially determined by the ratio of jump size over spot volatility[45]. For this reason, it empirically happens that a jump in DJIA is detected at the 1% level on November 16, 2006, although the estimate $\widehat{\mathrm{SSJ}}^{(l)}$ is very small at $0.03\,\%^2$, the detection made possible by the very low volatility on that day, as estimated by $\widehat{\mathrm{IV}}^{(l)} = 0.038\,\%^2$, corresponding to an annual volatility of 3.082 %. In general, we find the volatility of individual stocks to be significantly higher than that of market indexes: for instance, the average daily estimate of diffusive volatility, as given by $\widehat{\mathrm{IV}}^{(l)}$, equals $2.411\,\%^2$ for GE (corresponding to an annual volatility of 24.551 %),

[43]At the 5% significance level.
[44]Additional results for other assets are available from the author upon request.
[45]cf. Lee & Ploberger (2009).

but only $0.282\,\%^2$ for DJIA (corresponding to an annual volatility of $8.396\,\%$), which explains why it is possible to detect more jumps for DJIA than for GE.

8 Concluding Remarks

In this paper, we have taken a closer look at economic jumps: from an econometric point of view, these are jumps in prices, e.g. due to the release of important new information, while in the data, they may manifest themselves not as instantaneous price movements, but rather as gradual, possibly continuous adjustments of prices to their new level. Completely overlooked by the econometric literature on measuring and detecting jumps in high-frequency data, there is a significant amount of papers documenting this phenomenon, under the heading 'partial price adjustment', stating that it typically takes several minutes until important news are completely incorporated into prices. This necessitates both sparse sampling as well as the development of new estimators and tests for jumps, esp. since we have documented that most existing estimators and tests do not cope well with gradual jumps. In particular, when estimating SSJ, we recommend to sample at frequencies not exceeding 5 minutes, with the additional benefit of avoiding microstructure noise contaminating the estimators[46]. By exploiting the information hidden in intradaily highs and lows, we have constructed new methods to disentangle volatility into its diffusive part and the contributions of positive and negative jumps. These estimators can be computed separately for each intradaily subperiod, thereby enabling us to locate jump times within days, to compute diurnal patterns of diffusive volatility and jump activity, and to compute measures for volatility of volatility.

As far as testing for jumps is concerned, using the test procedure of Balter & Klößner (2010), we have found rather large proportions of jump days, both for GE and DJIA. The reader should notice, however, that these jumps may by surprisingly small, yet still detectable thanks to the test's remarkable power, stemming to a substantial part from the use of intradaily highs and lows. As the word 'jump' is often associated with large price movements, it would be useful to have a test that can detect days with 'large' jumps, by deciding on whether the sum SSJ of squared jumps exceeds a given threshold, e.g. $1\,\%^2$. To develop such a test, we need a limit theory for some estimator of SSJ in the presence of economic jumps, the derivation of which we postpone

[46]For the purpose of avoiding noise distorting estimators, using data sampled at a frequency of 5 minutes seems to be quite popular recently, see e.g. Bollerslev et al. (2009b), Todorov (2010).

to future research.

Finally, we recall the reader's attention to the following issue: as we have pointed out several times, economic jumps might correspond to steep, yet continuous price paths in the data, which we in accordance with Barndorff-Nielsen et al. (2009) interpret as a noisy version of the unobserved discontinuous true prices. Obviously, the same data might stem from a latent continuous process exhibiting a rather large drift rate for a few minutes, and it is impossible to decide purely from the data which of these possibilities holds true. Therefore, the new procedures developed in this paper react sensitively to intraday periods of large drift, which is sensible provided that these periods are interpreted as belonging to jumps. For applications dictating another view on these periods, other procedures will prove more appropriate: for instance, testing for purely mathematical jumps, not treating periods of large drift as jumps, should rather be done by applying the tests of Aït-Sahalia/Jacod or Barndorff-Nielsen/Shephard, as these tests treat periods of large drift and mathematical jumps differently.

References

Adams, G., McQueen, G., & Wood, R. (2004). The Effects of Inflation News on High Frequency Stock Returns. *Journal of Business*, 77(3), 547–574.

Aït-Sahalia, Y. & Jacod, J. (2009a). Estimating the Degree of Activity of Jumps in High Frequency Financial Data. *Annals of Statistics*, 37, 2202–2244.

Aït-Sahalia, Y. & Jacod, J. (2009b). Testing for Jumps in a Discretely Observed Process. *Annals of Statistics*, 37, 184–222.

Andersen, T. G., Bollerslev, T., & Diebold, F. X. (2007). Roughing It Up: Including Jump Components in the Measurement, Modeling and Forecasting of Return Volatility. *The Review of Economics and Statistics*, 89(4), 701–720.

Andersen, T. G., Bollerslev, T., & Diebold, F. X. (2010). Parametric and Nonparametric Volatility Measurement. In Y. Aït-Sahalia & L. P. Hansen (Eds.), *Tools and Techniques*, volume 1 of *Handbook of Financial Econometrics* chapter 2, (pp. 67–137). Elsevier.

Andersen, T. G., Dobrev, D., & Schaumburg, E. (2008). Jump-Robust

Volatility Estimation using Nearest Neighbor Truncation. Working Paper.

Balter, J. & Klößner, S. (2010). Testing separately for positive and negative jumps in financial data with high volatility of volatility. Working Paper.

Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A., & Shephard, N. (2008a). Designing realised kernels to measure the ex-post variation of equity prices in the presence of noise. *Econometrica*, 76, 1481–1536.

Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A., & Shephard, N. (2009). Realised Kernels in Practice: Trades and Quotes. *Econometrics Journal*, 12(3), C1–C32.

Barndorff-Nielsen, O. E., Kinnebrock, S., & Shephard, N. (2008b). Measuring downside risk — realised semivariance. Working paper.

Barndorff-Nielsen, O. E. & Shephard, N. (2004). Power and Bipower Variation with Stochastic Volatility and Jumps. *Journal of Financial Econometrics*, 2(1), 1–37.

Barndorff-Nielsen, O. E. & Shephard, N. (2006). Econometrics of Testing for Jumps in Financial Econometrics Using Bipower Variation. *Journal of Financial Econometrics*, 4(1), 1–30.

Barndorff-Nielsen, O. E., Shephard, N., & Winkel, M. (2006). Limit theorems for multipower variation in the presence of jumps. *Stochastic Processes and their Applications*, 116(5), 796–806.

Bollerslev, T., Kretschmer, U., Pigorsch, C., & Tauchen, G. (2009a). A Discrete-Time Model for Daily S&P500 Returns and Realized Variations: Jumps and Leverage Effects. *Journal of Econometrics*, 150(2), 151–166.

Bollerslev, T., Tauchen, G., & Zhou, H. (2009b). Expected Stock Returns and Variance Risk Premia. *Expected Stock Returns and Variance Risk Premia*, 22(11), 4463–4492.

Borodin, A. N. & Salminen, P. (2002). *Handbook of Brownian Motion - Facts and Formulae*. Probability and its Applications. Basel et al.: Birkhäuser, 2. edition.

Carr, P. & Wu, L. (2009). Variance Risk Premiums. *The Review of Financial Studies*, 22(3), 1311–1341.

Chen, C. X. & Rhee, S. G. (2010). Short sales and speed of price adjustment: Evidence from the Hong Kong stock market. *Journal of Banking & Finance*, 34(2), 471–483.

Corsi, F., Pirino, D., & Renò, R. (2008). Volatility forecasting: the jumps do matter. Working paper, University of Siena.

Ederington, L. H. & Lee, J. H. (1995). The Short-Run Dynamics of the Price Adjustment to New Information. *Journal of Financial and Quantitative Analysis*, 30(1), 117–134.

Engle, R. F. & Gallo, G. M. (2006). A multiple indicators model for volatility using intra-daily data. *Journal of Econometrics*, 131(1), 3–27.

Fair, R. C. (2002). Events That Shook the Market. *Journal of Business*, 75(4), 713–731.

Huang, X. & Tauchen, G. (2005). The Relative Contribution of Jumps to Total Price Variance. *Journal of Financial Econometrics*, 3(4), 456–499.

Jiang, G. J. & Oomen, R. (2008). Testing for jumps when asset prices are observed with noise- A "swap variance" approach. *Journal of Econometrics*, 144(2), 352–370.

Klößner, S. (2008). Estimating Volatility Using Intradaily Highs and Lows. Working Paper.

Klößner, S. (2009). Separating risk due to diffusion, positive jumps, and negative jumps. Working Paper.

Lee, S. S. & Mykland, P. A. (2008). Jumps in Financial Markets: A New Nonparametric Test and Jump Dynamics. *The Review of Financial Studies*, 21(6), 2535–2563.

Lee, T. & Ploberger, W. (2009). Rate-optimal Tests for Jumps in Diffusion Processes. Working Paper.

Merton, R. C. (1976). Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics*, 3(1–2), 125–144.

Mykland, P. A. & Zhang, L. (2009). Inference for Continuous Semimartingales Observed at High Frequency. *Econometrica*, 77(5), 1403–1445.

Patton, A. J. & Sheppard, K. (2009). Good Volatility, Bad Volatility: Signed Jumps and the Persistence of Volatility. Working Paper.

Rasmussen, T. B. (2009). Jump Testing and the Speed of Market Adjustment. CREATES Research Paper 2009-8.

Shiryaev, A. N. (1999). *Essentials of Stochastic Finance: Facts, Models, Theory*, volume 3 of *Advanced Series on Statistical Science & Applied Probability*. Singapur et al.: World Scientific.

Smidt, S. (1968). A New Look at the Random Walk Hypothesis. *The Journal of Financial and Quantitative Analysis*, 3(3), 235–261.

Taylor, S. J. (2005). *Asset Price Dynamics, Volatility, and Prediction*. Princeton, Oxford: Princeton University Press.

Todorov, V. (2010). Variance Risk-Premium Dynamics: The Role of Jumps. *The Review of Financial Studies*, 23(1), 345–383.

Veraart, A. E. D. (2010). Inference for the jump part of quadratic variation of Itô semimartingales. *Econometric Theory*, 26(2), 331–368.

High volatility ($\sigma^2 = 10^{-4}$):

	1 min	3 min	5 min	10 min	15 min	30 min
BPV	**0.080**	0.169	0.185	0.201	0.223	0.273
	(0.131)	(0.255)	(0.302)	(0.383)	(0.457)	(0.641)
TPV	**0.083**	0.140	0.143	0.163	0.186	0.244
	(0.139)	(0.237)	(0.278)	(0.374)	(0.458)	(0.673)
MinRV	**0.094**	0.181	0.160	0.134	0.145	0.195
	(0.161)	(0.294)	(0.322)	(0.396)	(0.475)	(0.694)
MedRV	**0.079**	0.162	0.152	0.135	0.152	0.207
	(0.134)	(0.261)	(0.296)	(0.361)	(0.436)	(0.643)
$\widehat{\mathrm{IV}}^{(l)}$	-0.009	-0.010	-0.010	-0.004	0.004	0.031
	(0.044)	(0.075)	(0.096)	(0.138)	(0.171)	(0.258)
$\widehat{\mathrm{IV}}^{(m)}$	**0.002**	0.006	0.015	0.042	0.066	0.111
	(0.027)	(0.049)	(0.066)	(0.106)	(0.141)	(0.214)

Low volatility ($\sigma^2 = 9 \cdot 10^{-6}$):

	1 min	3 min	5 min	10 min	15 min	30 min
BPV	**0.085**	0.161	0.155	0.121	0.116	0.130
	(0.101)	(0.188)	(0.186)	(0.158)	(0.157)	(0.181)
TPV	0.089	0.106	0.076	**0.066**	0.070	0.087
	(0.104)	(0.127)	(0.096)	**(0.092)**	(0.101)	(0.134)
MinRV	0.124	0.188	0.135	0.075	0.058	**0.050**
	(0.151)	(0.226)	(0.183)	(0.137)	**(0.127)**	(0.129)
MedRV	0.086	0.170	0.134	0.075	0.059	**0.054**
	(0.105)	(0.207)	(0.184)	(0.139)	**(0.128)**	(0.133)
$\widehat{\mathrm{IV}}^{(l)}$	-0.002	-0.001	-0.001	-0.001	0.000	0.004
	(0.004)	(0.007)	(0.009)	(0.013)	(0.017)	(0.029)
$\widehat{\mathrm{IV}}^{(m)}$	**0.000**	0.003	0.006	0.018	0.027	0.047
	(0.003)	(0.006)	(0.011)	(0.024)	(0.036)	(0.061)

Table 2: Bias (and RMSE) of different estimators for IV (in $\%^2$)

High volatility ($\sigma^2 = 10^{-4}$):

	1 min	3 min	5 min	10 min	15 min	30 min
RV − BPV	-0.423	-0.415	-0.361	-0.291	**-0.280**	-0.299
	(0.490)	(0.486)	(0.438)	**(0.398)**	(0.414)	(0.499)
RV − TPV	-0.427	-0.385	-0.319	-0.253	**-0.244**	-0.270
	(0.497)	(0.459)	(0.403)	**(0.390)**	(0.426)	(0.555)
RV − MinRV	-0.438	-0.426	-0.335	-0.224	**-0.202**	-0.222
	(0.513)	(0.517)	(0.447)	**(0.418)**	(0.460)	(0.614)
RV − MedRV	-0.423	-0.408	-0.328	-0.225	**-0.209**	-0.233
	(0.491)	(0.487)	(0.423)	**(0.384)**	(0.415)	(0.547)
$\widehat{\text{SSJ}}^{(l)}$	-0.335	-0.235	-0.166	-0.087	-0.062	**-0.060**
	(0.401)	(0.333)	**(0.318)**	(0.389)	(0.471)	(0.682)
$\widehat{\text{SSJ}}^{(m)}$	-0.360	-0.275	-0.225	**-0.197**	-0.209	-0.249
	(0.418)	(0.341)	**(0.314)**	(0.355)	(0.415)	(0.549)

Low volatility ($\sigma^2 = 9 \cdot 10^{-6}$):

	1 min	3 min	5 min	10 min	15 min	30 min
RV − BPV	-0.431	-0.408	-0.332	-0.212	-0.176	**-0.158**
	(0.496)	(0.473)	(0.395)	(0.290)	(0.256)	**(0.239)**
RV − TPV	-0.435	-0.354	-0.253	-0.157	-0.130	**-0.116**
	(0.500)	(0.413)	(0.303)	(0.223)	(0.204)	**(0.201)**
RV − MinRV	-0.470	-0.435	-0.312	-0.166	-0.118	**-0.078**
	(0.543)	(0.509)	(0.386)	(0.269)	(0.233)	**(0.207)**
RV − MedRV	-0.432	-0.418	-0.311	-0.167	-0.119	**-0.083**
	(0.497)	(0.488)	(0.386)	(0.270)	(0.234)	**(0.210)**
$\widehat{\text{SSJ}}^{(l)}$	-0.345	-0.246	-0.176	-0.091	-0.061	**-0.033**
	(0.397)	(0.289)	(0.216)	(0.155)	**(0.144)**	(0.156)
$\widehat{\text{SSJ}}^{(m)}$	-0.349	-0.256	-0.194	-0.134	**-0.124**	-0.136
	(0.401)	(0.298)	(0.234)	(0.191)	**(0.189)**	(0.219)

Table 3: Bias (and RMSE) of different estimators for SSJ (in $\%^2$)

High volatility ($\sigma^2 = 10^{-4}$):

	1 min	3 min	5 min	10 min	15 min	30 min
$\widehat{\text{SSpJ}}^{(l)}$	-0.167	-0.117	-0.081	-0.042	-0.033	**-0.033**
	(0.295)	**(0.280)**	(0.302)	(0.401)	(0.493)	(0.722)
$\widehat{\text{SSpJ}}^{(m)}$	-0.179	-0.137	-0.112	**-0.099**	-0.103	-0.123
	(0.292)	(0.241)	**(0.227)**	(0.260)	(0.307)	(0.410)

Low volatility ($\sigma^2 = 9 \cdot 10^{-6}$):

	1 min	3 min	5 min	10 min	15 min	30 min
$\widehat{\text{SSpJ}}^{(l)}$	-0.172	-0.123	-0.088	-0.046	-0.031	**-0.016**
	(0.276)	(0.202)	(0.153)	(0.114)	**(0.109)**	(0.125)
$\widehat{\text{SSpJ}}^{(m)}$	-0.174	-0.128	-0.097	-0.067	**-0.062**	-0.068
	(0.279)	(0.208)	(0.165)	**(0.137)**	(0.138)	(0.163)

Table 4: Bias (and RMSE) of different estimators for SSpJ (in $\%^2$)

High volatility ($\sigma^2 = 10^{-4}$):

	1 min	3 min	5 min	10 min	15 min	30 min
$\widehat{\text{SSnJ}}^{(l)}$	-0.168	-0.118	-0.086	-0.044	-0.030	**-0.027**
	(0.297)	**(0.284)**	(0.305)	(0.402)	(0.493)	(0.720)
$\widehat{\text{SSnJ}}^{(m)}$	-0.181	-0.138	-0.113	**-0.098**	-0.106	-0.126
	(0.295)	(0.244)	**(0.227)**	(0.263)	(0.310)	(0.412)

Low volatility ($\sigma^2 = 9 \cdot 10^{-6}$):

	1 min	3 min	5 min	10 min	15 min	30 min
$\widehat{\text{SSnJ}}^{(l)}$	-0.173	-0.123	-0.088	-0.045	-0.030	**-0.016**
	(0.277)	(0.202)	(0.152)	(0.113)	**(0.109)**	(0.126)
$\widehat{\text{SSnJ}}^{(m)}$	-0.175	-0.128	-0.097	-0.067	**-0.062**	-0.068
	(0.280)	(0.208)	(0.164)	**(0.137)**	(0.138)	(0.164)

Table 5: Bias of different estimators for SSnJ (in $\%^2$)

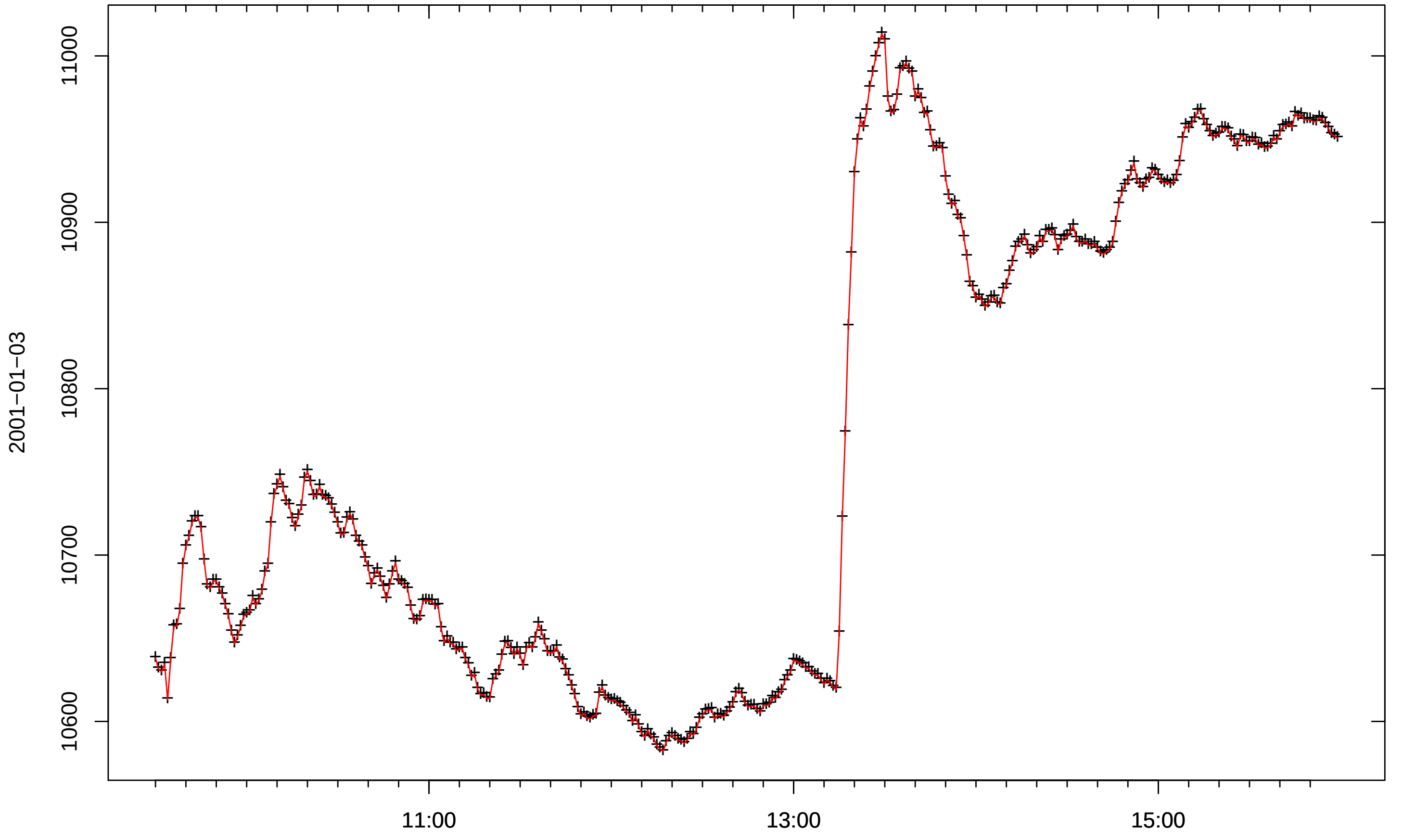

Figure 1: DJIA, 2001-01-03

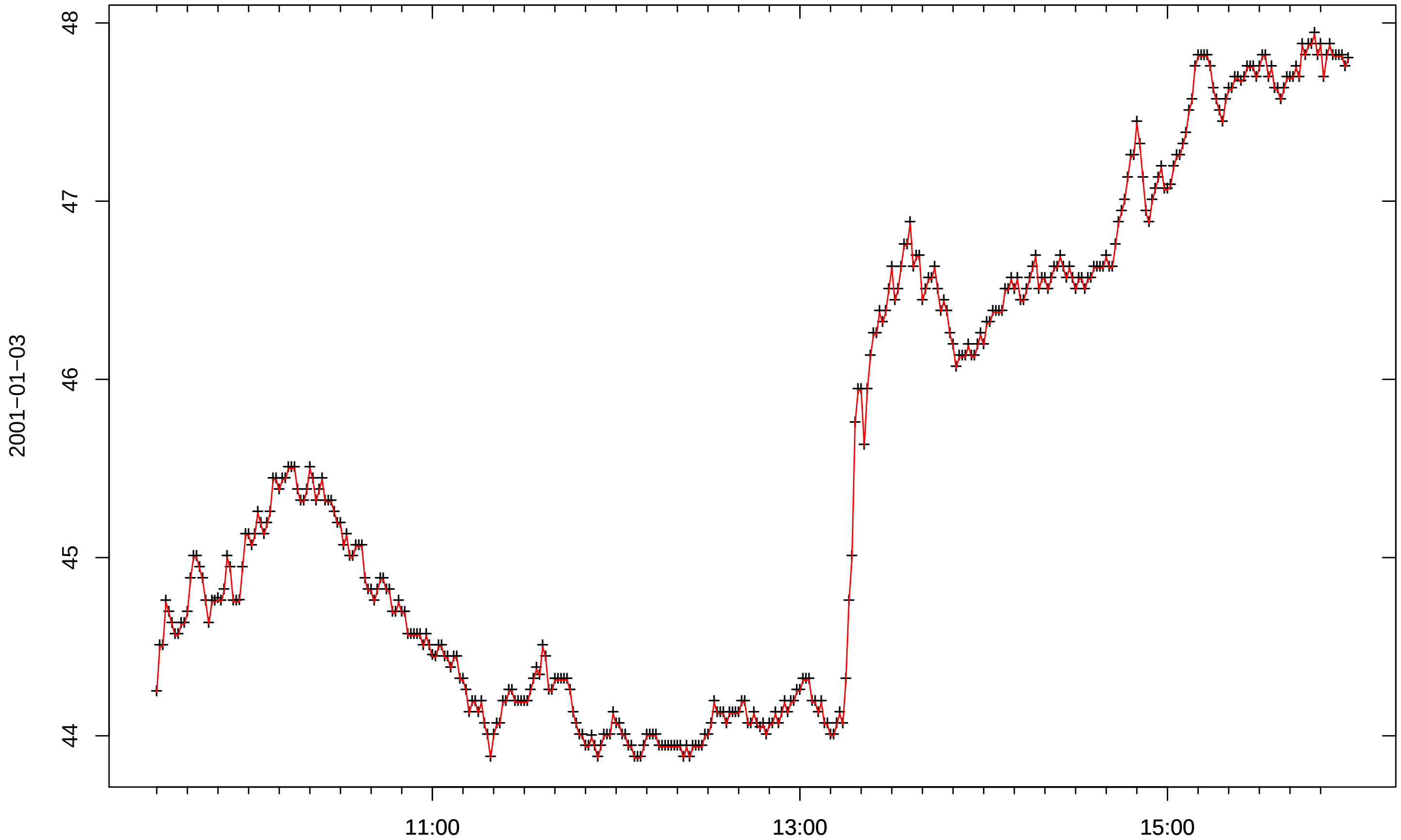

Figure 2: GE, 2001-01-03

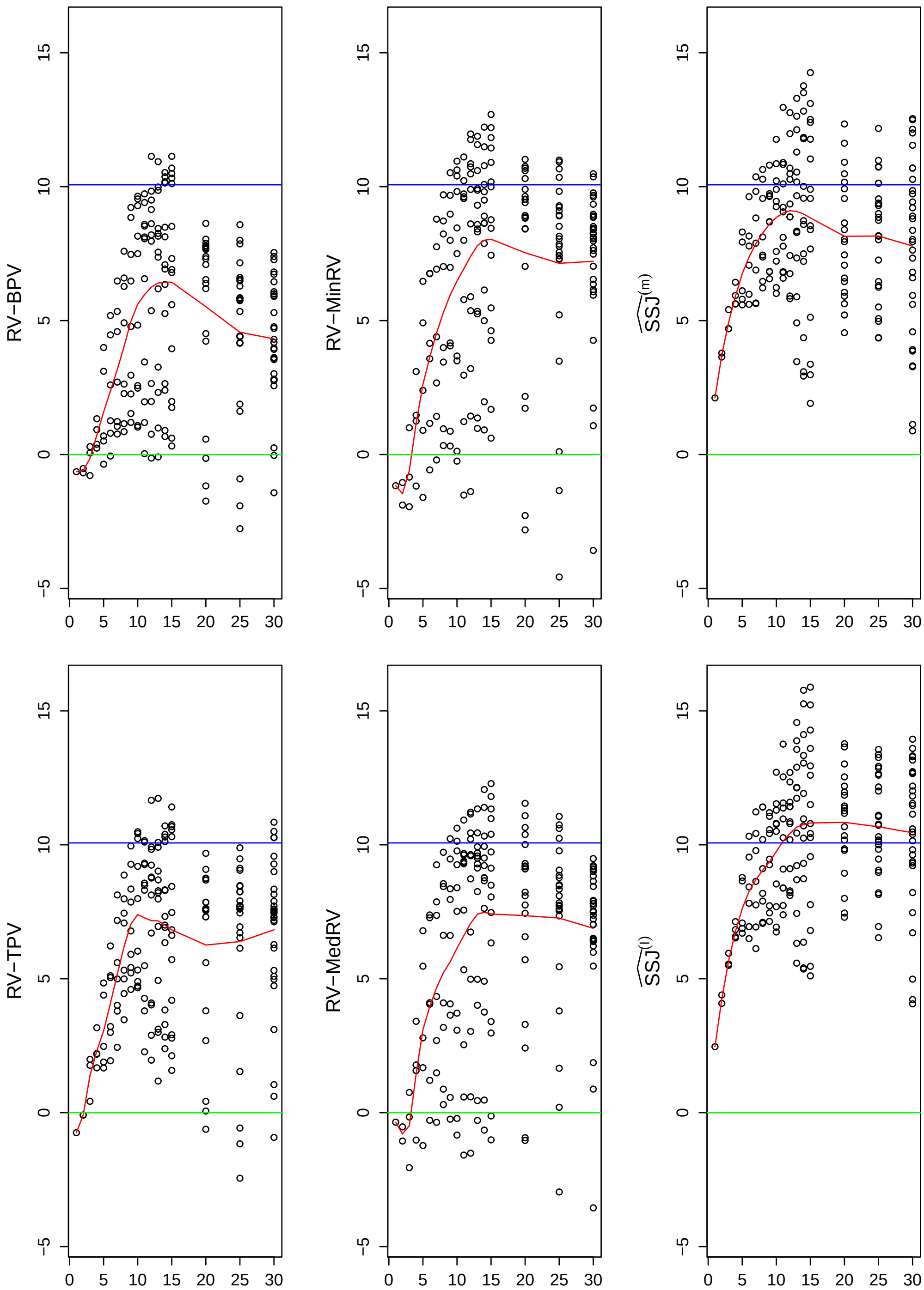

Figure 3: DJIA: Existing (first two columns) and new (last column) estimators for SSJ (in $\%^2$), 2001-01-03

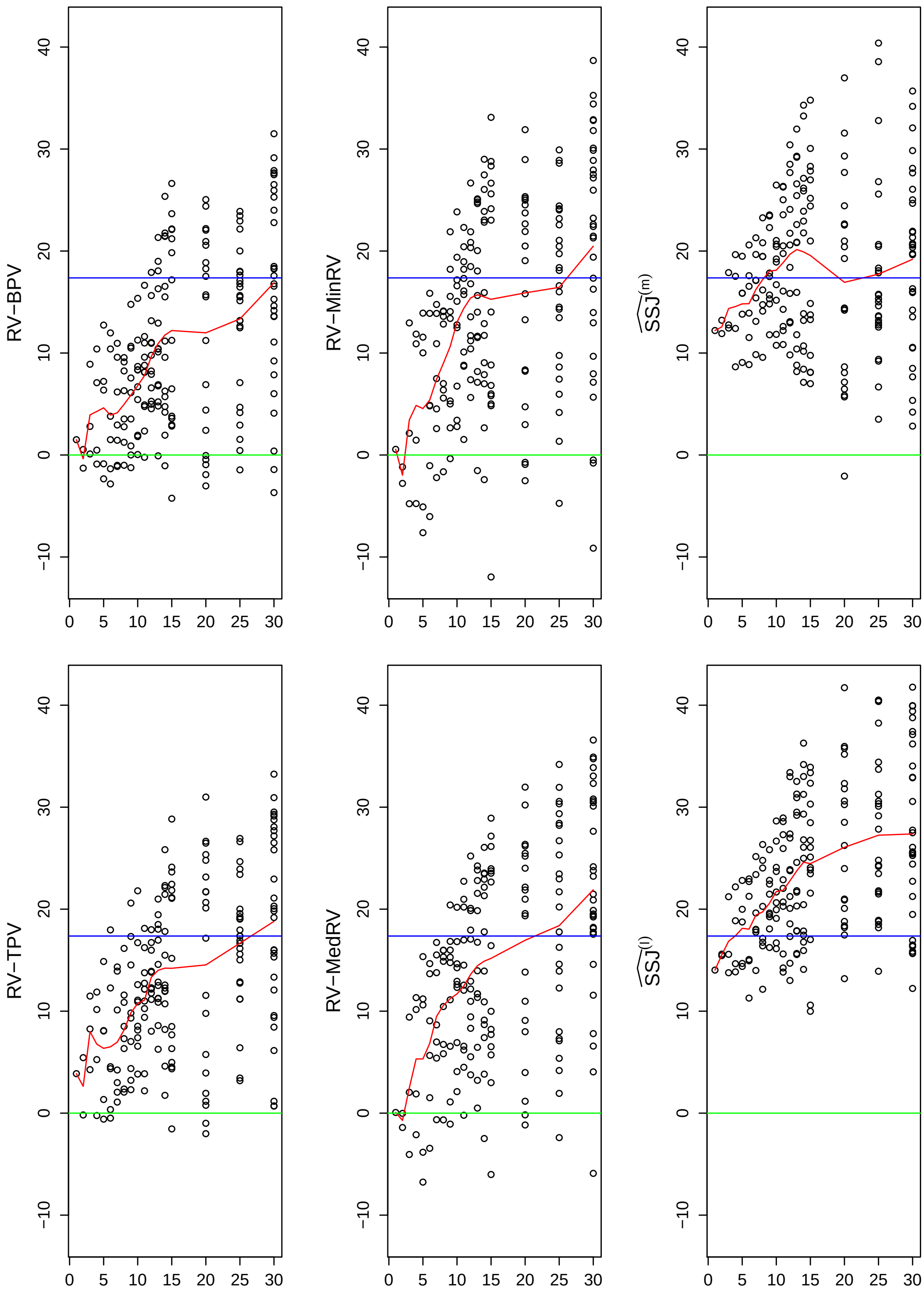

Figure 4: GE: Existing (first two columns) and new (last column) estimators for SSJ (in $\%^2$), 2001-01-03

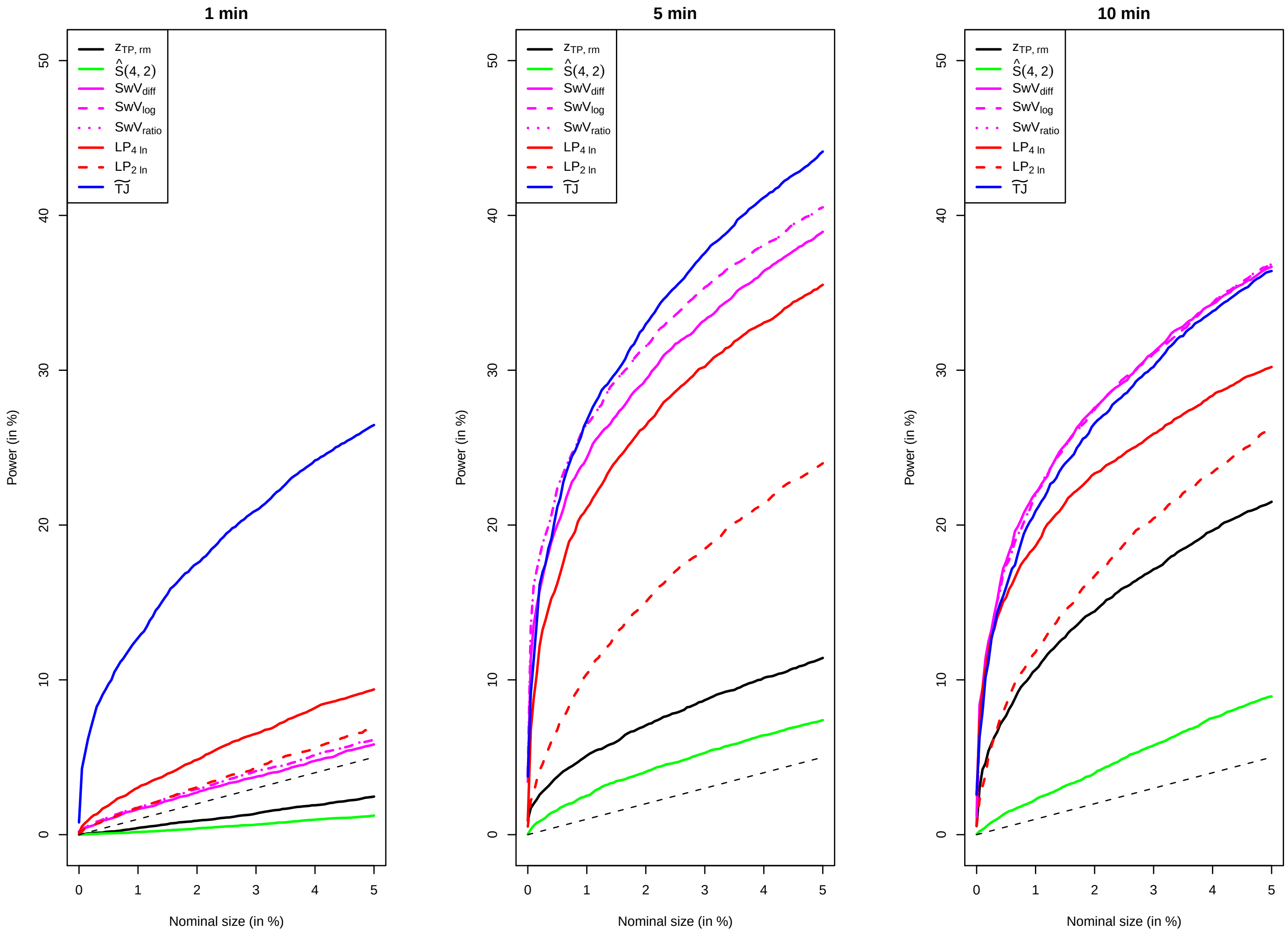

Figure 5: Size-adjusted (w.r.t. Brownian motion) power of different tests for Brownian motion with gradual jumps

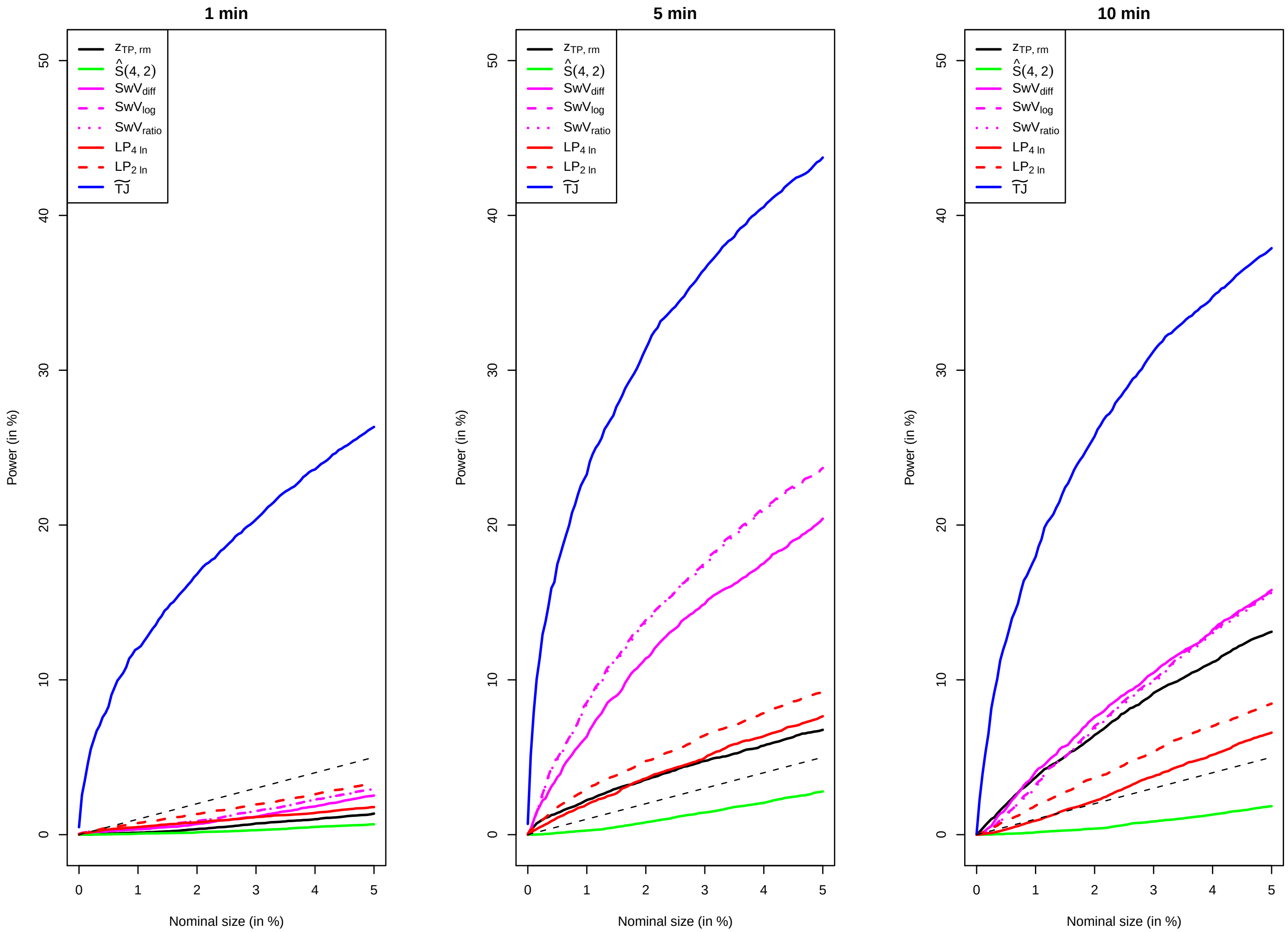

Figure 6: Size-adjusted (w.r.t. two-factor stochastic volatility model) power of different tests for Brownian motion with gradual jumps

Figure 7: Candlestick chart for simulated exponential Brownian motion, frequency: 5 minutes

Figure 8: Candlestick chart for DJIA, 2001-01-03, frequency: 5 minutes

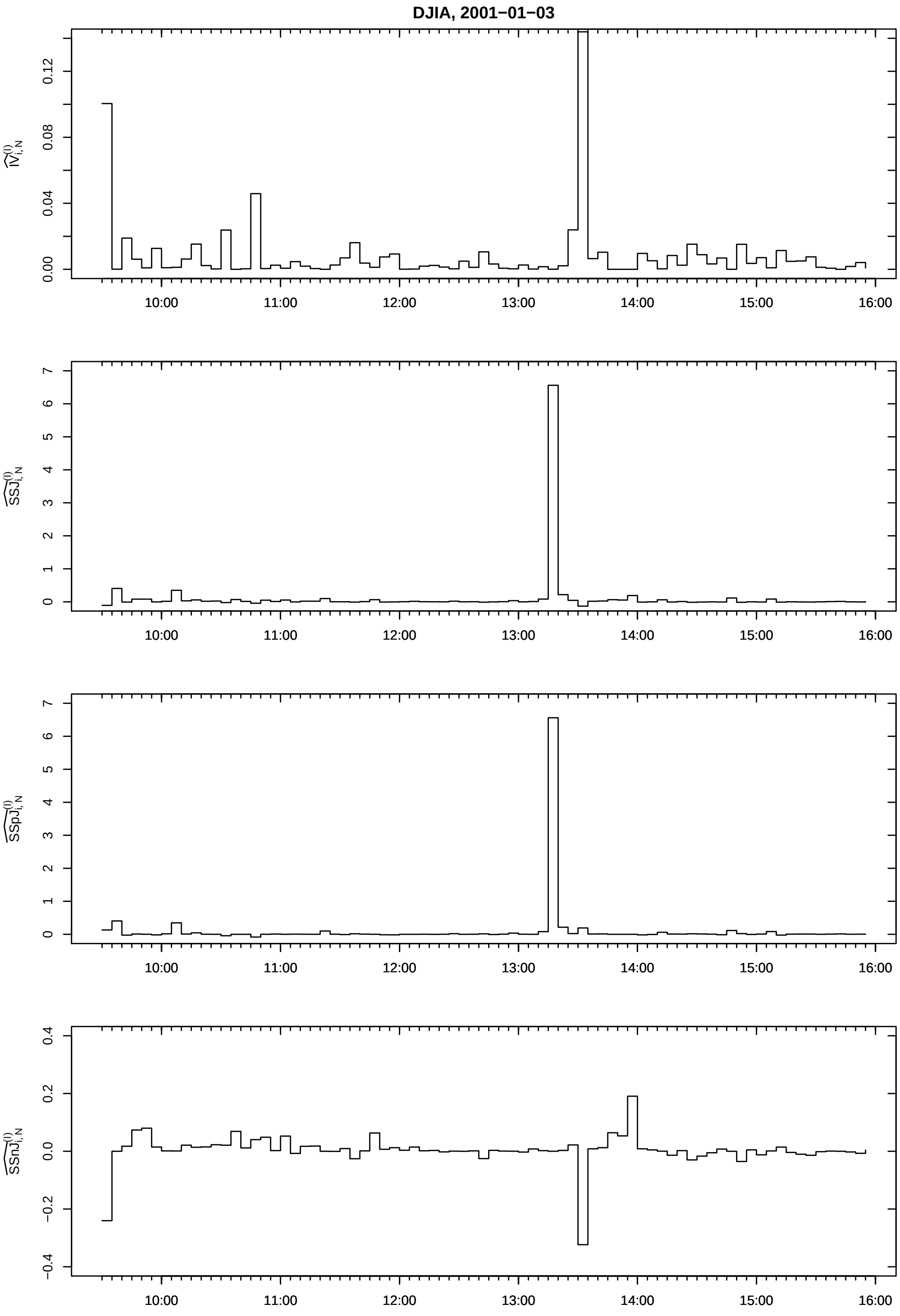

Figure 9: Intraday estimators suited to accomodate large jumps (in $\%^2$), frequency: 5 minutes

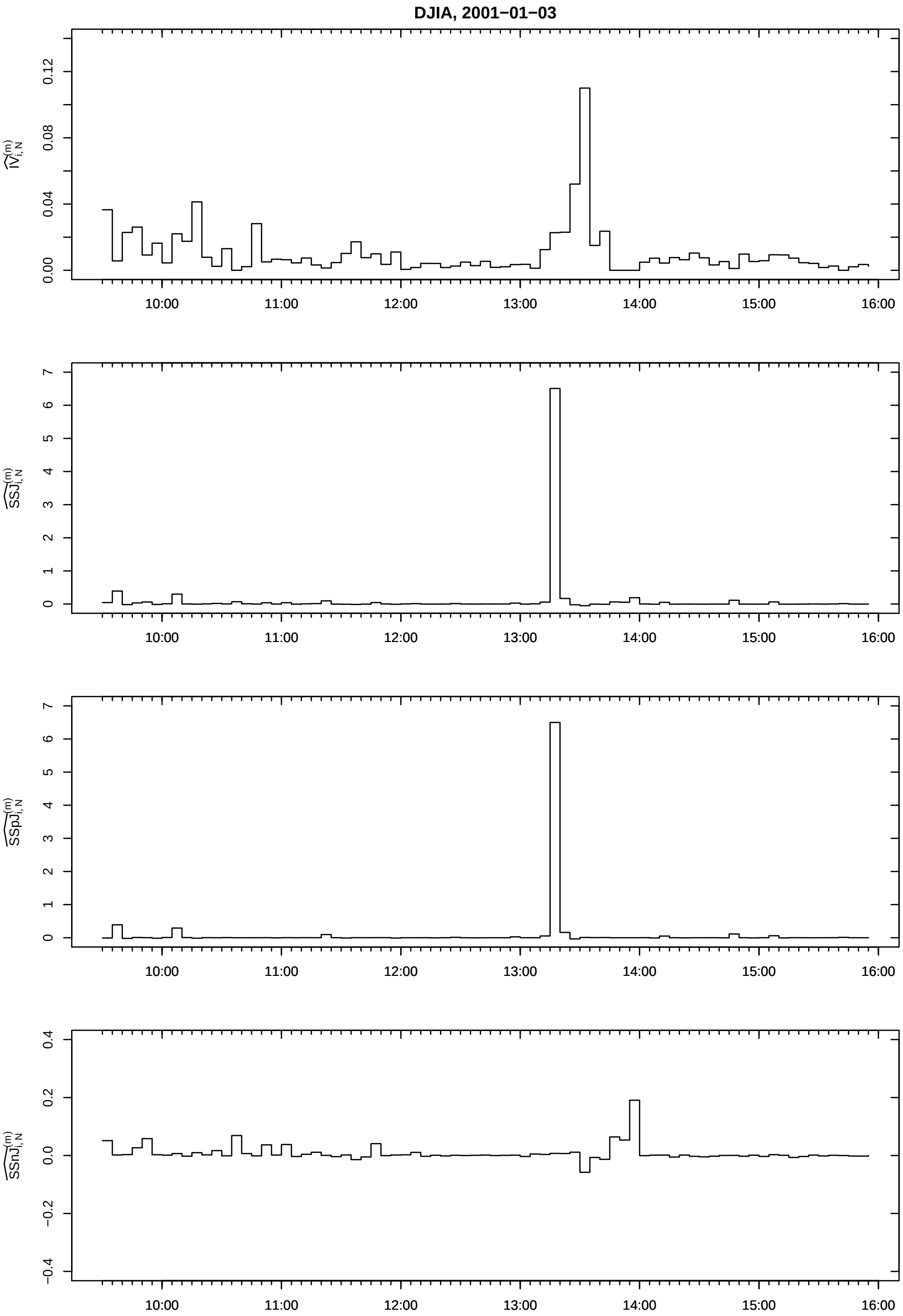

Figure 10: Intraday estimators suited to accomodate moderate jumps (in $\%^2$), frequency: 5 minutes

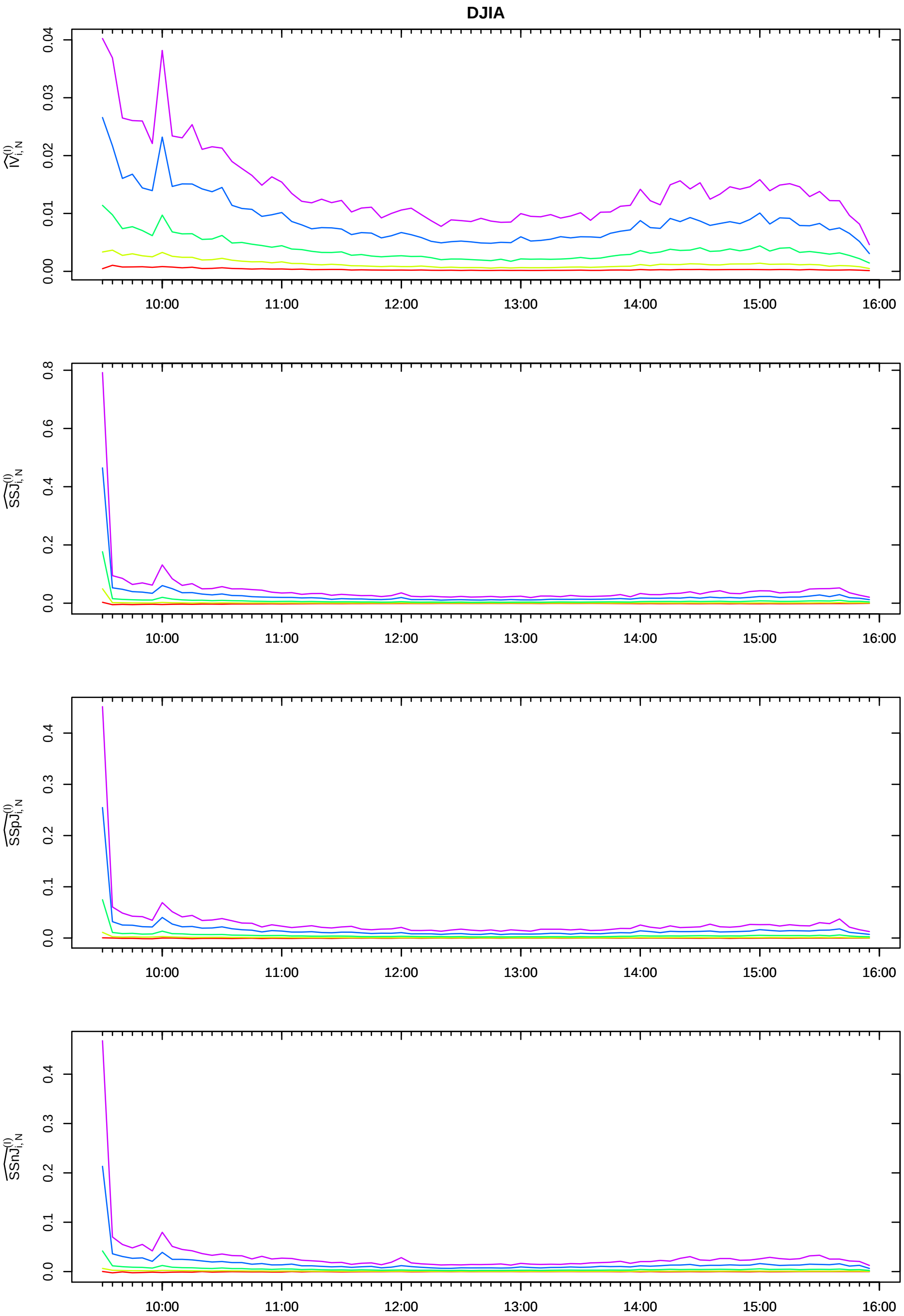

Figure 11: Diurnal patterns for intraday estimators suited to accomodate large jumps (in $\%^2$). Depicted are the following quantiles: 25%, 50%, 75%, 90%, 95% (from red to magenta).

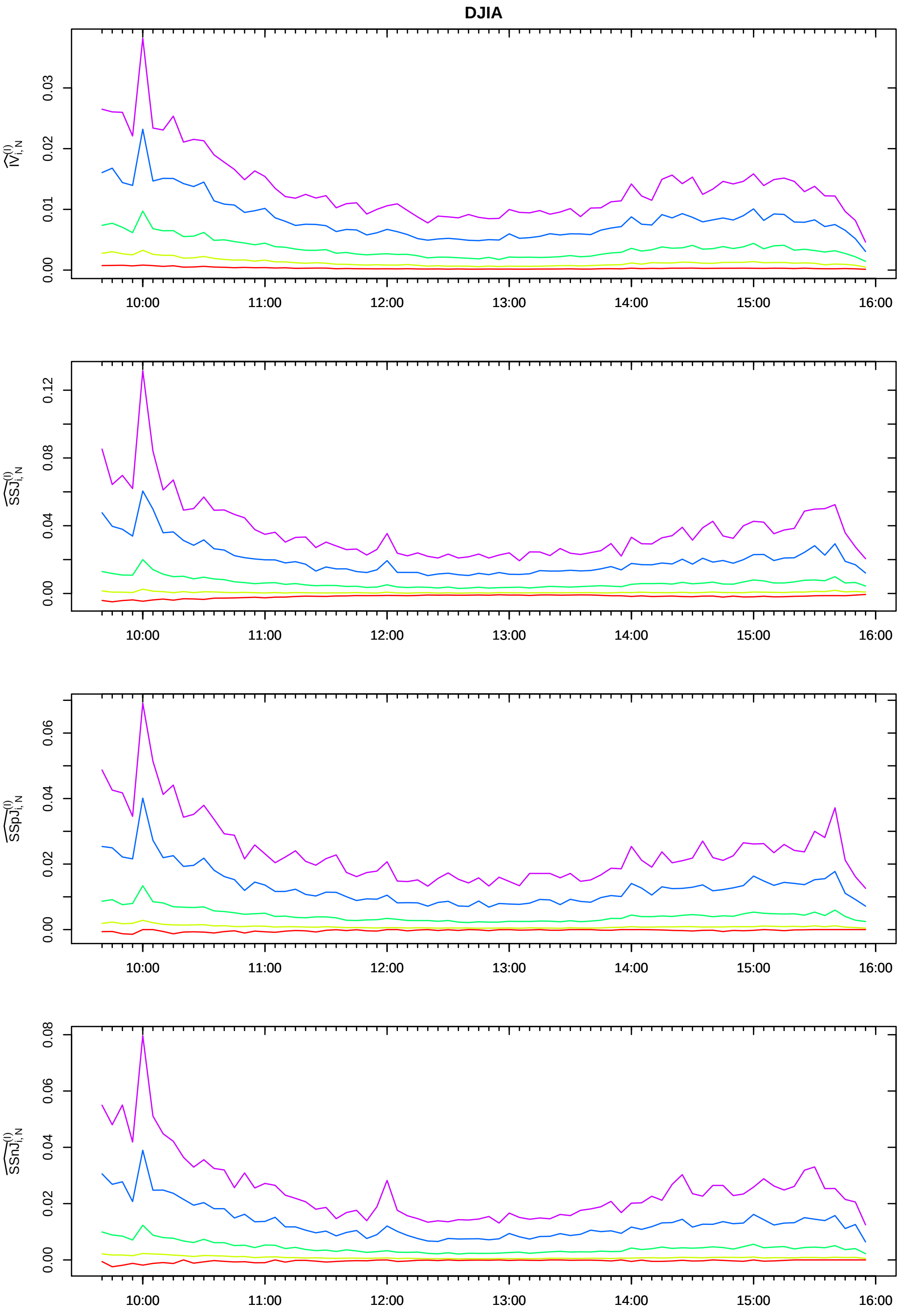

Figure 12: Diurnal patterns for intraday estimators suited to accomodate large jumps (in $\%^2$), with opening subperiod removed. Depicted are the following quantiles: 25%, 50%, 75%, 90%, 95% (from red to magenta).

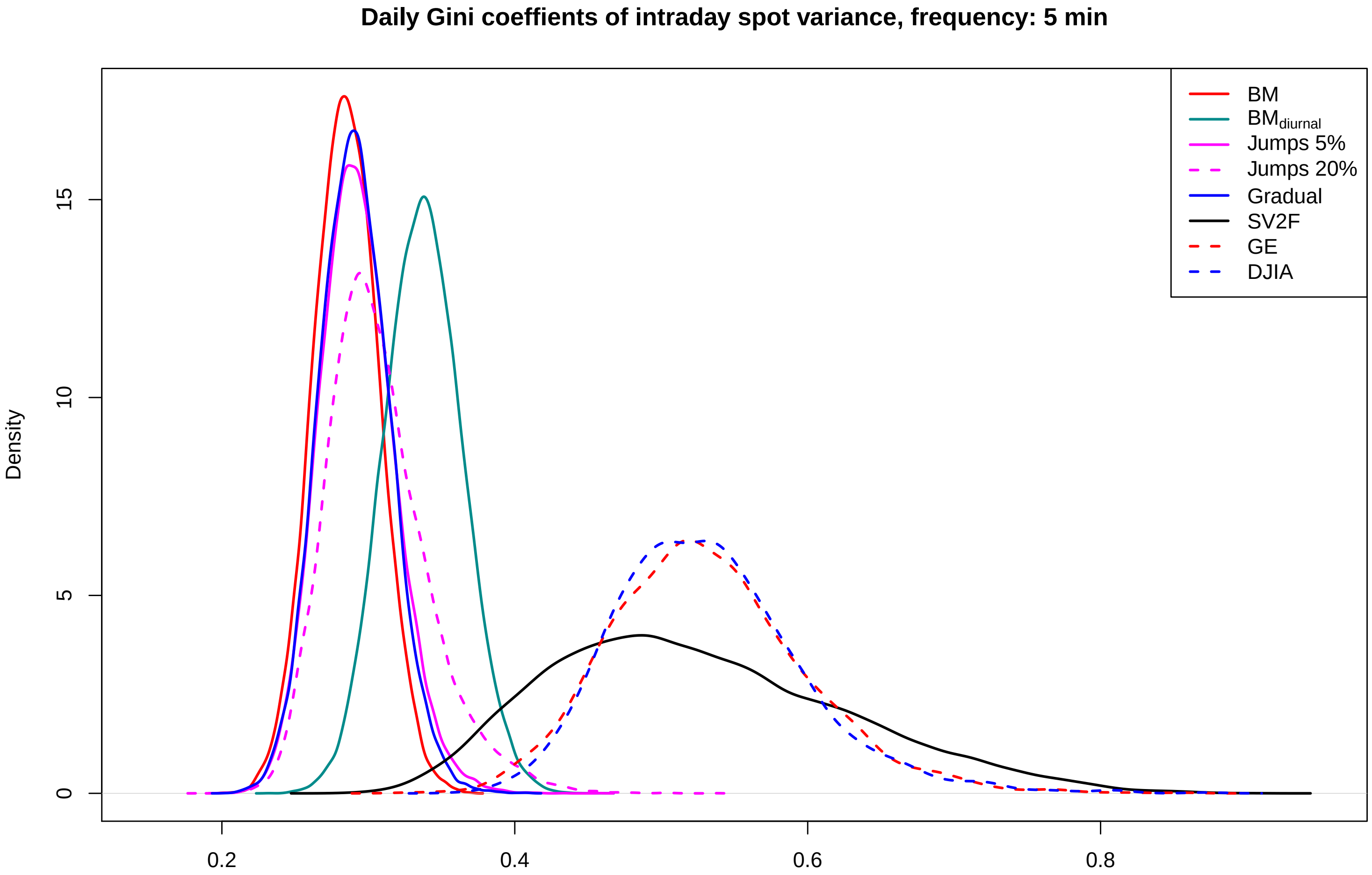

Figure 13: Daily Gini coefficients of intraday spot variance, frequency: 5 minutes

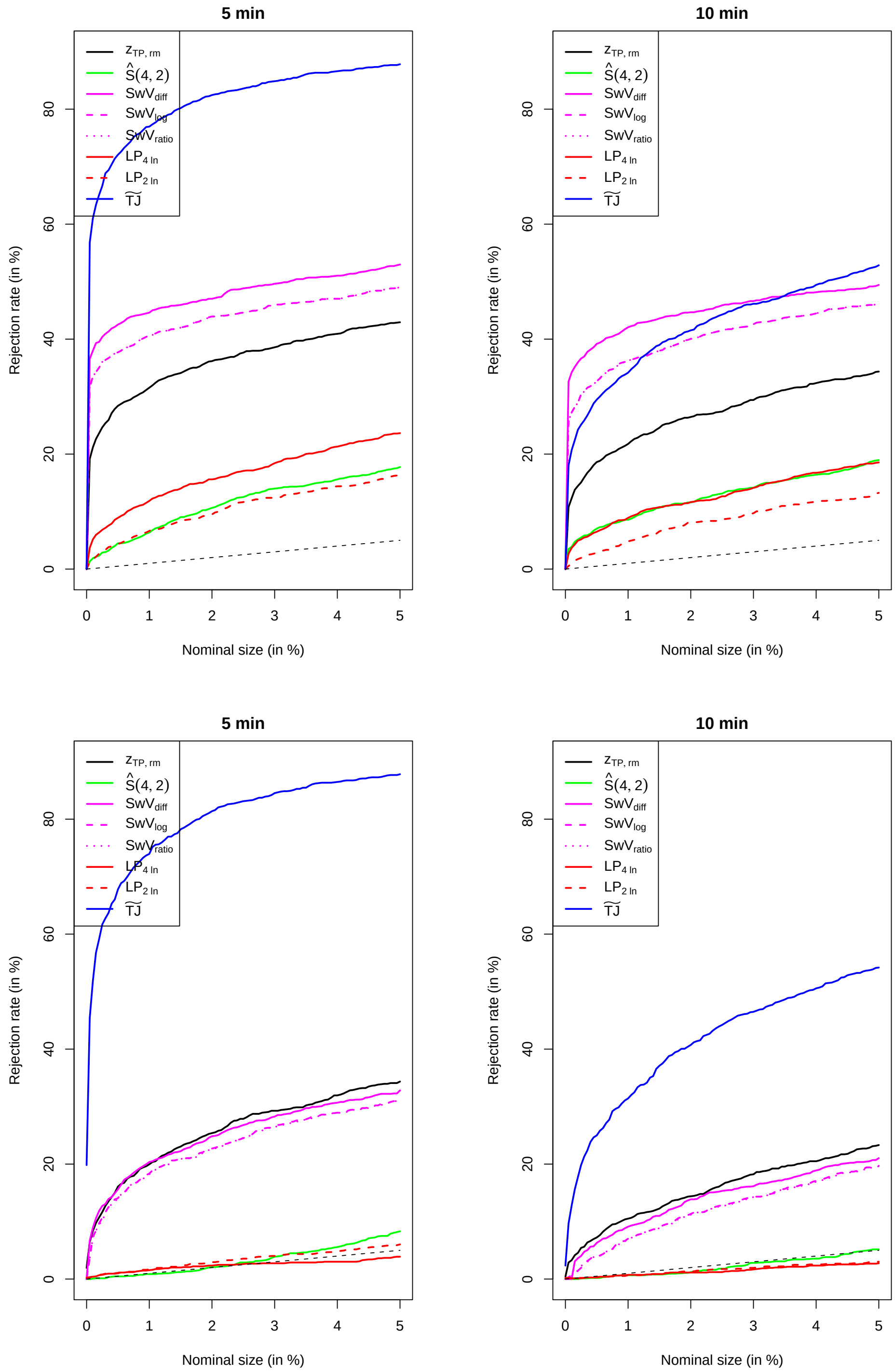

Figure 14: Empirical rejection rates for DJIA without (upper row) and with (lower row) size adjustment (w.r.t. two-factor stochastic volatility model) for different tests

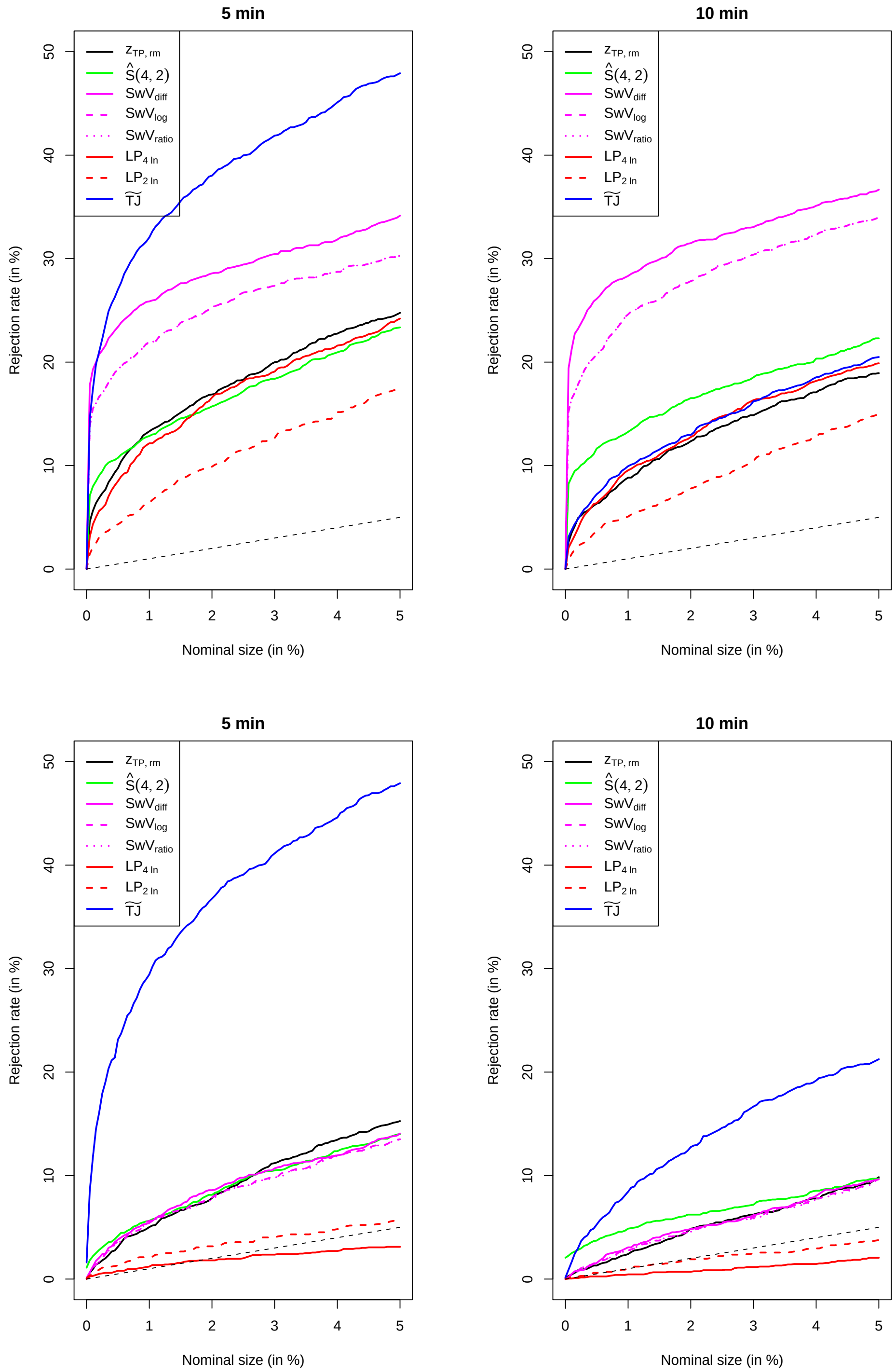

Figure 15: Empirical rejection rates for GE without (upper row) and with (lower row) size adjustment (w.r.t. two-factor stochastic volatility model) for different tests