Economics Bulletin

Volume 31, Issue 2

Algorithmic complexity theory detects decreases in the relative efficiency of stock markets in the aftermath of the 2008 financial crisis

Cleiton Taufemback
Federal University of Santa Catarina

Ricardo Giglio
Kiel University

Sergio Da Silva
Federal University of Santa Catarina

Abstract

The relative efficiency of financial markets can be evaluated using algorithmic complexity theory. Using this approach we detect decreases in efficiency rates of the major stocks listed on the Sao Paulo Stock Exchange in the aftermath of the 2008 financial crisis.

Citation: Cleiton Taufemback and Ricardo Giglio and Sergio Da Silva, (2011) "Algorithmic complexity theory detects decreases in the relative efficiency of stock markets in the aftermath of the 2008 financial crisis", *Economics Bulletin*, Vol. 31 no.2 pp. 1631-1647.
Submitted: May 17 2011. **Published:** June 05, 2011.

1. Introduction

A market is efficient if its price conveys nonredundant information (Mantegna and Stanley 2000). Econometric tests are usually employed to assess whether a market is efficient or not (Beechey *et al.* 2001). Market efficiency is thus considered in absolute terms (Campbell *et al.* 1997). Unlike economists, physicists are interested in the *relative* efficiency of a system. An efficiency rate refers, for example, to the relative proportion of energy converted to work. Here, algorithm complexity theory can be used to provide a relative efficiency interpretation for markets (Mantegna and Stanley 2000).

Algorithmic complexity theory (Kolmogorov 1965; Chaitin 1966) tells us that the price series of an idealized efficient market shows statistical features that are indistinguishable from those observed in genuinely random time series (Mantegna and Stanley 2000). As a result, measuring the deviation from randomness yields the relative efficiency of an actual market.

This deviation can be measured by the Lempel-Ziv algorithm (Lempel and Ziv 1976; Kaspar and Schuster 1987). Our previous work (Giglio *et al.* 2008a; 2008b; Giglio and Da Silva 2009) shows how this can be accomplished. Here, we apply the technique to the unique data provided by the 2008 financial crisis. Our finding suggests that stock markets had their efficiency rates reduced after the crisis.

Ranking financial assets in terms of relative efficiency using algorithm complexity theory is to be viewed as offering an alternative method of dealing with the hierarchy of related complex systems. Indeed, there are also other ways of doing such rankings (see Mantegna 1999; Cajueiro and Tabak 2004; Xu *et al.* 2005; Gligor and Ausloos 2008; Zunino *et al.* 2009, and references therein).

The rest of the paper is organized as follows. Section 2 elaborates further on the measure of relative efficiency based on algorithm complexity; Section 3 presents the data and discusses the method; Section 4 shows results, while Section 5 concludes the study.

2. The measure of algorithm complexity

In Shannon's entropy of information theory, the expected information content of a series is maximized if the series is genuinely random. Here, there is maximum uncertainty, and no redundancy in the series. The algorithmic complexity of a string is the length of the shortest computer program that can reproduce the string. But the shortest algorithm cannot be computed. However, there are several methods to circumvent the problem. Lempel and Ziv (1976) suggest a useful measure that does not rely on the shortest algorithm. And Kaspar and Schuster (1987) give an easily calculable measure of the Lempel-Ziv index, which runs as follows.

A program either inserts a new digit into the binary string $S = s_1, \dots, s_n$ or copies the new digit to S. The program then reconstructs the entire string up to the digit $s_r < s_n$ that has been newly inserted. Digit s_r does not originate in the substring $s_1, \dots, s_{r-1}$; otherwise, s_r could simply be copied from $s_1, \dots, s_{r-1}$. To learn whether the rest of S can be reconstructed by either simply copying or inserting new digits, s_{r+1} is initially chosen and subsequently confirmed as to whether it belongs to one of the substrings of S; in such a case, it can be obtained by simply copying it from S. If s_{r+1} can indeed be copied, the routine continues until a new digit (which once again needs to be inserted) appears. The number of newly inserted digits plus one (if the last copying step is not followed by insertion of a digit) yields the complexity measure c of the string S.

Consider the following three strings of 10 binary digits each.

A 0000000000
B 0101010101
C 0110001001

One might correctly guess that A is less random, so A is less complex than B which in turn is less complex than C. The complexity index c agrees with such an intuition. In the string A, one has only to insert the first zero and then rebuild the entire string by copying this digit; thus, $c = 2$, where c is the number of steps necessary to create a string. In the string B, one has to additionally insert digit 1 and then copy the substring 01 to reconstruct the entire string; thus, $c = 3$. In the string C, one has to further insert 10 and 001, and then copy 001; thus, $c = 5$.

The complexity of a string grows with its length. The genuinely random string asymptotically approaches its maximum complexity r as its length n grows following the rule $\lim_{n\to\infty} c = r = \frac{n}{\log_2 n}$ (Kaspar and Schuster 1987). One may thus compute a positive finite normalized complexity index $LZ = \frac{c}{r}$ to obtain the complexity of a string relative to that of a genuinely random one. Under the broad definition of complexity proposed by Lempel and Ziv (1976), almost all sequences of sufficiently large length are found to be complex. To obtain a useful measure of complexity, they then consider a De Bruijn sequence which is commonly viewed as a good finite approximation of a complex sequence (Lempel and Ziv 1976). After proving that the De Bruijn sequence is indeed complex according to their definition, and that its complexity index cannot be less than one, they decided to fix it as a benchmark against which other sequences could be compared. Thus, a finite sequence with a complexity index greater than one is guaranteed to be more complex than (or at least as complex as) a De Bruijn sequence of the same size. Note that the *LZ* index is not an absolute measure of the complexity (which is perhaps nonexistent), nor is the index ranged between zero and one. We provide more details on the *LZ* index in our previous work (Giglio *et al.* 2008a; 2008b; Giglio and Da Silva 2009).

To find the *LZ* index of a time series, sliding time windows are considered in our previous work. The index for every window is calculated and then the average is obtained. For example, in a time series of 2,000 data points and a chosen time window of 1,000 observations, the *LZ* index of the window from 1 to 1,000 is first computed; then the index of the window from 2 to 1,001 is derived, and so on, up to the index of the window from 1,001 to 2,000. Then the average of the indices is calculated. In this work, we consider windows of 5,000 observations.

3. Data and methods

We take high-frequency tick-by-tick stock return data from 43 companies listed on the Sao Paulo Stock Exchange (Bovespa) from the beginning of January 2007 to the final of December 2008, thus including September 15, 2008 (Lehman Brothers bankruptcy) and its aftermath. For the Petrobras corporate stock, for example, this means more than 5,000,000 data points. We first picked the 50 companies listed on the IBrX-50 index in 2008, but seven companies that went public only after 2007 were disregarded. Table 1 gives an overview of the data considered.

As the De Bruijn series is only an approximation of a truly random series, some efficiency values higher than one can emerge. However, this problem is minimized as one increases the data windows size (Giglio *et al.* 2008a; 2008b). Here, we consider data windows size of 5,000 data points, as observed. This allows for only 0.1% of values to be greater than one. We consider a simple ternary coding, and then assign bit 0 for zero returns, bit 1 for positive returns, and bit 2 for negative returns.

4. Results

Figure 1 shows the relative efficiency of four selected stocks (Aracruz, Petrobras PN, Sadia, and Vale PNA N1) over the period 2007–2008. As can be seen, the average *LZ* index is reduced dramatically after the Lehman Brothers bankruptcy. Figure 2 presents the histograms related to Figure 1.

Table 2 shows the average *LZ* index and its standard deviation for all the corporate stocks. Note that the magnitude of the variances is much lower than the magnitude of the averages for all the stocks. The data in Table 2 do confirm the pattern shown in Figure 1.

To assess whether the smaller average *LZ* index after the Lehman Brothers collapse does not come by chance, we carry out a nonparametric hypothesis test of means. Without resorting to any assumption of a normal distribution we estimate 95 percent confidence intervals from the bootstrap technique using 99 resamples for mean and 25 for variance (Efron and Tibshirani 1993). We then test

Hypothesis 0: mean (after the Lehman Brothers collapse) $\geq$ mean (before)
Hypothesis 1: mean (after) < mean (before)
Significance level = 0.05

For 37 out of 43 stocks the null is rejected, thus showing that we cannot discard the result that the 2008 financial crisis contributed to reduce market efficiency. The exceptions are the six stocks Ultrapar, BRF Foods, Petrobras ON, Gafisa, Ambev, and Usiminas ON N1 (at the bottom of Table 3). Figure 3 shows the histograms for the entire sample considered in this study.

5. Conclusion

Using high frequency data for the years of 2007 and 2008 of the corporate stocks listed on Bovespa, we detect a reduced efficiency rate for the great majority of the stocks in the aftermath of the 2008 crisis.

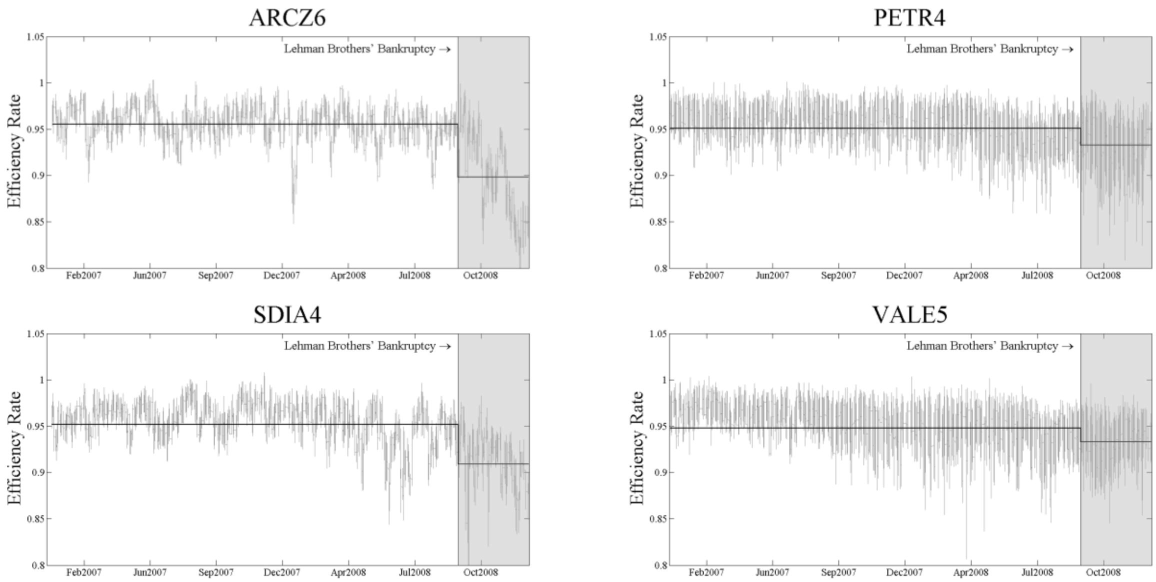


Figure 1. Relative efficiency of four selected corporate stocks listed on Bovespa. After the Lehman Brothers collapse there is a marked reduction in the average *LZ* index calculated from data. The solid lines represent the mean for each period.

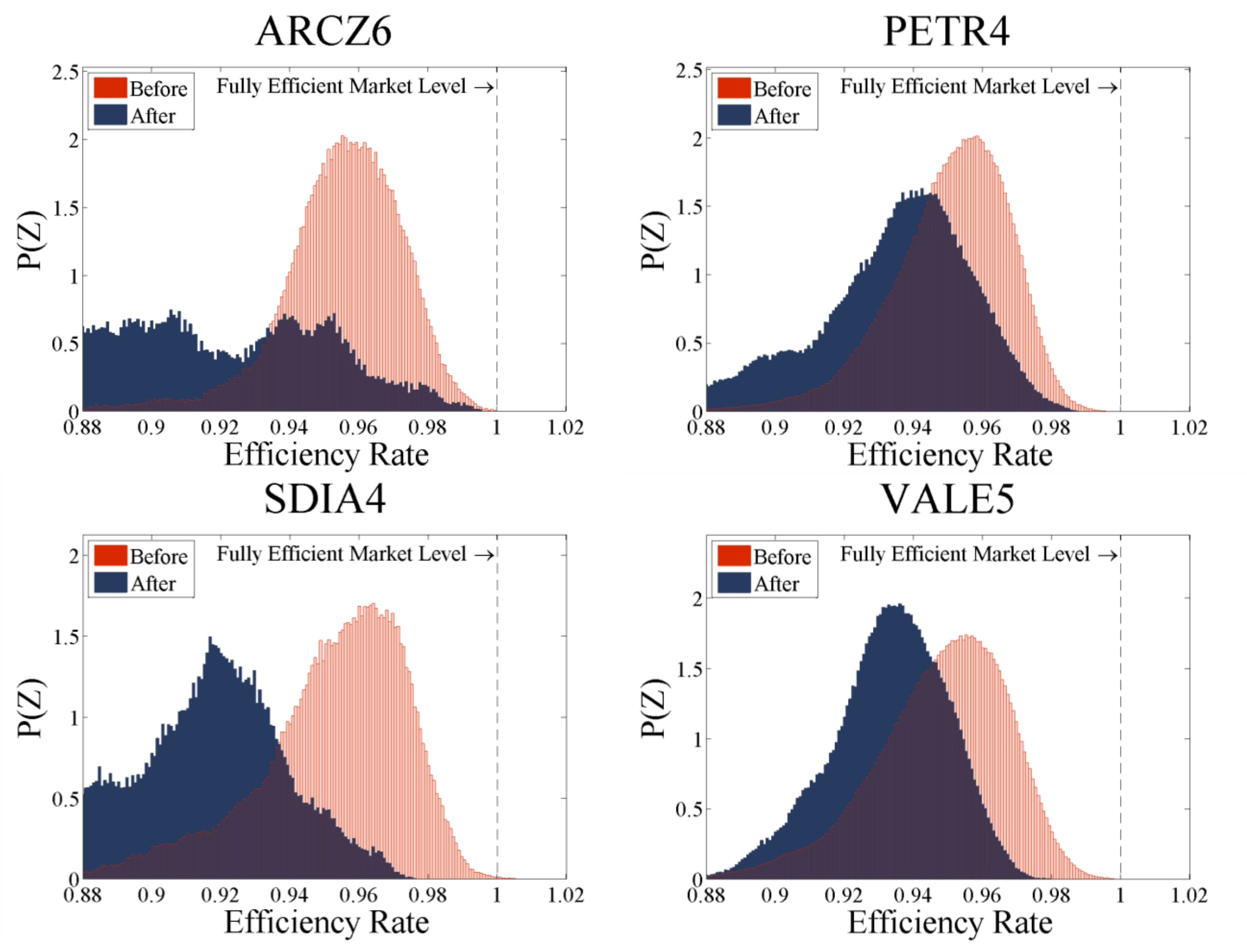
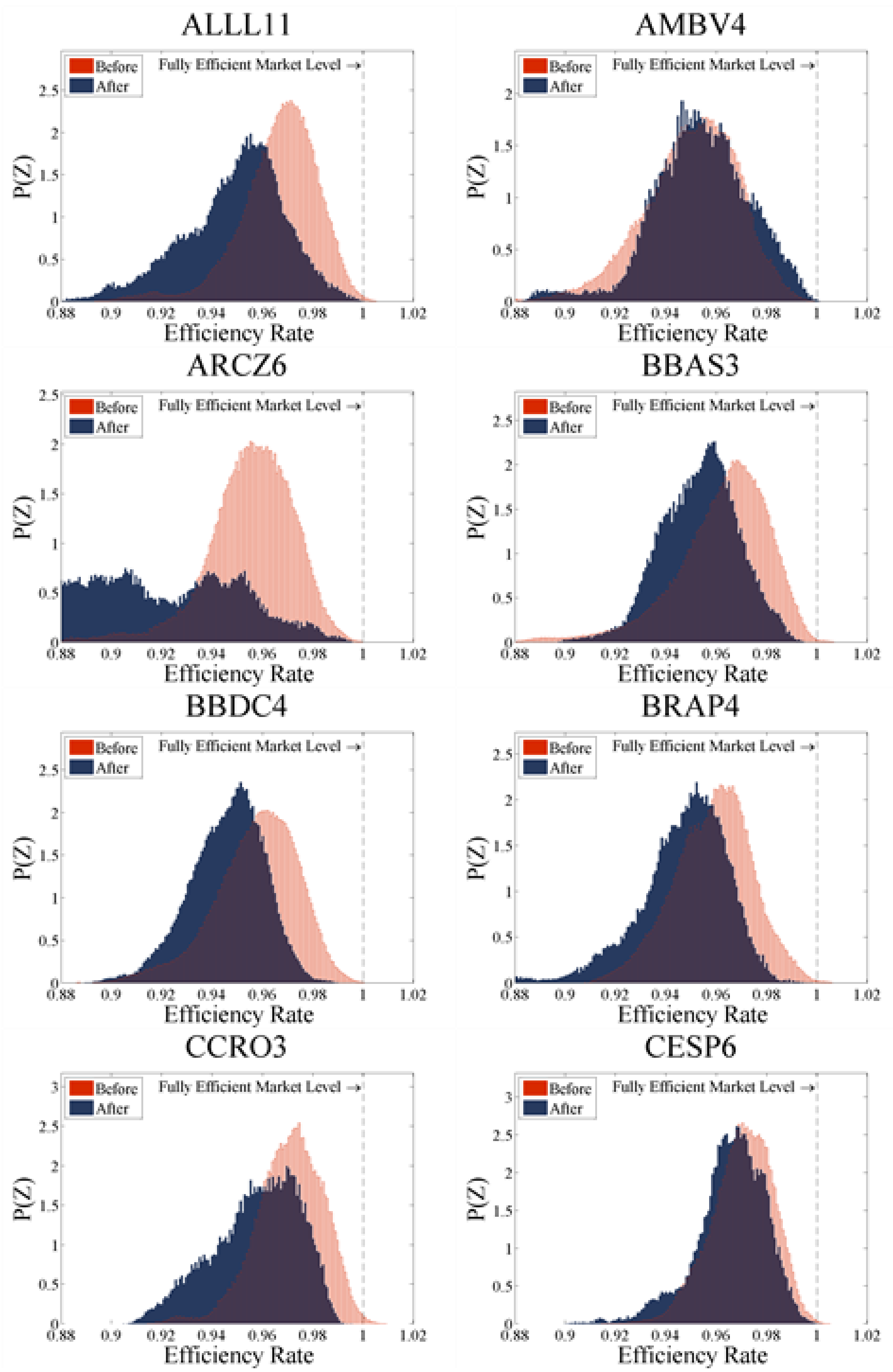
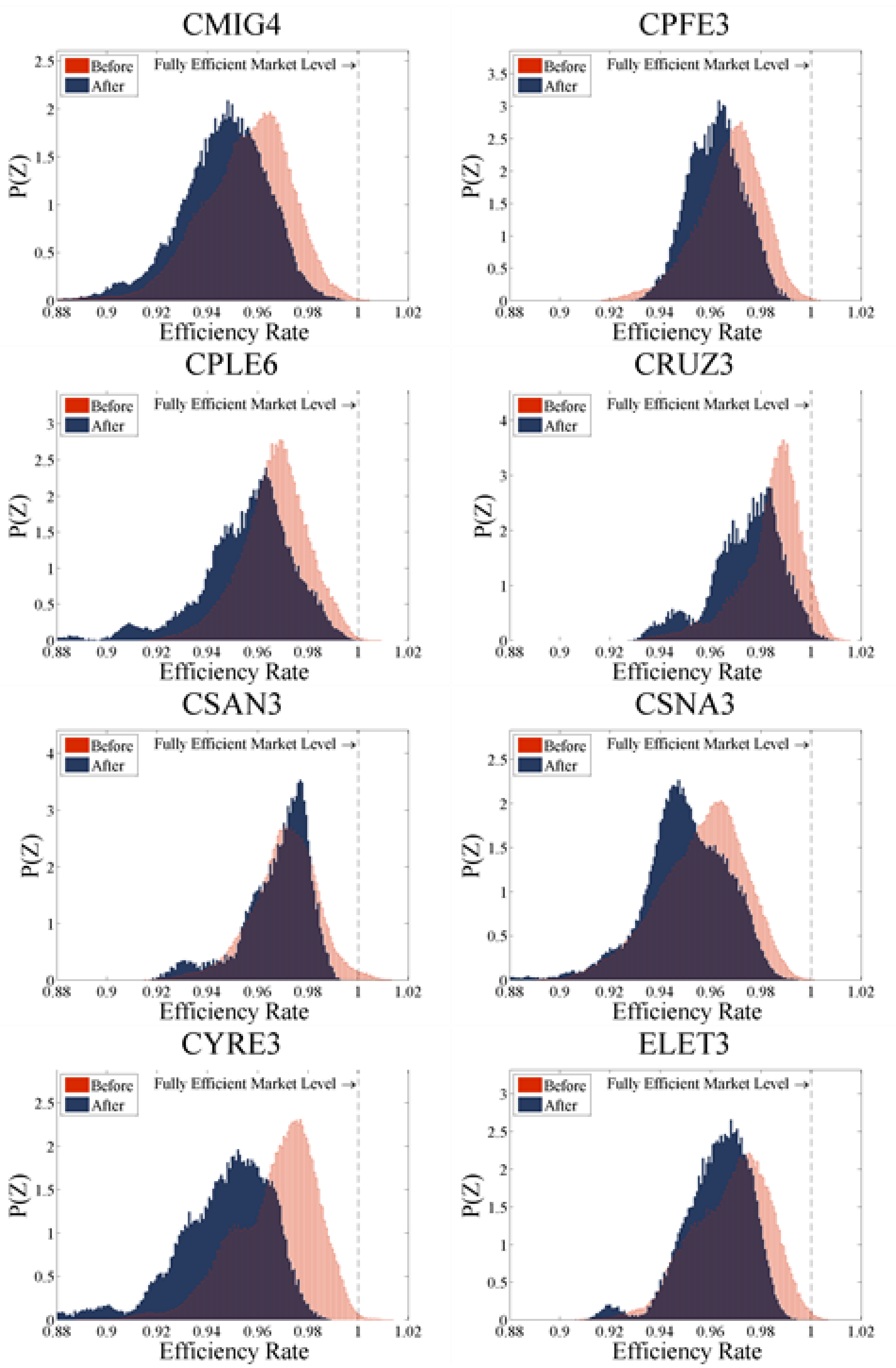
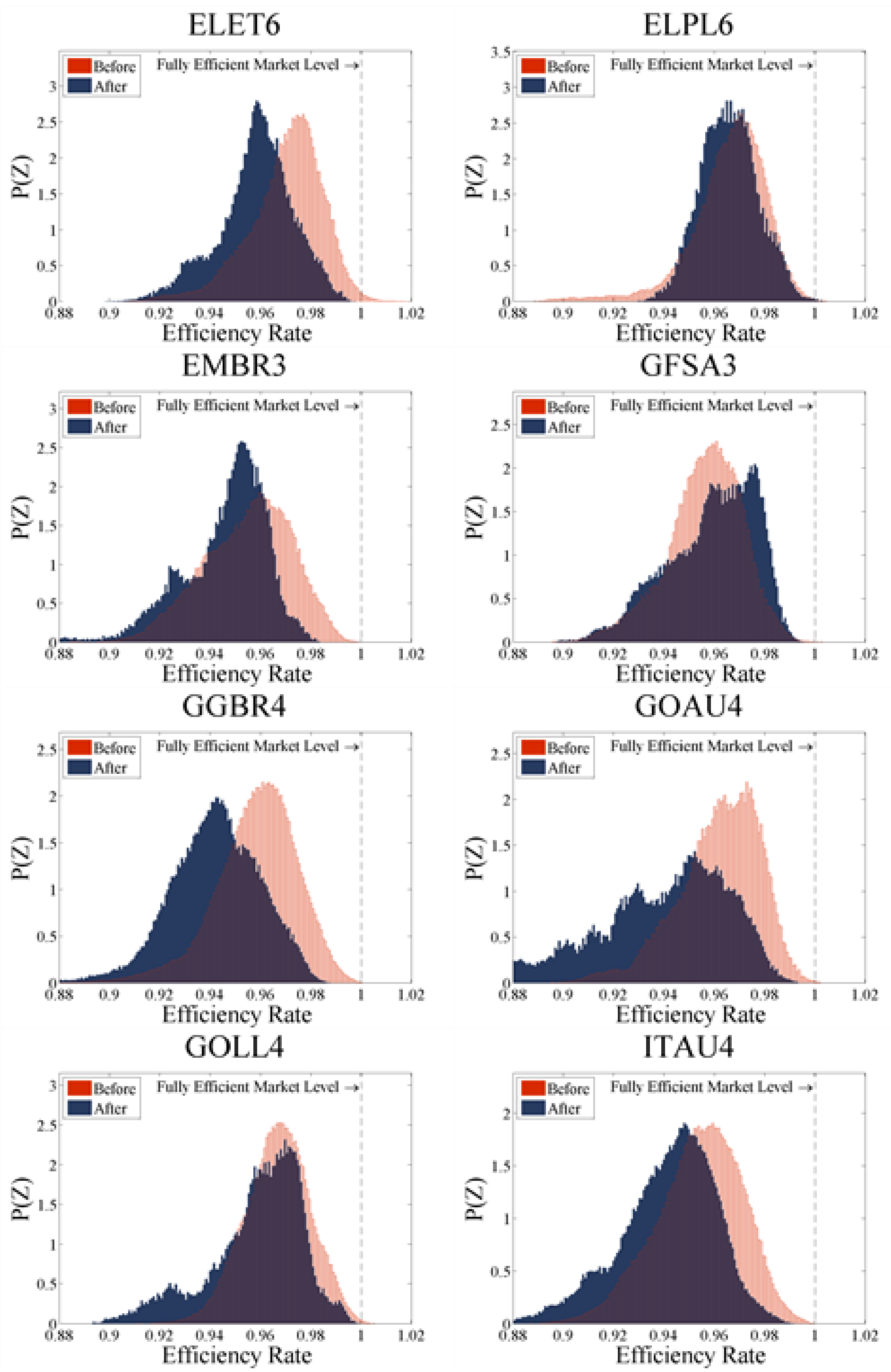
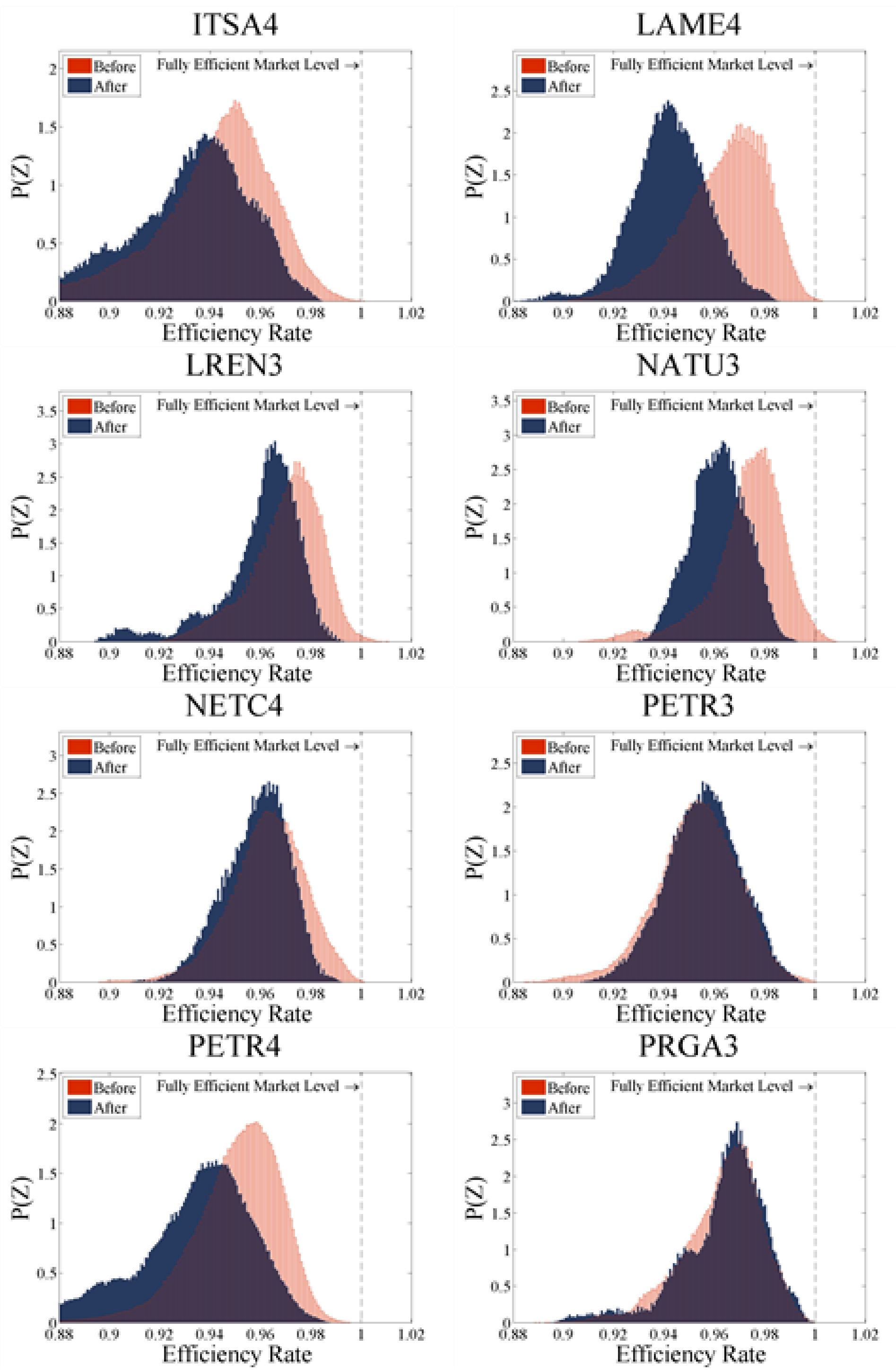


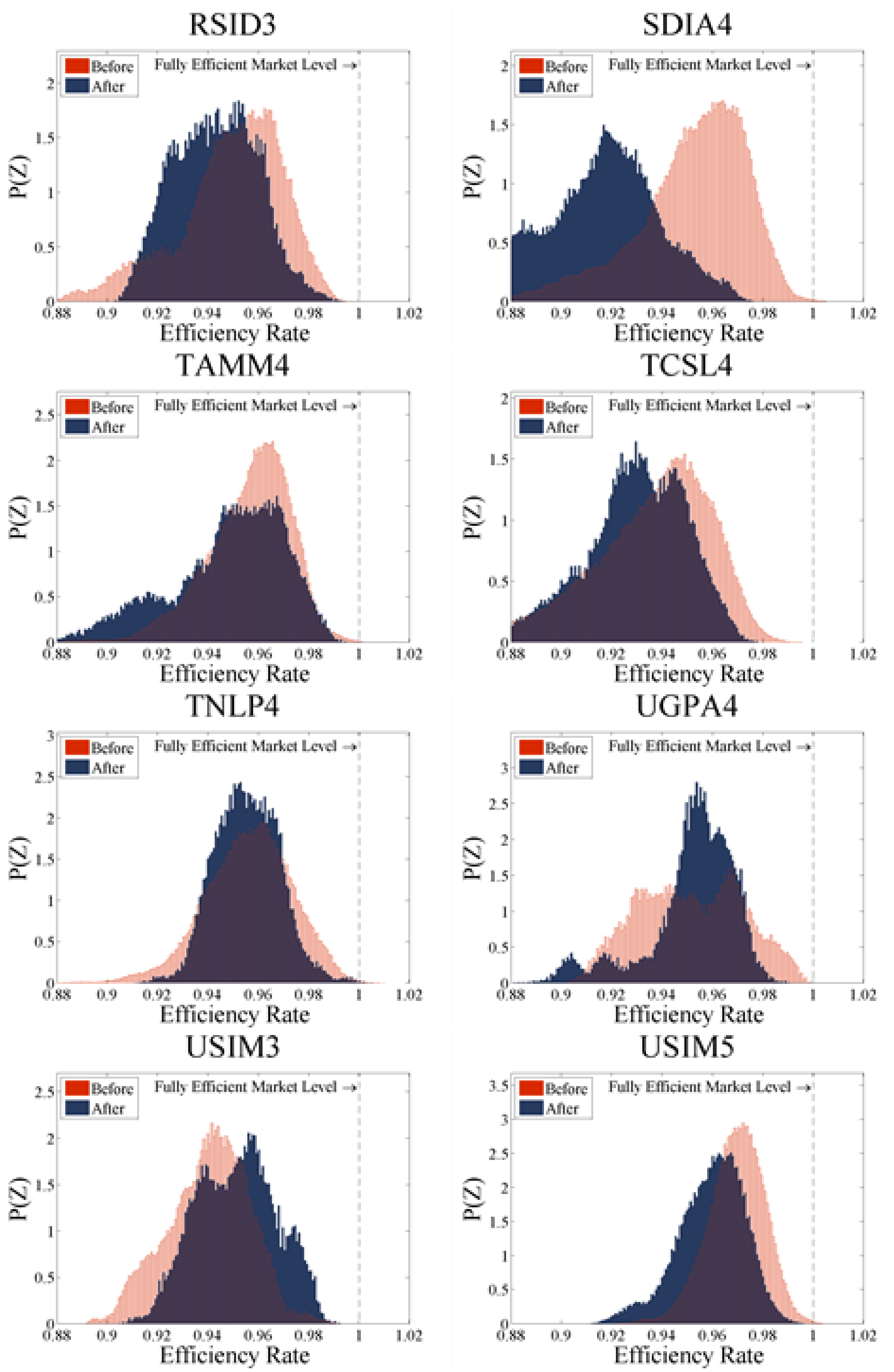
Figure 2. Histograms of the relative efficiency of four selected corporate stocks listed on Bovespa before and after the Lehman Brothers collapse.


ALLL11
AMBV4
ARCZ6
BBAS3
BBDC4
BRAP4
CCRO3
CESP6
Before
After
Fully Efficient Market Level →
P(Z)
Efficiency Rate
0.88
0.9
0.92
0.94
0.96
0.98
1
1.02


CMIG4
CPFE3
CPLE6
CRUZ3
CSAN3
CSNA3
CYRE3
ELET3
Before
After
Fully Efficient Market Level →
P(Z)
Efficiency Rate
0.88 0.9 0.92 0.94 0.96 0.98 1 1.02


ELET6
Before
After
Fully Efficient Market Level →
P(Z)
Efficiency Rate
ELPL6
Before
After
Fully Efficient Market Level →
P(Z)
Efficiency Rate
EMBR3
Before
After
Fully Efficient Market Level →
P(Z)
Efficiency Rate
GFSA3
Before
After
Fully Efficient Market Level →
P(Z)
Efficiency Rate
GGBR4
Before
After
Fully Efficient Market Level →
P(Z)
Efficiency Rate
GOAU4
Before
After
Fully Efficient Market Level →
P(Z)
Efficiency Rate
GOLL4
Before
After
Fully Efficient Market Level →
P(Z)
Efficiency Rate
ITAU4
Before
After
Fully Efficient Market Level →
P(Z)
Efficiency Rate


ITSA4
Before
After
Fully Efficient Market Level →
P(Z)
Efficiency Rate
LAME4
Before
After
Fully Efficient Market Level →
P(Z)
Efficiency Rate
LREN3
Before
After
Fully Efficient Market Level →
P(Z)
Efficiency Rate
NATU3
Before
After
Fully Efficient Market Level →
P(Z)
Efficiency Rate
NETC4
Before
After
Fully Efficient Market Level →
P(Z)
Efficiency Rate
PETR3
Before
After
Fully Efficient Market Level →
P(Z)
Efficiency Rate
PETR4
Before
After
Fully Efficient Market Level →
P(Z)
Efficiency Rate
PRGA3
Before
After
Fully Efficient Market Level →
P(Z)
Efficiency Rate


RSID3
SDIA4
TAMM4
TCSL4
TNLP4
UGPA4
USIM3
USIM5
Before
After
Fully Efficient Market Level →
P(Z)
Efficiency Rate

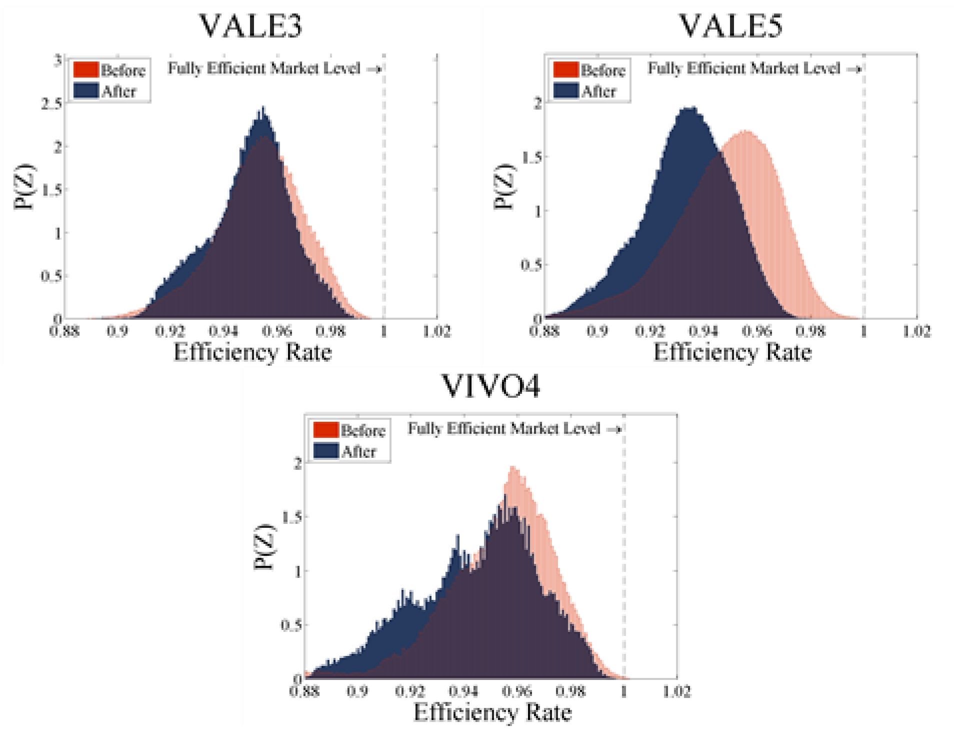


Figure 3. Histograms of the relative efficiency of the corporate stocks listed on Bovespa before and after the Lehman Brothers collapse.

Table 1. The data from the companies listed on Bovespa considered in this work

Corporate label	*Stock acronym*	*Type*
ALL AMER LAT	ALLL11	UNT N2
AMBEV	AMBV4	PN
ARACRUZ	ARCZ6	PNB N1
BRASIL	BBAS3	ON NM
BRADESCO	BBDC4	PN N1
BRADESPAR	BRAP4	PN N1
CCR RODOVIAS	CCRO3	ON NM
CESP	CESP6	PNB N1
CEMIG	CMIG4	PN N1
CPFL ENERGIA	CPFE3	ON NM
COPEL	CPLE6	PNB N1
SOUZA CRUZ	CRUZ3	ON ED
COSAN	CSAN3	ON NM
SID NACIONAL	CSNA3	ON
CYRELA REALT	CYRE3	ON NM
ELETROBRAS	ELET3	ON N1
ELETROBRAS	ELET6	PNB N1
ELETROPAULO	ELPL6	PNB N2
EMBRAER	EMBR3	ON NM
GAFISA	GFSA3	ON NM
GERDAU	GGBR4	PN N1
GERDAU MET	GOAU4	PN N1
GOL	GOLL4	PN N2
ITAUSA	ITSA4	PN EDJ N1
ITAUUNIBANCO	ITUB4	PN EX N1
LOJAS AMERIC	LAME4	PN
LOJAS RENNER	LREN3	ON NM
NATURA	NATU3	ON NM
NET	NETC4	PN N2
PETROBRAS	PETR3	ON
PETROBRAS	PETR4	PN
BRF FOODS	PRGA3	ON NM
ROSSI RESID	RSID3	ON NM
SADIA S/A	SDIA4	PN N1
TAM S/A	TAMM4	PN N2
TIM PART S/A	TCSL4	PN
TELEMAR	TNLP4	PN
ULTRAPAR	UGPA4	PN N1
USIMINAS	USIM3	ON N1
USIMINAS	USIM5	PNA N1
VALE	VALE3	ON N1
VALE	VALE5	PNA N1
VIVO	VIVO4	PN

Table 2. Average *LZ* index and standard deviation of the corporate stocks listed on Bovespa before and after the Lehman Brothers collapse

Stock	*Period*	*Mean*	*Std*	*Stock*	*Period*	*Mean*	*Std*
ALLL11	2007-2008	0.9629	0.0183	GOLL4	2007-2008	0.9656	0.0143
	Before	0.9668	0.0157		Before	0.9670	0.0130
	After	0.9493	0.0200		After	0.9587	0.0185
AMBV4	2007-2008	0.9509	0.0196	ITAU4	2007-2008	0.9506	0.0188
	Before	0.9498	0.0197		Before	0.9542	0.0171
	After	0.9542	0.0192		After	0.9415	0.0198
ARCZ6	2007-2008	0.9389	0.0386	ITSA4	2007-2008	0.9370	0.0279
	Before	0.9554	0.0182		Before	0.9408	0.0245
	After	0.8984	0.0450		After	0.9242	0.0341
BBAS3	2007-2008	0.9608	0.0184	LAME4	2007-2008	0.9588	0.0187
	Before	0.9627	0.0188		Before	0.9640	0.0168
	After	0.9538	0.0152		After	0.9427	0.0149
BBDC4	2007-2008	0.9551	0.0170	LREN3	2007-2008	0.9684	0.0151
	Before	0.9577	0.0167		Before	0.9705	0.0140
	After	0.9462	0.0148		After	0.9603	0.0165
BRAP4	2007-2008	0.9561	0.0169	NATU3	2007-2008	0.9715	0.0147
	Before	0.9590	0.0154		Before	0.9736	0.0147
	After	0.9472	0.0179		After	0.9621	0.0108
CCRO3	2007-2008	0.9684	0.0151	NETC4	2007-2008	0.9622	0.0147
	Before	0.9710	0.0133		Before	0.9629	0.0149
	After	0.9578	0.0169		After	0.9585	0.0127
CESP6	2007-2008	0.9697	0.0127	PETR3	2007-2008	0.9540	0.0159
	Before	0.9705	0.0122		Before	0.9533	0.0164
	After	0.9660	0.0140		After	0.9557	0.0145
CMIG4	2007-2008	0.9531	0.0187	PETR4	2007-2008	0.9465	0.0210
	Before	0.9554	0.0183		Before	0.9509	0.0174
	After	0.9465	0.0180		After	0.9328	0.0250
CPFE3	2007-2008	0.9667	0.0128	PRGA3	2007-2008	0.9634	0.0159
	Before	0.9677	0.0131		Before	0.9633	0.0157
	After	0.9624	0.0108		After	0.9640	0.0168
CPLE6	2007-2008	0.9644	0.0154	RSID3	2007-2008	0.9480	0.0199
	Before	0.9670	0.0132		Before	0.9494	0.0207
	After	0.9562	0.0184		After	0.9434	0.0164
CRUZ3	2007-2008	0.9819	0.0136	SDIA4	2007-2008	0.9409	0.0310
	Before	0.9841	0.0126		Before	0.9519	0.0225
	After	0.9745	0.0142		After	0.9090	0.0301
CSAN3	2007-2008	0.9704	0.0133	TAMM4	2007-2008	0.9545	0.0194
	Before	0.9707	0.0132		Before	0.9565	0.0169
	After	0.9679	0.0135		After	0.9461	0.0261

CSNA3	2007-2008	0.9549	0.0173	TCSL4	2007-2008	0.9308	0.0308
	Before	0.9565	0.0172		Before	0.9336	0.0277
	After	0.9502	0.0169		After	0.9206	0.0383
CYRE3	2007-2008	0.9630	0.0189	TNLP4	2007-2008	0.9563	0.0164
	Before	0.9672	0.0165		Before	0.9564	0.0172
	After	0.9471	0.0189		After	0.9558	0.0128
ELET3	2007-2008	0.9669	0.0150	UGPA4	2007-2008	0.9519	0.0196
	Before	0.9682	0.0152		Before	0.9518	0.0204
	After	0.9629	0.0133		After	0.9520	0.0177
ELET6	2007-2008	0.9678	0.0156	USIM3	2007-2008	0.9444	0.0168
	Before	0.9705	0.0147		Before	0.9408	0.0161
	After	0.9584	0.0149		After	0.9516	0.0158
ELPL6	2007-2008	0.9662	0.0143	USIM5	2007-2008	0.9675	0.0124
	Before	0.9662	0.0150		Before	0.9697	0.0111
	After	0.9662	0.0113		After	0.9598	0.0133
EMBR3	2007-2008	0.9532	0.0179	VALE3	2007-2008	0.9524	0.0163
	Before	0.9551	0.0174		Before	0.9533	0.0165
	After	0.9457	0.0178		After	0.9499	0.0155
GFSA3	2007-2008	0.9571	0.0157	VALE5	2007-2008	0.9442	0.0204
	Before	0.9565	0.0150		Before	0.9479	0.0199
	After	0.9592	0.0176		After	0.9329	0.0174
GGBR4	2007-2008	0.9541	0.0184	VIVO4	2007-2008	0.9515	0.0221
	Before	0.9586	0.0164		Before	0.9532	0.0216
	After	0.9422	0.0180		After	0.9451	0.0230
GOAU4	2007-2008	0.9547	0.0242				
	Before	0.9622	0.0162				
	After	0.9372	0.0301				

Table 3. Ranks of the corporate stocks listed on Bovespa according to their efficiency rate decreases before and after the Lehman Brothers collapse. Only the six stocks at the bottom had their relative efficiency increased after the collapse

Rank	*Corporate stock*	*Before minus after*	*Before minus after, %*	*Rank*	*Corporate stock*	*Before minus after*	*Before minus after, %*
1	ARCZ6	–0.0570	–5.9663	23	EMBR3	–0.0095	–0.9925
2	SDIA4	–0.0429	–4.5056	24	CMIG4	–0.0089	–0.9347
3	GOAU4	–0.0250	–2.5952	25	BBAS3	–0.0089	–0.9266
4	LAME4	–0.0214	–2.2147	26	GOLL4	–0.0083	–0.8563
5	CYRE3	–0.0201	–2.0802	27	VIVO4	–0.0081	–0.8508
6	PETR4	–0.0180	–1.8972	28	CSNA3	–0.0063	–0.6534
7	ALLL11	–0.0175	–1.8142	29	RSID3	–0.0061	–0.6393
8	ITSA4	–0.0167	–1.7729	30	ELET3	–0.0053	–0.5474
9	GGBR4	–0.0164	–1.7130	31	CPFE3	–0.0053	–0.5467
10	VALE5	–0.0149	–1.5762	32	CESP6	–0.0045	–0.4658
11	CCRO3	–0.0133	–1.3645	33	NETC4	–0.0044	–0.4549
12	TCSL4	–0.0130	–1.3935	34	VALE3	–0.0035	–0.3619
13	ITAU4	–0.0128	–1.3362	35	CSAN3	–0.0028	–0.2915
14	ELET6	–0.0120	–1.2406	36	TNLP4	–0.0007	–0.0711
15	BRAP4	–0.0118	–1.2315	37	ELPL6	0.0000	–0.0041
16	NATU3	–0.0116	–1.1873	38	UGPA4	0.0002	0.0221
17	BBDC4	–0.0115	–1.2050	39	PRGA3	0.0007	0.0696
18	CPLE6	–0.0108	–1.1189	40	PETR3	0.0024	0.2518
19	TAMM4	–0.0104	–1.0904	41	GFSA3	0.0027	0.2791
20	LREN3	–0.0102	–1.0551	42	AMBV4	0.0044	0.4590
21	USIM5	–0.0099	–1.0230	43	USIM3	0.0109	1.1565
22	CRUZ3	–0.0095	–0.9695				

References

Beechey, M., D. Gruen, and J. Vickery (2001) "The efficient market hypothesis: a survey" Reserve Bank of Australia research discussion paper number 2000-01.

Cajueiro, D.O., and B.M. Tabak (2004) "The Hurst exponent over time: testing the assertion that emerging markets are becoming more efficient" *Physica A* **336**(3-4), 521–537.

Campbell, J.Y., A.W. Lo, and A.C. MacKinlay (1997) *The Econometrics of Financial Markets*, Princeton University Press: Princeton.

Chaitin, G.J. (1966) "On the length of programs for computing finite binary sequences" *Journal of the Association for Computing Machinery* **13**(4), 547−569.

Efron, B., and R.J. Tibshirani (1993) *An Introduction to the Bootstrap*, Chapman & Hall: London.

Giglio, R., R. Matsushita, and S. Da Silva (2008a) "The relative efficiency of stockmarkets" *Economics Bulletin* **7**(6), 1–12.

Giglio, R., R. Matsushita, A. Figueiredo, I. Gleria, and S. Da Silva (2008b) "Algorithmic complexity theory and the relative efficiency of financial markets" *Europhysics Letters* **84**(4), 48005-1−48005-6.

Giglio, R., and S. Da Silva (2009) "Ranking the stocks listed on Bovespa according to their relative efficiency" *Applied Mathematical Sciences* **3**(43), 2133–2142.

Gligor, M., and M. Ausloos (2008) "Clusters in weighted macroeconomic networks: the EU case. Introducing the overlapping index of GDP/capita fluctuation correlations" *European Physical Journal B* **63**(4), 533–539.

Kaspar, F., and H.G. Schuster (1987) "Easily calculable measure for the complexity of spatiotemporal patterns" *Physical Review A* **36**(2), 842–848.

Kolmogorov, A.N. (1965) "Three approaches to the quantitative definition of information" *Problems of Information Transmission* **1**(1), 4−7.

Lempel, A., and J. Ziv (1976) "On the complexity of finite sequences" *IEEE Transactions on Information Theory* **22**(1), 75–81.

Mantegna, R.N. (1999) "Hierarchical structure in financial markets" *European Physical Journal B* **11**(1), 193–197.

Mantegna, R.N., and E. Stanley (2000) *An Introduction to Econophysics: Correlations and Complexity in Finance*, Cambridge University Press: Cambridge.

Xu, L., P.C. Ivanov, K. Hu, Z. Chen, A. Carbone, and H.E. Stanley (2005) "Quantifying signals with power-law correlations: a comparative study of detrending and moving average techniques" *Physical Review E* **71**(5), 051101-1−051101-14.

Zunino, L., M. Zanin, B.M. Tabak, D.G. Perez, and O.A. Rosso (2009) "Forbidden patterns, permutation entropy and stock market inefficiency" *Physica A* **388**(14), 2854–2864.