Detecting Jumps in the Daily and Intraday Returns of PZU

Daniele Melotti

I choose to work with PZU stock data with 10-minutes frequency. The first thing to be done is the transformation of the dataset into a matrix: this will permit to have the intraday returns displayed on rows, namely all the observations collected within one day will be located on a unique row after the transformation. Such transformation can be performed thanks to the "series_into_matrix.ox" code from the Moodle.

After the transformation is ready, there are three days that only have 0 value of returns. I eliminate these three days (2016-01-15, 2016-03-21, 2016-03-25) in order to avoid possible issues when using periodicity filters.
The jump detection process can start now.

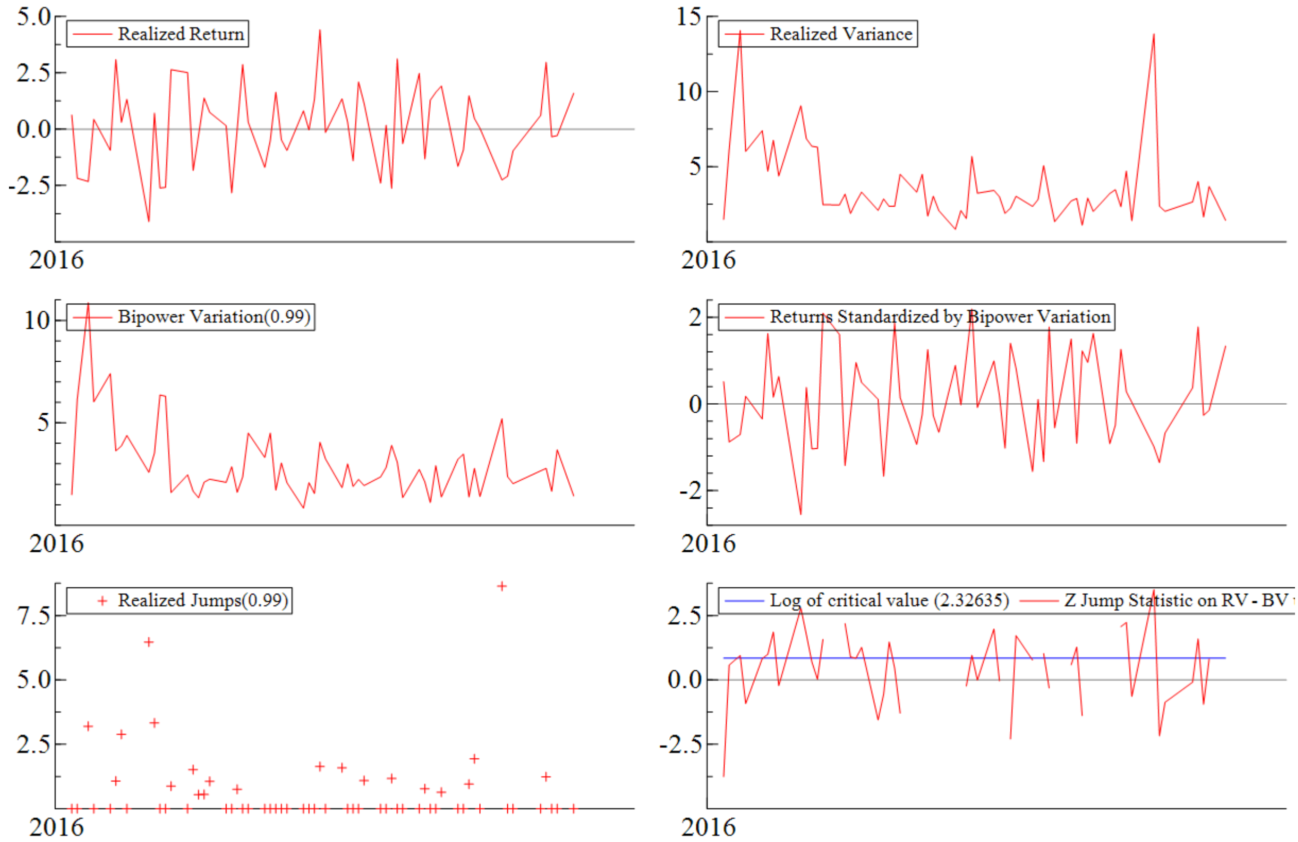
Barndorff-Nielsen and Shepard (BNS) approach

Models from the BNS approach come from the "Realized Volatility using G@RCH" class. We can check for Realized Volatility and Realized Jumps, using the Bipower Variation as an estimator of Integrated Volatility.
Then, we pick Tripower Quarticity as a proxy for Integrated Quarticity, choose Z on RV-BV as a Test Statistics, and set an alpha of 0.99, which would mean having critical level at 1%. Here's the output:

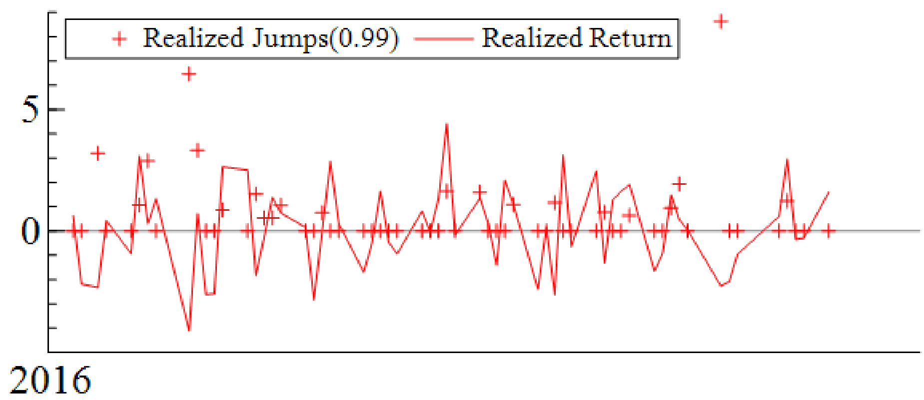
```
Expected number of spurious detected jumps (under H0=no jumps): 0.61
Number of detected jumps: 21
Proportion of detected jumps: 0.344262
Critical level: 0.01
Critical value: 2.32635
```

A lot of information is found in the above lines. First of all, the Expected number of spurious detected jumps is 0.61, and it is obtained by multiplying the number of days within the sample (the model was run for the whole sample, therefore 61 days) by the critical level (1%).
The Number of detected jumps with this method is 21, and the Proportion of detected jumps is 0.344262, which means that in the 34.43% of the days considered (21 out of 61 days) there is a jump. The critical level is what we assumed initially (0.01 = 1%) and the Critical Value equals 2.32635. If our empirical statistics are higher than the critical value, then it will mean that we have the jumps detected.

Next, here is a Graphical analysis:



The Realized Jumps are calculated by subtracting the Bipower Variation from the Realized Variance. In the case that their values are the same, there is no jump, therefore the symbols on the graph lay on the x-axis, while if the value of Realized Variance is greater, then we will observe some jumps, namely all the symbols which lay above the x-axis. If we copy the Realized Returns on the Realized Jumps plot:



We can see more clearly which returns were considered jumps and which not.

We can store the values of Realized Variance, Bipower Variation, Realized Jumps and Z Jump Statistics in the database. This will allow us to see in which precise days the overall volatility was high enough we to be considered as a jump.

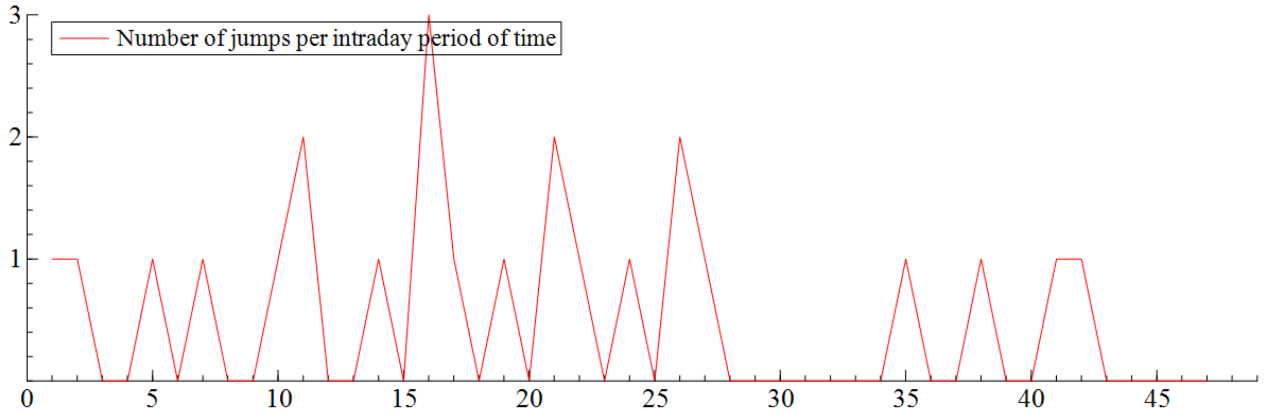
Lee and Mykland approach

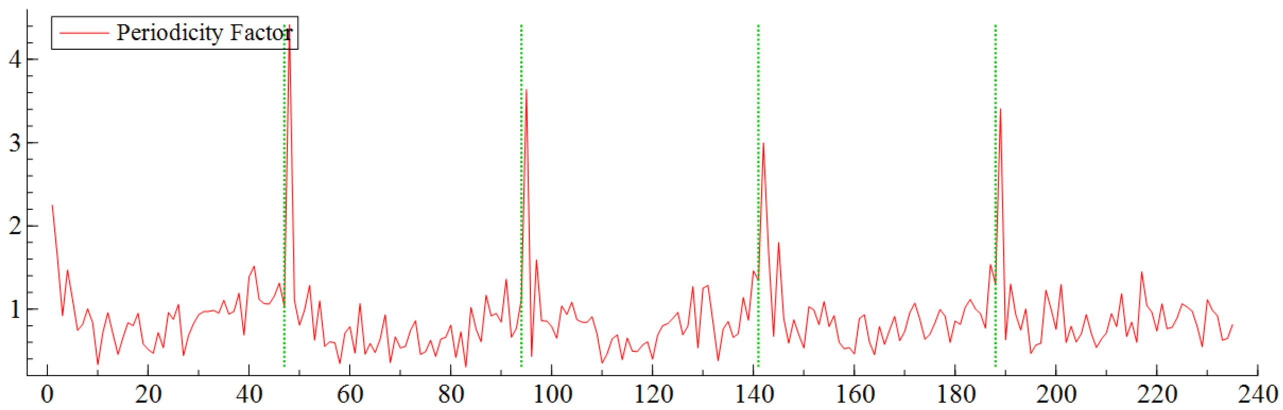
Models from the Lee and Mykland's approach can be formed using the "Lee and Mykland tests for jumps using G@RCH" class.
We can choose the same critical level as in BNS approach, hence 1%. Here we also need to choose a periodicity filter, let's pick WSD. The window size n can be kept as default (intraday observations). Here is the output:

```
Lee and Mykland type of test for jump arrival times.
Local robust variance = Average Bipower Variation
Robust non-parametric periodicity filter: WSD
Critical level of the test: 0.99
n: 2867

Number of detected jumps: 24
Proportion of detected jumps: 0.00837112
Number of periods (typically days) with at least one significant jump: 21
Proportion of periods with at least one significant jump: 0.344262
Critical value, i.e. G(Beta)*Sn+Cn: 4.7397
Expected number of spurious detected jumps (under H0=no jumps): 0.01
```

As we can see, the Number of detected jumps using this method is 24. The Proportion of detected jumps is much smaller, as here we are considering intraday data. In the BNS approach we were considering days (61), here we are considering the total number of observations (2867). The Number of periods with at least one significant jump is 21; this means that there are 21 days within which there is at least 1 jump.
The Proportion of periods with at least one significant jump is 0.344262, indicating that within 34,43% of the sampled days there is at least a jump. The critical value here is 4.7397.
Here is a Graphic analysis:





The first plot shows the 47 observations obtained within a single day (every 10 minutes, from 9:10 to 15:50). And we can see that in the 16th interval, for example, there was a total of 3 jumps. Taking into account that the 16th interval represents 11:40 as a time, it means that within the 61 considered days it happened 3 times that a return observed at 11:40 was considered abnormal, therefore as a jump. And it happened 2 times for the 11th, 21st and 26th intervals, which indicate 10:50, 12:30 and 13:20 respectively.
There are also some intervals when we observe no jumps, or just 1 jump. The Periodicity Factor plot shows 5 days within our sample, so we can see some differences between particular days, for example, we can see a higher value at the beginning of Tuesdays.

We can store the results in the database now. By doing so we will obtain two new files, one containing statistics, and another one containing the jumps. Also, we will get the Lee Mykland Realized Jumps statistics (LM_RJ) at the end of our initial dataset. When such statistic is different from 0, it means that in the related day a jump was detected. Hence, it is now easy and possible to observe how the two approaches of BNS and Lee and Mykland detected jumps.

Here is a table representing the statistics from BNS approach and Lee and Mykland's approach. In green colour are highlighted the values of RJ_BV and Z statistics indicating a jump within the related day for the BNS approach; in light blue colour are highlighted the values of LM_RJ indicating a jump within the related day using the Lee and Mykland approach. Also, dates when jumps were detected using BNS approach are bolded:

	date	*RV*	*BV(0.99)*	*RJ_BV(0.99)*	*RJ-Stat*	*LM_RJ*
1	2016-01-04	1.491829746	1.491829746	0	0.02333707	0
2	2016-01-05	6.153509638	6.153509638	0	1.788568319	3.350494069
3	**2016-01-07**	14.07247153	10.87076613	3.201705398	2.584879031	0
4	2016-01-08	6.023926085	6.023926085	0	0.398506799	0
5	2016-01-11	7.405254308	7.405254308	0	2.292801479	0
6	**2016-01-12**	4.703736369	3.629801561	1.073934807	2.708859484	0.644759387
7	**2016-01-13**	6.762253367	3.875499129	2.886754238	6.448558359	0
8	2016-01-14	4.383753941	4.383753941	0	0.802967173	0
9	**2016-01-18**	9.05548256	2.581603833	6.473878727	16.29309329	7.155258656
10	**2016-01-19**	6.859978572	3.529687667	3.330290905	5.824739313	0
11	2016-01-20	6.362917091	6.362917091	0	2.021781327	1.155494317
12	2016-01-21	6.302539731	6.302539731	0	1.02605285	2.751764573
13	**2016-01-22**	2.471595212	1.593975279	0.877619934	4.840029467	0.297172417
14	2016-01-25	2.461956798	2.461956798	0	-0.862755987	0
15	**2016-01-26**	3.179041802	1.662980011	1.516061791	8.966242868	0.166964468
16	**2016-01-27**	1.887749302	1.341614213	0.546135089	2.435860646	0
17	**2016-01-28**	2.648905883	2.098071012	0.550834871	2.328302268	0.473251357
18	**2016-01-29**	3.313580396	2.253184251	1.060396146	3.554059365	0
19	2016-02-01	2.09746571	2.09746571	0	0.211968792	0
20	2016-02-02	2.860433918	2.860433918	0	0.574075764	0.861309748
21	**2016-02-03**	2.359038814	1.606603961	0.752434853	4.377894866	0
22	2016-02-04	2.361273312	2.361273312	0	1.582216231	0
23	2016-02-05	4.505170902	4.505170902	0	0.274859533	0
24	2016-02-08	3.309388986	3.309388986	0	-0.321404755	0
25	2016-02-09	4.503608355	4.503608355	0	-0.195846573	0
26	2016-02-10	1.712553126	1.712553126	0	1.32745874	0.174499138

27	2016-02-11	3.040927602	3.040927602	0	-0.135237473	0
28	2016-02-12	2.077144538	2.077144538	0	0.525005777	0
29	2016-02-15	0.836965478	0.836965478	0	-0.534451179	0
30	2016-02-16	2.080967327	2.080967327	0	-0.000218135	0
31	2016-02-17	1.550230795	1.550230795	0	0.789309406	0
32	**2016-02-18**	5.691215292	4.054289563	1.636925729	2.612440207	0
33	2016-02-19	3.240927167	3.240927167	0	0.992870092	1.02457851
34	**2016-02-22**	3.428531472	1.840577484	1.587953988	7.257749504	0
35	2016-02-23	3.000677101	3.000677101	0	0.965324396	0
36	2016-02-24	1.894821192	1.894821192	0	-1.855056084	0
37	2016-02-25	2.240940216	2.240940216	0	0.101125964	0
38	**2016-02-26**	3.033238471	1.939945516	1.093292955	5.609117892	0
39	2016-02-29	2.360569723	2.360569723	0	2.171644778	0
40	2016-03-01	2.821941965	2.821941965	0	-0.587059835	0
41	**2016-03-02**	5.077121069	3.902465671	1.174655398	2.78491984	0.279291517
42	2016-03-03	3.083106208	3.083106208	0	0.731097675	0
43	2016-03-04	1.348663774	1.348663774	0	-0.072447917	0.336065734
44	2016-03-07	2.726316656	2.726316656	0	1.80650751	0.330296441
45	**2016-03-08**	2.885660223	2.112040073	0.773620149	3.58963592	0
46	2016-03-09	1.109155333	1.109155333	0	0.249859434	0
47	2016-03-10	2.912577666	2.912577666	0	-0.434523244	0
48	**2016-03-11**	2.024398273	1.383229027	0.641169246	4.474500414	0.151319004
49	2016-03-14	3.22333542	3.22333542	0	-0.976104254	0
50	2016-03-15	3.47300916	3.47300916	0	-0.798863869	0.254362893
51	**2016-03-16**	2.337602213	1.383509179	0.954093034	7.948976405	0.606222251
52	**2016-03-17**	4.713718331	2.777868586	1.935849745	9.350744853	0
53	2016-03-18	1.40459766	1.40459766	0	0.525921196	0
54	**2016-03-22**	13.84891551	5.206720459	8.642195051	33.33906841	0
55	2016-03-23	2.375940746	2.375940746	0	0.113796451	0.574008564
56	2016-03-24	2.034552612	2.034552612	0	0.421037623	0.762327729
57	2016-03-29	2.658287923	2.658287923	0	0.917224322	0
58	**2016-03-30**	4.01710578	2.778314911	1.238790869	4.936487272	0.290447649
59	2016-03-31	1.654626253	1.654626253	0	0.388742822	0.2041689
60	2016-04-01	3.685943052	3.685943052	0	2.253495853	0
61	2016-04-04	1.432705048	1.432705048	0	-0.179054036	0

It is quite interesting to notice how often a jump that is detected using BNS approach is not confirmed by the Lee and Mykland’s approach and vice versa.

If we open the file with statistics, we will see that it contains the statistics for each return. So, we have the J value calculated for each particular day, calculated by dividing absolute returns by the Bipower Variation, which was our Integrated Volatility estimator.

And if we open the file with jumps, we can check at which exact time within a particular day a jump was detected.
As we had seen in the text output, we have at least one jump in 21 days, but a total of 24 jumps. If we go to see the cells where we have these jumps, but in our original starting dataset, we will see that the values contained in those cells are all higher than the critical value from the text output, which is 4.7397 in this case. Therefore, when in the jumps’ file there is a zero, it means that in the corresponding cell of the original dataset the value is lower than the critical one, therefore there is

no jump. If we consider columns, we can get the total number of jumps observed within each interval.
Using WSD as a periodicity filter, we can confirm everything what the plots had shown us, for example the fact that at the time 11:40 there are 3 jumps in total: this happens on the days 2016-01-22, 2016-02-10 and 2016-03-02.

Here is a small table representing the values of the 5 largest (positive and negative) jumps found using Lee and Mykland's approach:

Date	*Time*	*Jump*
2016-01-12	10:40	0.802969107
2016-01-28	15:20	0.687932669
2016-02-02	13:30	0.635932191
2016-01-22	11:40	0.545135228
2016-03-15	12:30	0.504344022
2016-01-18	9:10	-2.674931524
2016-01-05	12:30	-1.830435486
2016-01-21	15:50	-1.658844348
2016-01-20	16:00	-1.074939216
2016-02-19	13:00	-1.012214656

In green are the five most positive jumps, while in red there are the five most negative ones.