Workshop on Financial Econometrics, 23-24 April 2010

Identifying Jumps in Financial Assets with a Comparison between Nonparametric Jump Tests

(with A-M Dumitru)

Giovanni Urga

Cass Business School, London (U.K.) & Bergamo University (Italy)

Outline

- Introduction and motivation
- Jump tests (?)
- Monte Carlo analysis: design and main findings
- -Extensions:
 - (a) Approximate finite sample distributions for ABD and LM
 - (b) Cross-performances of the tests
- Empirical application (?)

Conclusion and further work

Introduction and motivation

Modeling return dynamics requires the specification of a stochastic volatility component (accommodates the persistence in volatility) and of a jump component (takes care of the unpredictable large movements in the price process).

Identification of the time and the size of jumps has profound implications in risk management, portfolio allocation, derivatives pricing (AIT-Sahalia, 2004).

The use of jump diffusion models proved very difficult, as there are no closed forms of the likelihood function and in addition, the number of parameters to estimate is very high.

Approach 1: to focus on the popular class of affine models (Duffie et al., 2000): allow tractable estimation, but impose restrictive set of assumptions.

Approach 2: to use nonlinear volatility models. However, the estimation procedure, based on simulation methods, such as the Gallant and Tauchen (2002)'s efficient method of moments, is computationally demanding and too much dependent on the choice of an auxiliary model (Chernov et al., 2003; Andersen et al., 2002, for instance).

Development of nonparametric procedures to test for the presence of jumps in the path of a price process during a certain time interval or at certain point in time.

Such methods are

- 1) very simple to apply, they just require high frequency transaction prices or mid-quotes.
- 2) developed in a model free framework, incorporating different classes of stochastic volatility models.

Test procedures proposed:

Barndorff-Nielsen and Shephard (2006a, BNS), Andersen, Bollerslev and Dobrev. (2007, ABD), Lee and Mykland (2008, LM), Art-Sahalia and Jacod (2008, AJ), Jiang and Oomen (2008, JO) and Podolskij and Ziggel (2008, PZ).

All tests are based on CLT-type results that require an intraday sampling frequency that tends to infinity.

The test statistics are based on robust to jumps measures of variation in the price processes which are estimated by using either realized multi-power variations (Barndorff-Nielsen and Shephard, 2004, 2003, see) or threshold estimators (Mancini, 2009).

ABD(2007) and LM (2008) tests have the null hypothesis of continuity of the sample path at a certain moment, allowing for the exact identification of the time of a jump.

The other procedures have a null of continuity within a certain time period, such as a trading day.

Variety of nonparametric methodologies to identify jumps: which procedure should be preferred, and is the choice dependend upon data characteristics?.

Is the performance of the tests related to specific features of the data: sampling frequencies, levels of volatility, varying persistence in volatility, varying contamination with microstructure noise, varying jump size and jump intensity.

Such characteristics vary between classes of assets, as well as between different time periods.

For instance, equity prices are 'jumpier' than bond prices and markets in general have been more volatile and at the same time 'jumpier' during the last 2 years than before.

The main objective of this paper:

perform a thorough comparison among the six testing procedures via comprehensive Monte Carlo simulations, which embodies important features of financial data.

Size of all tests: our simulations are based on both constant and stochastic volatility models with varying persistence.

Power of all tests: we consider stochastic volatility models with jumps of different sizes arriving with varying intensity.

The simulation exercise may provide a set of guidelines to users of nonparametric tests for jumps

Additional contributions of this paper to the existing literature:

- (1) for ABD. (2007) and LM (2008) tests, we explore the benefits from using approximate finite sample distributions, generating critical values based on simulations (White (2000)'s Monte Carlo Reality Check approach).
- (2) we propose a procedure that combines tests and frequencies to reduce the probability of detecting spurious jumps.
- (3) we apply the tests to high frequency data on the US Treasury 2-, 5- 10- and 30- year bonds over a period lying between January 2003 and March 2004, sampled at different frequencies

Jump tests:

None of the procedures can test for the absence or presence of jumps in the model or data generating process.

They supply us with information on whether within a certain time interval or at a certain moment, the realization of the process is continuous or not.

ABD, LM assume the null of continuity of the sample path at time t_j .

AJ, BNS, JO, PZ: the null is of continuity of the sample path during a certain period, such as a trading day.

The alternative hypothesis implies discontinuity of the sample path, that is the occurrence of at least one jump.

For all procedures, under the null, the test statistics are asymptotically standard normal, though in some cases (ABD and LM) standard normal thresholds, like 99% or 95% quantiles, appear too liberal and more restrictive thresholds need to be used.

Apart from the procedures proposed by Ait-Sahalia and Jacod (2008) and Podolskij and Ziggel (2008), all other procedures work only when a finite number of jumps occur within a certain time interval. This is due to the fact that in all cases, the construction of the test statistics is based on realized multi-power variation estimators, which are robust only to a finite number of jumps.

For this reason, in the simulation set-up, we only consider processes with a finite number jumps (compound Poisson) and compare tests under this scenario.

In the light that the ABD and LM differ only in terms of the choice of the critical values, for a large part of our simulation exercise, we do not distinguish between the two of them .

Monte Carlo analysis: design...

First, we simulated a simple stochastic process with constant volatility. Second, simulated several stochastic volatility processes with leverage effects, with and without jumps and different levels of persistence in volatility as well as jump intensities and jump variances

- Under the null: benchmark model (SV1F):

$$dp(t) = \mu dt + \exp \left[\beta_0 + \beta_1 v(t)\right] dw_p(t)$$

$$dw_p(t) = \alpha_v v(t) dt + dw_v(t)$$

$$corr(dw_p, dw_v) = \rho$$

p(t) =log-price process, w standard B motions, v(t) volatility factor, μ drift of the process, α_v drift of the volatility process, ρ is the leverage effect.

SV1F: stochastic volatility model+one volatility factor+jumps under the alternative of discountinuos sample paths

SV2F: two volatility factors: one controls persistence of the process, the second controls high tails via volatility feedback component (Chernov et al, 2003)

Volatility factor enters in an exponential form (Chernov et al, 2003)

- Under le alternative: SVF1J model=SV1F+jumps
- Compound Poisson process with jump intensity λ
- Jump size distributed as $N(\mathbf{0}, \sigma_{jump}^2)$

further

- data simulated for 10000 trading days, for all models and under both hypotheses (continuity and discontinuity).
- the simulation of each path, we use an Euler discretization scheme based on increments of 1 second.
- sampling frequency: 1 sec, 1, min, 5min, 15 min, 30 min.

Remarks

- 1. For the intraday jump detection procedures, two sets of results:
- (a) Comparing ABD and LM with the other tests applied on time intervals equal to one trading day, we compute the test statistics for every moment t_i within a trading day and then pick up the maximum statistic as the final test for that day. We do not distinguish between the two procedures: ABD-LM. To contrast our results to the ones reported in Lee and Mykland (2008), we also adopt their strategy to calculate for the intraday tests both overall probabilities, as well as means and standard deviations of the intraday probabilities of spurious and nonspurious detection of jumps.
- (b) LM (2008) use critical values from the Gumbel distribution, ABD. (2007) suggest the Sidak approach, with the advantage of considering

the daily number of observations and is expect to work better in finite samples. However, it requires very low nominal sizes (10-5), whereas for all other procedure, we use 5% and 1% significance levels. In order to assure comparability with the other procedures and to gain better finite sample properties for the intraday tests, we approximate, by means of Monte Carlo simulations the finite sample distribution of the max for each window size considered in estimating the realized bipower variation which enters the test statistics. Then, critical values are picked up from this approximate finite-sample distribution. We dedicate a separate section (4.1) to compare results based on this procedure with results for the standard LM test.

2. AJ (2008) suggest two possible ways to estimate the variance of their test statistic in a robust to jumps manner. The fiirst on threshold estimators (Mancini 2009), the second one on realized multipower variations (Barndorff-Nielsen and Shephard, 2004, 2006b, 2003). In our simulations, we employ both versions.

3. In the paper we report results for a 5% significance level. Results for lower significance levels, i.e. 1%, 0.1% and 0.01%, are in line with the ones at 5%.

Monte Carlo analysis: ..and main findings

Constant volatility model

A very simple stochastic process with a diffusion parameter that remains constant through time. When we evaluate the power of the tests, we add to the diffusion term jumps of different sizes. Under this setup, extreme dynamics are caused only by jumps. Consequently, this analysis enables us to understand how well tests can disentangle jumps at different sampling frequencies.

Stochastic volatility models

We simulate the stochastic processes for every second.

We sampled the process every 1, 5, 15 and 30 minute(s) and compare the size of the tests.

For the process with one volatility factor (SV1F), we consider three (low, medium, high) alternative values for the mean reversion parameter of the volatility factor.

Size. If we look at all the sampling frequencies, the biggest size distortion is encountered in the case of the JO test, where, for a 1 second sampling frequency, we have a size equal to 0.095, which grows fast when we diminish the sampling frequency.

A similar pattern can be seen for the PZ procedure, which displays a size close to the nominal one when sampling is performed every second, but then gets rapidly and highly oversized. Thus, these tests should be applied only when data sampled as frequent as possible.

The best behavior in terms of size is found for the BNS classic test and for the intraday ABD - LM procedures. In all cases, size does not change very much over the sampling frequency and the size distortion is not very high.

TABLE 5

Table 5: Size of the tests for jumps for the SV1F model with medium mean reversion $\,$

Procedure	Nominal size: 5%						
	1 sec	1 min	$5 \min$	$15 \min$	$30 \min$		
AJ (threshold)	0.047	0.038	0.031	0.027	0.014		
AJ (power var)	0.048	0.046	0.051	0.088	0.150		
BNS	0.048	0.054	0.053	0.057	0.063		
JO	0.095	0.091	0.099	0.119	0.151		
ABD-LM	0.074	0.066	0.074	0.063	0.059		
PZ	0.049	0.065	0.083	0.100	0.121		

Power We added to the continuous stochastic volatility process SV1F jump processes with different intensities and jump sizes.

Varying jump intensity In order to examine how jump detection changes as the number of jumps grows, we consider Poisson jump arrival times depending on the following varying jump intensities (λ): .014, .058, .089, .118, .5, 1, 1.5, 2, and 2.5. For all these scenarios, we consider a jump size that is $N(0, \sigma_{jump}^2 = 1.5\%)$.

TABLE 10

Table 10: Power of daily jump tests for a 5% significance level. We consider the SV1F model with medium mean reversion for the volatility factor and with a varying number of jumps, as a result of varying jump intensities

	Test	1 sec	1 min	5 min	15 min	30 min
	AJ (threshold)	0.971	0.777	0.238	0.035	0.005
	AJ (power var)	0.969	0.775	0.280	0.185	0.236
$\lambda = 0.058$	BNS	0.954	0.819	0.685	0.496	0.366
	JO	0.478	0.445	0.386	0.353	0.321
	ABD-LM	0.989	0.859	0.742	0.612	0.505
	PZ	0.985	0.894	0.773	0.654	0.493
	AJ (threshold)	0.971	0.779	0.208	0.031	0.007
	AJ (power var)	0.970	0.796	0.301	0.187	0.246
$\lambda = 0.118$	BNS	0.954	0.831	0.704	0.544	0.368
	JO	0.503	0.450	0.402	0.368	0.343
	ABD-LM	0.981	0.863	0.762	0.651	0.521
	PZ	0.975	0.895	0.783	0.667	0.513
	AJ (threshold)	0.972	0.803	0.217	0.043	0.006
	AJ (power var)	0.972	0.811	0.323	0.220	0.246
$\lambda = 0.5$	BNS	0.959	0.854	0.729	0.561	0.402
	JO	0.492	0.461	0.421	0.379	0.347
	ABD-LM	0.987	0.888	0.784	0.652	0.523
	PZ	0.982	0.911	0.815	0.695	0.556
	AJ (threshold)	0.982	0.833	0.206	0.040	0.005
	AJ (power var)	0.982	0.852	0.351	0.228	0.255
$\lambda = 1$	BNS	0.970	0.890	0.783	0.609	0.430
	JO	0.502	0.474	0.441	0.391	0.356
	ABD-LM	0.989	0.912	0.828	0.688	0.535
	PZ	0.988	0.932	0.861	0.738	0.570
	AJ (threshold)	0.991	0.853	0.176	0.028	0.004
	AJ (power var)	0.992	0.899	0.410	0.259	0.279
$\lambda = 2$	BNS	0.984	0.933	0.854	0.687	0.488
	JO	0.535	0.516	0.484	0.443	0.397
	ABD-LM	0.996	0.952	0.881	0.734	0.555
	PZ	0.994	0.960	0.911	0.810	0.623

Varying jump size Fixed the number of jumps for the entire sample and vary the jump size, generated by a normal distribution with mean 0 and a standard deviation that ranges between 0 and 2 bs with a growth rate of 0.5.

TABLE 11

Table 11: Power of daily jump tests for a 5% significance level. We consider the SV1F model with medium mean reversion for the volatility factor and with a varying jump variance

		1 sec	1 min	5 min	15 min	30 min
	AJ (threshold)	0.921	0.490	0.101	0.024	0.012
	AJ (power var)	0.921	0.509	0.159	0.123	0.171
$\sigma = 0.5$	BNS	0.872	0.565	0.341	0.176	0.120
	JO	0.469	0.365	0.281	0.205	0.202
	ABD-LM	0.967	0.719	0.491	0.288	0.167
	PZ	0.950	0.725	0.509	0.303	0.200
	AJ (threshold)	0.972	0.713	0.189	0.029	0.006
	AJ (power var)	0.972	0.727	0.265	0.176	0.211
$\sigma = 1$	BNS	0.943	0.779	0.612	0.416	0.267
	JO	0.487	0.419	0.366	0.318	0.278
	ABD-LM	0.987	0.850	0.713	0.528	0.384
	PZ	0.982	0.867	0.733	0.557	0.394
	AJ (threshold)	0.976	0.797	0.214	0.039	0.007
	AJ (power var)	0.976	0.815	0.332	0.216	0.245
$\sigma = 1.5$	BNS	0.962	0.861	0.731	0.565	0.403
	JO	0.507	0.475	0.430	0.385	0.344
	ABD-LM	0.986	0.890	0.794	0.648	0.517
	PZ	0.984	0.914	0.819	0.691	0.534
	AJ (threshold)	0.983	0.847	0.223	0.038	0.004
	AJ (power var)	0.983	0.857	0.376	0.249	0.274
$\sigma = 2$	BNS	0.970	0.890	0.799	0.659	0.503
	JO	0.519	0.489	0.459	0.418	0.385
	ABD-LM	0.991	0.907	0.831	0.715	0.597
	PZ	0.988	0.932	0.862	0.752	0.625
	AJ (threshold)	0.985	0.878	0.233	0.039	0.002
	AJ (power var)	0.984	0.890	0.418	0.267	0.288
$\sigma = 2.5$	BNS	0.976	0.915	0.840	0.721	0.578
	JO	0.506	0.484	0.462	0.430	0.401
	ABD-LM	0.990	0.922	0.861	0.759	0.639
	PZ	0.987	0.945	0.893	0.811	0.690

Overall, the performance of all tests increases with the size of the jumps. The ranking of the tests is in line with what was found for the case of varying jump intensity.

- Very good ability of the the ABD-LM and PZ tests to detect jumps, with powers around 98% and 99% at 1 second, which gradually decreases with the sampling frequency.
- BNS shows power around 97% and 98% at 1 second, which decays when sampling less frequently, but to lower numbers than for the intraday and PZ test.
- AJ does again very well for the highest frequency, with a dramatic decrease in power at 5 and 15 minutes frequencies.

- JO show very modest performance

The behavior of the different tests for jumps in the presence of microstructure noise (see TABLE 12 in the paper)

Extensions to the jump testing procedures

(a)Advantages of approximate finite sample distributions for the ABD and LM tests

The difference between the ABD and LM procedures resides in the choice of the critical values.

Sidak approach for the ABD, which has the advantage of taking into consideration the daily number of observations and the size of the window on which the local volatility estimator that enters the test statistic is computed.

LM test makes use of the asymptotic distribution of the maximum and is characterized by simplicity in comparison with the ABD approach.

Although we expect better finite sample properties for the ABD approach, it requires very smal nominal sizes, not allowing for comparisons with all other tests considered here.

In this section, we propose to use simulated critical values for the maximum of the tests statistics.

This approach enables us to attain comparability with the other tests and to benefit from accounting for the sample size in the inference process

White (2000) suggests the procedure in the context of model selection, where so-called "Monte Carlo Reality Check" defined as a simulation based method for "obtaining a consistent estimate of a p-value for the null in the context of a specification search" (White, 2000, pp 1102) is proposed.

In our case, this "Reality Check" implies the following.

Let M be the number of daily observations and $\hat{\sigma}_j$ the local volatility estimate at time t_j . At each time, t_j , we simulate from $N(0,\hat{\sigma}_j)$ 10,000 paths, where each path includes a number of M observations. For every path, we take the maximum over the M observations, resulting in an approximate finite sample distribution of the maximum. From this empirical distribution, we select the critical values.

We compare the results of the RC with the those based on the asymptotic distribution of the maximum.

Daily size = % of days when tests erroneously identified the occurrence of at least 1 jump.

Overall size = proportion of incorrectly identified jumps in the total number of observations.

Size distortion = overall size minus adequate nominal size.

Figures 6 depicts all the above measures together with the corresponding nominal sizes for different sampling frequencies for the SV1F model with medium mean reversion.

The "Reality Check" approximation of the critical values works very well for a nominal size of 5%.

At a lower 0.01% significance level the "Reality Check" still performs better, except for the case of 30 minutes data, whereas at .1%, we notice that the departures of the "Reality Check" size from the nominal one becomes greater compared to the asymptotic size starting from 15 minutes data.

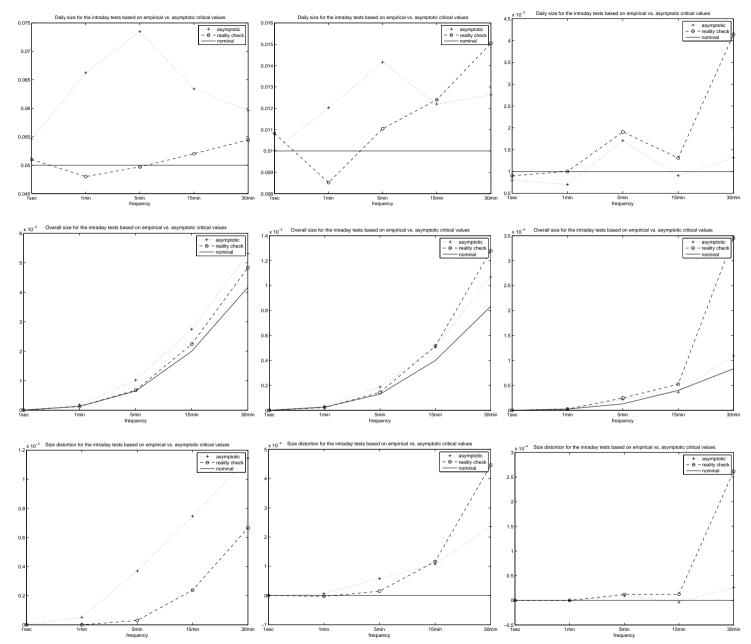


Figure 6: Daily size, overall size and size distortion for the SV1F model and for different significance levels: from left to right: 5%, 1% and .1%

FIGURE 6 (size, SV1F)

The poor performance of the "Reality Check" procedure at low frequencies and for very small significance levels can be explained by considering the combination of two different effects.

First, the number of observations per day, M, is decreasing with the sampling frequency, i.e. 23399 for data sampled every second, 389 at 1 minute, 77 at 5 minutes and only 25 and 12 for 15 and 30 minutes respectively. As the "Reality Check" procedure requires taking maxima over M normal variables, these maxima generally decrease with the frequency, as they are taken over fewer values.

Second, under the null of jumps, returns are assumed to be conditionally normal. When applying the intraday tests, we rely on nonparametric estimators of the local volatility, which require a sampling frequency that

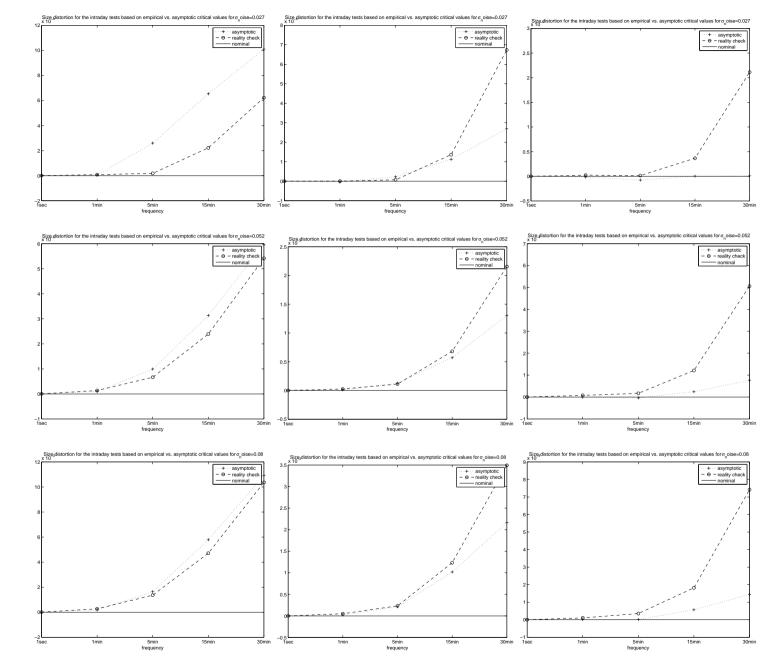


Figure 7: Size distortion for the SV1F model plus microstructure noise for different significance levels: from left to right: 5%, 1% and .1%

becomes higher and higher, i.e. $M \to 1$. For volatile processes and low frequencies, the realized bipower variation seems to underestimate the integrated variance of the price process.

Thus, the standardized intraday returns are no longer i.i.d. normal and consequently, the critical values obtained as a result of the "Reality Check" procedure become too permissive, leading to a higher size for lower frequencies and very small significance levels.

Power Daily power= % of days the procedures were able to correctly signal that at least one jump occurred, overall power= proportion of observations correctly classified as jumps.

We report daily power for the SV1F model with jumps with $\lambda=0.5$ and $\sigma_{jump}=1.5\%$, to which we further add microstructure noise with noise $\sigma_{noise}=0.052$.

FIGURE 8 (power)

The main conclusion of this section is that the "Reality Check" approach can lead to more correctly sized intraday tests with respect to the asymptotic counterpart, accompanied by an increase in power, provided that the number of observations per day is not very low (over 15 minutes).

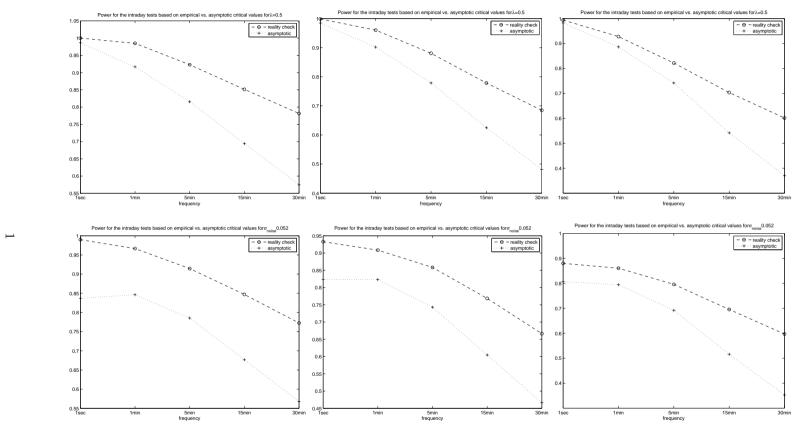


Figure 8: Power for the SV1F + jumps model in the absence/ presence of microstructure noise for different significance levels: from left to right: 5%, 1% and .1%

(b) Cross-performances of the tests

This simple approach is meant to show that combinations of tests and/or sampling frequencies can do better than just applying one single procedure.

It preserves a high percentage of rightly classified jumps, with significant decrease in the percentage of spurious jumps.

Since the BNS test is the most utilized in applied work, we analyze here all combinations of this test with the other four tests. For each case we consider, we compute the percentages of correctly identified days with jumps, correctly identified days without jumps, as well as spuriously detected discontinuities.

TABLE 14

Table 14: Cross-performances of the BNS test coupled with the following tests: LM,PZ, JO and AJ, at a sampling frequency of 15 minutes

	Intersection	BNS-ABDLM	BNS-PZ	BNS-JO	AJ (power var)
	'Jump'	0.4940	0.5323	0.2692	0.1418
	'No Jump'	0.9082	0.0258	0.8507	0.8502
	'Spurious'	0.0069	0.0237	0.0132	0.0064
	Reunion	BNS-ABDLM	BNS-PZ	BNS-JO	AJ (power var)
Ī	'Jump'	0.6856	0.6851	0.6473	0.6134
	'No Jump'	0.9931	0.9289	0.9868	0.9936
	'Spurious'	0.4550	0.4838	0.5125	0.5130

First column in the upper panel of Table 14 (intersection of two jump detection criteria)

0.4940 means that 49.4% of jumps were identified by both the BNS and intraday procedures,

0.9082 indicates that in 90.82% of the days without jumps, both procedures did not identify jumps.

0.0069 shows that there are .69% spurious jumps detected when both procedures are simultaneuosly considered

First column in the lower panel of the table (reunion of two jump detection criteria):

in 68.56% of the days with jumps at least one of the two above procedures identifies jumps,

in 99.31% of the days without jumps at least one of the two tests did not identify jumps,

in 45.5% of the days without jumps at least one of the two tests spuriously identifies jumps.

Combine intersections and reunions across procedures and across frequencies. Table 15 reports the results for some of the combinations

For instance, for the results in column 3, we applied the BNS test on both 5 and 15 minutes simulated data, as well as the ABD-LM procedure based on 15 minutes data.

We adopted the following decision rule: on a certain trading day, the path of the price process is considered discontinuous if one or more jumps is/are detected by the ABD-LM test and at least by one of the two BNS tests.

The results suggest that this procedure manages to average the power over frequencies and/or tests, combined with a substantial decrease in the percentage of spurious jumps.

Third column of Table 15, we observe that the percentage of spuriously detected jumps becomes very low and is combined with a very high proportion (94.98%) of days that were rightly classified as without jumps and a high proportion of correctly identified jumps (approximately 60.70%, defined % averages of the powers of individual tests, i.e. around 54% and 69% for BNS and 64% for ABD-LM as reported in Table 13.

TABLE 15

Table 15: Cross-performance for different combinations of tests

	$\begin{array}{c} (BNS5 \cap BNS1) \cup \\ (BNS5 \cap BNS15) \end{array}$	$ \begin{array}{c} (ABDLM5 \cap ABDLM1) \cup \\ (ABDLM5 \cap ABDLM15) \end{array} $	$\begin{array}{c} (BNS5 \cap ABDLM15) \cup \\ (BNS15 \cap ABDLM15) \end{array}$	$\begin{array}{c} (PZ15 \cap ABDLM15) \cup \\ (ABDLM15 \cap BNS15) \end{array}$
'Jump'	0.660199	0.756965	0.606965	0.597761
'No Jump'	0.954682	0.93796	0.949833	0.932274
'Spurious'	0.001672	0.024582	0.00903	0.014883

Conclusions...

We offer a robust and comprehensive comparison between alternative jump detection procedures based on high frequency data available in the literature

We offer some useful guidelines to potential users on which test/combinations of tests to use to detect jumps in the prices of financial assets

To potential users we recommend to use LM -ABD intraday procedure, as well as the PZ test, which have good power properties combined with a manageable size.

ABD-LM and PZ tests are the most robust to microstructure noise.

When the price processes are very volatile, as it might happen for some assets such as some derivatives, stocks, they become highly oversized. In this case, we recommend the use of the BNS test, as its size distortion is smaller and more stable across frequencies.

We propose as an alternative to the LM asymptotic test critical values calculated according to White (2000)'s "Monte Carlo Reality Check" approach. Provided that the frequency is lower than 15/30 minutes, this approach shows systematic improvements in terms of size and power.

We show that potential users of these procedures can gain advantages by combining them through both reunion and intersection across procedures and across sampling frequencies.

...further developments

There are at least three interesting developments from our work.

The first one concerns with measures of the integrated variance alternative to the Barndorff-Nielsen and Shephard (2004)' realized bipower variation. For instance, Andersen et al. (2009) propose the MinRV and MedRV estimators, whereas Christensen et al. (2009) propose the quantile based realized variance. Andersen et al. (2009) also provide an asymptotic result on the joint limit distribution of the realized variance, the bipower variation, MinRV and MedRV, which enables the testing for jumps.

The second development concerns with the extension of the simulation design to an infinite number of jumps. In this paper, we only considered processes that generate a finite number of jumps within a certain time interval, given that the available tests (the only exceptions being AJ and

PZ) are based on multipower variation-type estimators, which are robust only to a finite number of jumps.

Finally, to reduce the probability of detecting spurious jumps, the combination of tests could be enriched by considering test averaging procedures using Fisher (1925)'s method of combining p-values of different tests.

THANK YOU