RARE EVENTS ANALYSIS OF HIGH-FREQUENCY FINANCIAL DATA

Dragos Bozdog[1], Ionut Florescu[1], Khaldoun Khashanah[1,2], Jim Wang[2]

[1]*Department of Mathematical Sciences, Stevens Institute of Technology, USA*

[2]*Program in Financial Engineering, Stevens Institute of Technology, USA*

ABSTRACT: In this work we present a methodology to detect unusual trading activity defined as high price movement with relatively little volume traded. The study is performed using all the trade data for some five thousand equities for five days. We analyze what happens after this unusual activity is detected and we find that in the majority of the cases the price of equities tends to bounce back. The methodology developed is based on non-parametric statistics and makes no assumption about the distributions involved.

SPECIFIC OBJECTIVES OF THE STUDY

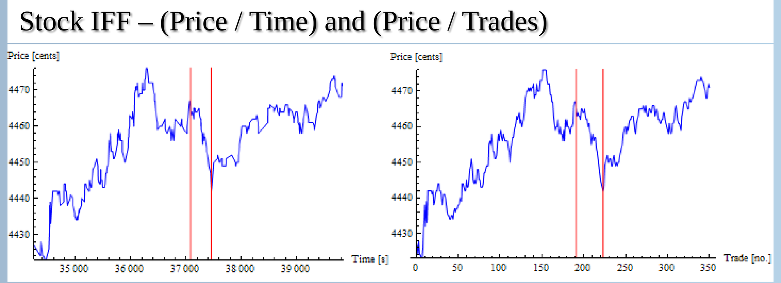
The main objectives this study are:

1. Develop a method to detect *large price movements* corresponding to *small volume of shares traded*.
2. Analyze the evolution of price *after* these unusual events and study the probability of price recovery.
3. Estimate the *expected return* if a trade is placed at the detected event.
4. Compare the result obtained with "normal" price evolution simulated using Monte Carlo technique.

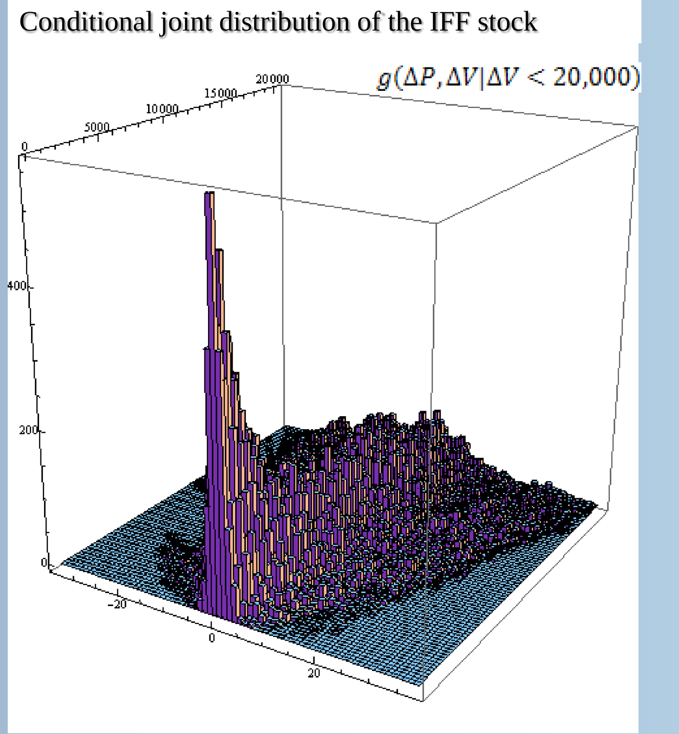
METHODOLOGY

In this analysis we use tick-by-tick data of 5,369 equities (TAQ).

Stock IFF – (Price / Time) and (Price / Trades)



Conditional joint distribution of the IFF stock



First we note that the trades are recorded in consecutive order and, though the time between trades is not necessarily constant, the price versus time and the price versus trade distributions are very similar. As a way to visualize this similarity, we present a small segment of the stock IFF from Apr. 14 2008 (the two images on the left).

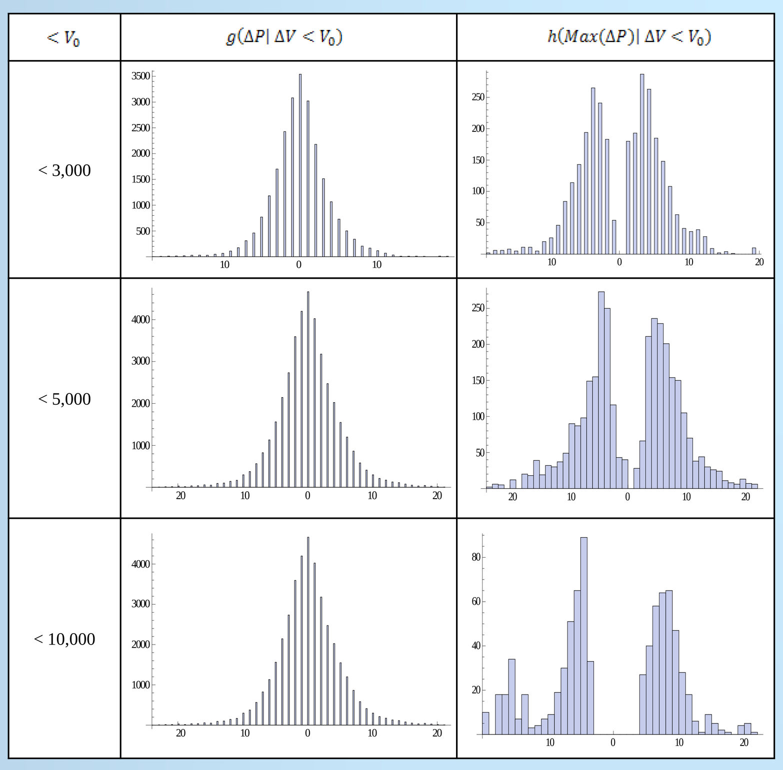
The 3D image on left represents the joint distribution of volume and price movement. The total number of pairs used for this distribution is 159,583. This is an extremely large number of data points to be calculated and analyzed for every stock and for every day.

To further simplify we retain only the maximum price movement with respect to change in volume for each window with $\Delta V < V_0$. Consequently, all the analysis that follows is based on the following distribution:

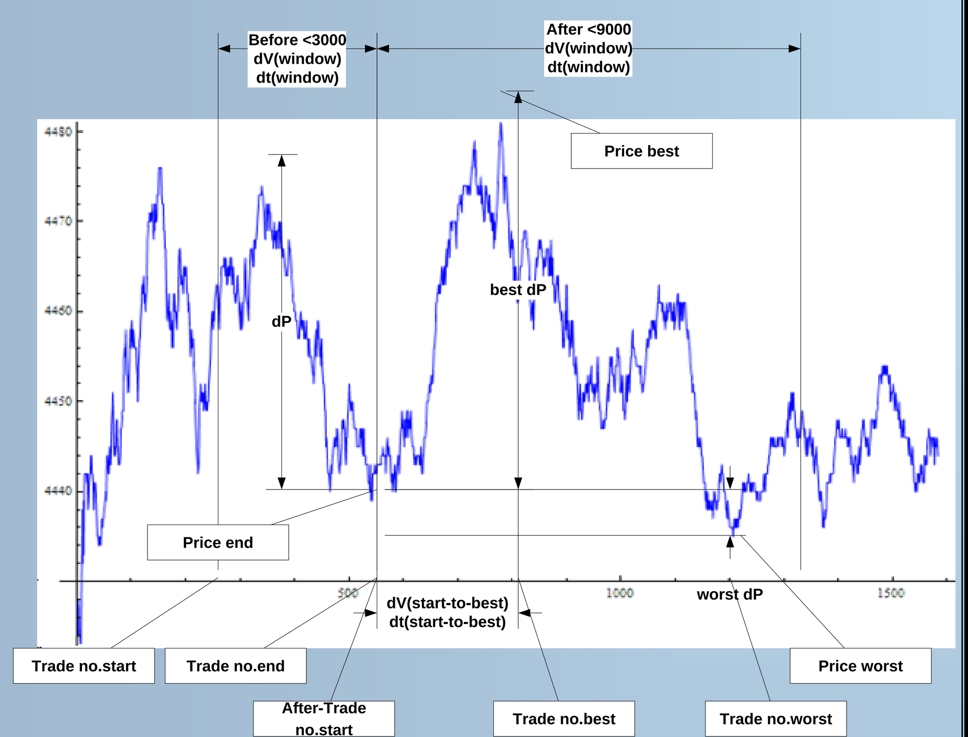
$$h(Max(\Delta P) | \Delta V < V_0)$$

For example for the IFF stock, we obtain 1,570 observations for $V_0 = 3{,}000$ shares, 1,562 observations for $V_0 = 5{,}000$ shares and 1,544 observations for $V_0 = 10{,}000$ shares.

The proposed sampling technique generates new distributions that share similar behavior of the tails and provides a computationally feasible approach to this analysis.



RARE EVENTS DETECTION AND ANALYSIS

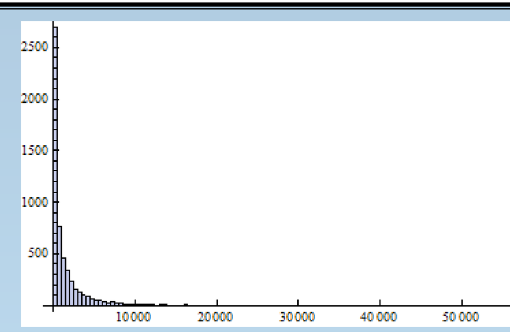


Consider the distribution $h(Max(\Delta P) | \Delta V < V_0)$ with $V_0 = 5{,}000$ for stock IFF on Apr. 14, 2008. The sampling distribution is performed by using a moving window with cumulative volume smaller than 5,000 shares. The observation that has the maximum price change within the window is recorded. A typical output of such analysis has the following form:

ΔP	ΔV	Start	End	PriceEnd	dt	r
-24	4900	189	222	4443	389	0.537273
-23	4900	191	223	4442	374	0.515118
-23	4900	188	221	4444	397	0.514887
-22	4900	192	224	4443	369	0.492721
-22	4800	188	220	4445	393	0.492501
21	4900	2	27	4444	453	0.474791
21	4900	1	26	4444	541	0.474791
21	4900	108	136	4471	261	0.47191
21	4900	104	134	4471	272	0.47191
20	4900	636	664	4462	537	0.450248
20	4800	634	663	4462	540	0.450248
20	4800	111	137	4470	257	0.449438
20	4900	107	135	4470	265	0.449438
20	4900	103	133	4470	277	0.449438
20	4900	102	132	4470	289	0.449438

COMPARATIVE STUDY OF STOCK GROUPS

Herein, we analyze the change in price from the volume perspective; therefore, we recognize the need for classifying stocks into classes based on the average daily traded volume. We refer to this classification as the *multi-scale volume classification*. The histogram corresponds to the average daily trading volume of the total universe of 5,369 stocks considered in this study.



Group	Average daily volume (shares)	Number equities
1	$ADV \le 30{,}000$	1,305
2	$30{,}000 < ADV \le 100{,}000$	1,088
3	$100{,}000 < ADV \le 1{,}000{,}000$	2,117
4	$1{,}000{,}000 < ADV \le 10{,}000{,}000$	799
5	$10{,}000{,}000 < ADV$	60

For a consistent approach to the rare events detection, we calculate the quantiles for all equities individually over a 5 day period. The tables present an example of quantile calculations for a few exemplifying stocks. Based on the previous classification JPM belongs to class 1, GS to class 2, IFF to class 3, and STAN to class 4.

Symbol	0.01	0.99	0.0075	0.9925	0.005	0.995	0.0025	0.9975	0.001	0.999	0.00075	0.99925	0.0005	0.9995	0.00025	0.99975	0.0001	0.9999
JPM	-4	4	-4	4	-5	5	-6	6	-7	7	-8	8	-9	9	-11	10	-14	12
GS	-22	22	-23	23	-25	24	-27	27	-31	32	-33	34	-34	36	-37	38	-40	42
IFF	-14	13	-15	14	-16	16	-19	18	-21	20	-21	20	-22	20	-23	23	-24	23
STAN	-42	44	-43	45	-51	46	-83	48	-83	51	-87	78	-87	78	-87	78	-87	78

Symbol	0.01	0.99	0.0075	0.9925	0.005	0.995	0.0025	0.9975	0.001	0.999	0.00075	0.99925	0.0005	0.9995	0.00025	0.99975	0.0001	0.9999
JPM	-5	5	-6	6	-6	6	-7	7	-9	9	-10	9	-11	10	-14	12	-20	14
GS	-28	27	-29	29	-31	31	-35	35	-39	39	-40	40	-42	42	-47	45	-52	47
IFF	-17	17	-19	18	-21	20	-23	21	-27	24	-29	24	-30	25	-31	27	-31	27
STAN	-56	49	-60	54	-63	59	-78	83	-78	85	-83	89	-83	89	-83	89	-83	89

Symbol	0.01	0.99	0.0075	0.9925	0.005	0.995	0.0025	0.9975	0.001	0.999	0.00075	0.99925	0.0005	0.9995	0.00025	0.99975	0.0001	0.9999
JPM	-7	8	-8	8	-9	9	-10	10	-12	12	-13	13	-13	14	-17	15	-23	18
GS	-38	38	-40	40	-43	42	-47	46	-52	49	-54	50	-57	52	-60	54	-64	57
IFF	-24	24	-25	26	-27	28	-30	31	-31	34	-31	35	-32	36	-33	38	-33	38
STAN	-68	57	-68	58	-69	60	-69	67	-71	89	-71	89	-71	89	-71	89	-71	89

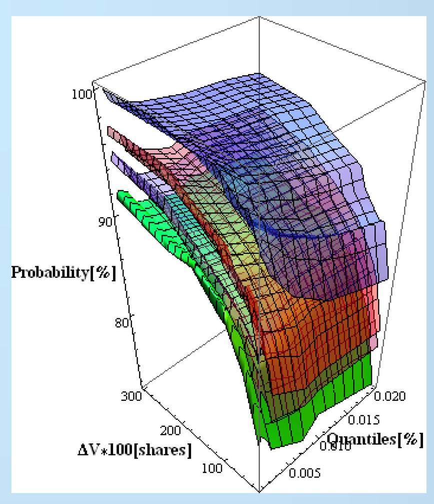
Definition 1: *We say that a favorable price movement occurred if either the price level within the after-event window raises above the event price for at least one trade if the event was generated by a negative quantile, or the price level within the after-event window decreases below the event price level for at least one trade if the event was generated by a positive quantile.*

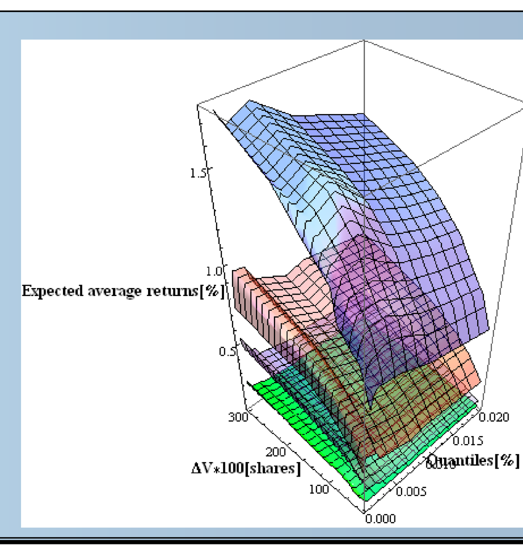
Probability of favorable price movement

Ave Daily Vol.	Quantile	3000	6000	9000	5000	10000	15000	10000	20000	30000
< 100,000 Small stocks	0.02	84.1287	88.9725	91.0485	89.0041	92.6829	94.056	93.7998	95.727	96.3583
	0.015	85.2982	89.9954	91.7798	90.1654	93.4329	94.6612	94.6836	96.462	97.0049
	0.01	86.6804	91.2253	92.6554	91.8279	94.6177	95.6836	95.0239	96.832	97.2953
	0.005	89.5197	93.4789	94.5852	94.0678	96.0766	97.1438	96.8093	97.7477	98.048
	0.002	92.6829	95.5865	96.6318	96.3255	98.1627	98.8189	98.2175	98.574	98.7522
	0.0015	93.7238	96.0251	96.8619	95.5294	97.6471	98.8235	98.7692	99.0769	99.0769
	0.001	94.5205	97.2603	98.6301	97.4576	99.1525	99.1525	98.6301	100.	100.
	0.0005	na	na	na	na	na	na	na	na	na
	0.0002	na	na	na	na	na	na	na	na	na
<1,000,000 Medium Stocks	0.02	78.4828	84.8194	87.5394	83.3873	88.2792	90.349	88.8366	92.1485	93.3952
	0.0075	78.852	85.0869	87.7105	83.6993	88.5493	90.5445	89.2813	92.4249	93.6237
	0.01	79.3456	85.433	88.0341	84.5752	89.2344	91.1732	90.0452	93.0639	94.1157
	0.0025	81.2395	86.9481	89.2847	86.3678	90.4421	92.3146	91.7597	94.2695	95.0263
	0.002	84.6479	89.3192	91.1972	89.6975	92.8197	94.3401	94.0815	96.1137	96.5557
	0.00075	85.8151	90.4196	92.1203	90.9554	93.5798	94.9642	94.7773	96.4453	96.8097
	0.001	86.9838	91.2461	92.7312	91.7077	94.0715	95.3609	95.563	97.0686	97.3751
	0.00025	88.914	92.7812	93.8769	93.2125	94.9259	96.112	96.4557	97.6492	97.8662
	0.0001	88.8655	92.2269	93.4874	94.2792	95.6522	97.254	97.5845	98.0676	98.0676
< 10,000,000 Large Stocks	0.02	76.5361	83.1437	86.1431	80.5519	86.1905	88.7339	85.4713	89.8481	91.7945
	0.0075	76.8222	83.356	86.2902	80.9934	86.4852	88.99	85.8024	90.1185	92.0426
	0.01	77.29	83.7233	86.5757	81.4613	86.7742	89.2262	86.2093	90.4458	92.367
	0.0025	78.3131	84.4568	87.0637	82.4018	87.4875	89.7823	86.9878	91.0876	92.895
	0.002	80.5029	85.9844	88.3024	84.0494	88.7074	90.7643	88.7757	92.6605	94.0936
	0.00075	81.4689	86.7155	88.8686	84.9381	89.3504	91.2462	89.6924	93.5121	94.7174
	0.001	82.6894	87.7362	89.8215	86.1956	90.3234	92.1687	91.2626	94.519	95.4282
	0.00025	85.4161	89.6189	91.5803	89.2837	92.5458	94.0866	93.0452	95.4923	96.1792
	0.0001	88.2302	92.0103	93.6426	92.6695	95.1695	96.2712	95.1715	96.6606	96.9314
Greater than 10,000,000 (Major indexes and super large cap stocks)	0.02	71.7507	79.7573	83.5179	77.3566	83.9935	87.0469	81.4939	86.9276	89.2139
	0.0075	72.3597	80.4332	84.0911	77.4632	84.0274	87.1246	81.832	87.2281	89.5866
	0.01	74.0957	81.8966	85.2831	77.9987	84.5674	87.6781	83.0278	88.035	90.2904
	0.0025	74.8716	82.7306	86.0694	78.7202	85.2405	88.1177	83.7288	88.3179	90.6386
	0.002	76.2729	83.2455	86.757	80.531	86.7735	89.5001	86.0767	90.1563	91.78
	0.00075	76.4362	83.2473	86.8593	80.9623	87.229	90.0505	86.2069	90.0383	91.6888
	0.001	77.5913	84.5015	88.154	82.5988	88.3235	90.5952	86.9584	90.7712	92.2877
	0.00025	79.3965	86.0562	88.7617	84.3602	89.5735	91.3744	87.2174	90.4348	92.087
	0.0001	81.5934	87.9121	90.1099	84.9741	89.6373	91.9689	91.4141	93.4343	94.9495

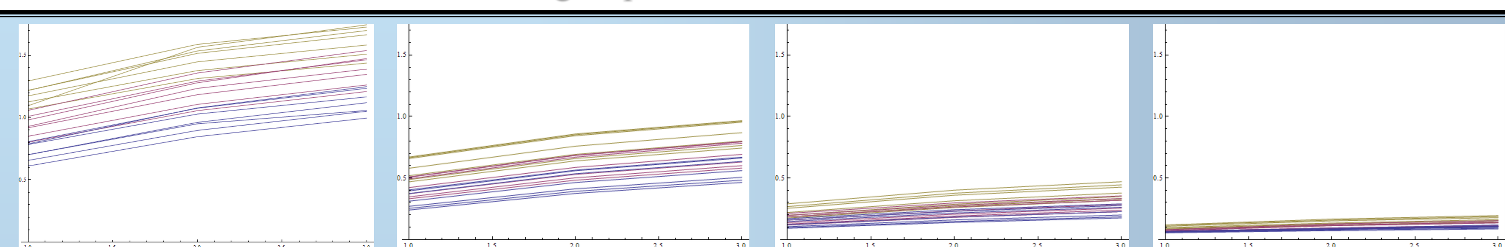
Expected return (%)

Ave Daily Vol.	Quantile	3000	6000	9000	5000	10000	15000	10000	20000	30000
< 100,000 Small stocks	0.02	0.611921	0.847285	0.996335	0.802997	1.05622	1.20965	1.06955	1.31429	1.43802
	0.015	0.656979	0.897619	1.05133	0.850725	1.10645	1.26259	1.12679	1.37813	1.50934
	0.01	0.702632	0.96203	1.11986	0.918936	1.18467	1.34749	1.17505	1.44852	1.58317
	0.005	0.789999	1.07835	1.24904	1.05854	1.35847	1.539	1.29473	1.58841	1.72662
	0.002	0.80718	1.07549	1.23589	1.01076	1.2934	1.46488	1.21951	1.51646	1.66556
	0.0015	0.784388	1.02919	1.16697	0.981992	1.28018	1.47456	1.21839	1.53559	1.69974
	0.001	0.703042	0.950619	1.05786	0.929868	1.23523	1.38982	1.09631	1.56492	1.74533
	0.0005	na	na	na	na	na	na	na	na	na
	0.0002	na	na	na	na	na	na	na	na	na
<1,000,000 Medium Stocks	0.02	0.239613	0.374475	0.464317	0.330938	0.482059	0.580856	0.468476	0.639295	0.743711
	0.0075	0.252937	0.39158	0.481936	0.346658	0.501376	0.600762	0.487116	0.65867	0.764661
	0.01	0.269312	0.411849	0.50513	0.371941	0.529415	0.630178	0.518186	0.69221	0.800753
	0.0025	0.311149	0.463261	0.560327	0.421791	0.586057	0.692732	0.57935	0.758693	0.868612
	0.002	0.373187	0.534782	0.635005	0.493077	0.669666	0.783159	0.658094	0.844148	0.954149
	0.00075	0.394853	0.559079	0.662608	0.506602	0.683459	0.794365	0.665208	0.851786	0.962188
	0.001	0.403372	0.563945	0.669373	0.50845	0.685649	0.795949	0.670629	0.858216	0.966528
	0.00025	0.382946	0.534452	0.636835	0.471952	0.639154	0.745375	0.61603	0.796165	0.895118
	0.0001	0.298788	0.423029	0.511236	0.405315	0.554774	0.649517	0.523982	0.673348	0.750551
< 10,000,000 Large Stocks	0.02	0.0906299	0.138484	0.175016	0.117096	0.179483	0.226264	0.174201	0.259594	0.319075
	0.0075	0.0952728	0.146138	0.183984	0.124835	0.190401	0.238656	0.182996	0.27066	0.331204
	0.01	0.103869	0.159264	0.19915	0.136226	0.205562	0.255211	0.19627	0.286725	0.348983
	0.0025	0.119761	0.182413	0.226254	0.156987	0.231691	0.284987	0.219136	0.314432	0.377666
	0.002	0.145761	0.215078	0.261853	0.184306	0.264321	0.318991	0.253299	0.359764	0.425617
	0.00075	0.157837	0.229132	0.27606	0.195605	0.278408	0.333475	0.267214	0.376421	0.444189
	0.001	0.169361	0.241843	0.28938	0.211685	0.29706	0.354674	0.289545	0.401081	0.469584
	0.00025	0.197835	0.277052	0.327418	0.247397	0.336652	0.396583	0.323411	0.43866	0.509101
	0.0001	0.228865	0.31888	0.377122	0.274156	0.376071	0.438603	0.359477	0.48209	0.558128
Greater than 10,000,000 (Major indices and super large cap stocks)	0.02	0.0542638	0.0721463	0.0859378	0.0666065	0.0898908	0.108068	0.0819398	0.112971	0.137848
	0.0075	0.0565237	0.0761864	0.0910154	0.0658548	0.0897141	0.107224	0.0839185	0.117814	0.143116
	0.01	0.0606778	0.0827955	0.0983685	0.0645788	0.0902117	0.108276	0.0896684	0.127246	0.153922
	0.0025	0.060112	0.0832626	0.101406	0.0636383	0.0942417	0.114349	0.096469	0.13342	0.161706
	0.002	0.0596466	0.0829418	0.10535	0.0707946	0.111681	0.134294	0.107295	0.15084	0.177329
	0.00075	0.0615032	0.0850912	0.109235	0.0727399	0.116729	0.140565	0.114852	0.157772	0.186487
	0.001	0.0658508	0.0911557	0.117165	0.079435	0.126301	0.152673	0.11749	0.165267	0.194223
	0.00025	0.0768166	0.111109	0.138084	0.0810528	0.128487	0.155467	0.13192	0.180729	0.2153
	0.0001	0.0881374	0.118497	0.150216	0.0876858	0.143562	0.167868	0.148991	0.196382	0.247034





We can see that unlike the probability plots these surfaces seem to have different curvatures. Each surface has a maximum level which is remarkably obtained using the same quantile level for all volume window sizes. The optimum quantile level is different within each surface. The 3D plots are providing details of the optimal after-event window size for stock groups considered.



At the top of the table to the right we present the realized probabilities of price rebound for JPM for various quantiles and window sizes. The bottom line estimates these probabilities using Monte Carlo simulations.

Quantile	Day 1 <5000	Day 1 <10000	Day 1 <15000	Day 2 <5000	Day 2 <10000	Day 2 <15000	Day 3 <5000	Day 3 <10000	Day 3 <15000	Day 4 <5000	Day 4 <10000	Day 4 <15000	Day 5 <5000	Day 5 <10000	Day 5 <15000
0.02	73.01	79.65	82.30	76.47	82.51	86.17	76.52	83.98	87.57	74.36	82.05	82.05	67.32	78.99	84.05
0.015	72.28	79.21	83.17	77.89	83.51	88.07	73.45	82.74	87.17	68.25	77.78	77.78	66.20	78.87	84.51
0.01	72.28	79.21	83.17	77.89	83.51	88.07	73.45	82.74	87.17	68.25	77.78	77.78	66.20	78.87	84.51
0.005	71.05	81.58	86.84	72.79	81.62	87.50	72.86	85.71	90.00	73.17	85.37	85.37	67.11	78.95	86.84
0.002	66.67	100.00	100.00	66.67	90.91	96.97	73.53	83.82	89.71	84.62	100.00	100.00	74.07	74.07	74.07
0.0015	0.00	100.00	100.00	66.67	88.89	96.30	73.85	83.08	89.23	66.67	100.00	100.00	66.67	66.67	66.67
0.001	0.00	100.00	100.00	76.92	100.00	100.00	78.85	84.62	90.38	100.00	100.00	100.00	66.67	66.67	66.67
0.0005	-	-	-	100.00	100.00	100.00	90.63	90.63	93.75	-	-	-	71.43	71.43	71.43
0.0002	-	-	-	-	-	-	100.00	100.00	100.00	-	-	-	100.00	100.00	100.00
Monte Carlo results	61.396	74.824	81.002	67.001	78.121	83.252	58.17	72.883	80.119	60.929	74.729	81.176	63.957	77.402	83.217

CONCLUSIONS

• As expected stocks behave differently depending on the daily average volume of trades. In particular, from the data under consideration we saw the expected return attains its maximum for the following quantile levels:

Class	Quantile
< 100,000 (ADV)	0.0025
< 1,000,000 (ADV)	0.0005
< 10,000,000 (ADV)	0.0001
> 10,000,000 (ADV)	less than 0.0001

• The estimation of optimal window size produced the following results (with the data under study).

Class	Before event window size	After event window size
< 100,000 (ADV)	5,000	15,000
< 1,000,000 (ADV)	3,000	9,000
< 10,000,000 (ADV)	3,000	6,000
> 10,000,000 (ADV)	10,000	10,000

• In more general terms we determined:

1. We observed the existence of information embedded in the stock movement
2. The events we discover exhibit increased probability of price recovery
3. The method is possibly more appropriate as a measuring tool for market reaction to singular events rather than a trading tool. These singular events may be viewed as suspicious or events that go against the information available to all market participants and in this light it could be developed further for forensic analysis.

REFERENCES

Alfonsi, A., A. Schied and A. Schultz (2007). Optimal execution strategies in limit order books with general shape functions. http://www.citebase.org/abstract?idoai:arXiv.org:0708.1756.

Beaver W.H. (1968). The information content of annual earnings announcements. Empirical research in Accounting: Selected Studies; suppl to Journal of Accounting Research, vol. 6, pp 67-92

Bollerslev, T. and Jubinski, D. (1999). Equity trading volume and volatility: latent information arrivals and common long-run dependencies. Journal of Business & Economic Statistics 17, pp. 9–21.

Gallant, A.R., Rossi, P.E. and Tauchen, G.E, (1992), Stock prices and volume. The Review of Financial Studies 5, pp. 199–242.

Karpoff, J. (1987). The relation between price change and trading volume: A survey, Journal of Financial and Quantitative Analysis, 22, March, pp 109-126.

Lo, A. W., H. Mamaysky, and J. Wang (2000). Foundation of technical analysis: Computational algorithms, statistical inference, and empirical implementation. The Journal of Finance 55 (4), pp 1705-1765.

Lo, A.W, and J. Wang. (2002) Trading volume: Implications of an intertemporal capital asset price model. Advances in Economic Theory: Eighth World Congress, pp 1-23

Osborne M.F.M. (1959) Brownian motion in the stock market, Operations Research, 7(2), pp 145-173.

Sun W. (2003) Relationship between Trading Volume and Security Prices and Returns, MIT LIDS Technical Report 2638, February 2003 Area Exam.

Tsay, A. S. and C. Ting (2006, January). Intraday stock prices, volume, and duration: a nonparametric conditional density analysis. Empirical Economics 30 (4), pp 253-268.

Tsay, R. (2005). Analysis of Financial Time Series. Wiley-Interscience.

Zhang, M. Y., Russell, J. R., & Tsay, R. S. (2008). Determinants of bid and ask quotes and implications for the cost of trading. Journal of Empirical Finance. 15 (4), pp 656-678.