SCHOOL OF ECONOMICS

Discussion Paper 2005-12

Extreme-Valued Distributional Relationships in Asset Prices

Nagaratnam Jeyasreedharan
(University of Tasmania)

ISSN 1443-8593
ISBN 1 86295 294 9

Extreme-Valued Distributional Relationships in Asset Prices

by

Nagaratnam Jeyasreedharan

School of Economics
University of Tasmania
Private Bag 85
Hobart Tasmania 7001
E-mail: nj.sreedharan@utas.edu.au
Phone: +61 03 62267671
Fax: +61 03 62267587

August 2005

Abstract

Daily closing prices are typically used as proxies for daily security prices and for other financial time series data. Empirical observations of financial data, however, usually not only contain the closing prices, but also frequently include the opening, the highest and the lowest prices for specific horizons such as days, weeks and months. Obviously, any insight or information as to the distribution of the Open, High and Low prices should contribute to a better understanding of the data generation process (DGP).

The insight of this paper is that an exact functional representation can be used to depict the relationship between the observed Close, High and Low price-changes. The underlying distribution is, however, not defined explicitly but numerically determined as an empirical cumulative distribution function (ECDF). This approach enables extreme value distributions to be defined and statistically tested using the Kolmogorov-Smirnov (KS) goodness-of-fit test.

The findings shed new light on the statistical behaviour of financial asset returns, especially the High and Low log-returns. The multivariate relationships identified in this paper could be used to improve our understanding of the returns generation process of financial assets.

Key words: empirical cumulative distribution function (ECDF); empirical probability density function (EPDF); extreme-value theory (EVT); asset return distributions.

Topic Areas: (1) Asset Pricing; (3) Capital markets; (6) Finance Theory and Evidence; (11) Quantitative Finance.

JEL Classifications: D30 – General; G19 – Other.

1 Introduction

Daily closing prices are typically used as proxies for daily security prices and for other financial time series data. Empirical observations of financial data, however, usually not only contain the closing prices, but also frequently include the opening, the highest and the lowest prices for specific horizons such as days, weeks and months. The opening price refers to the price at the opening of the market; whereas the High and Low prices correspond to the two extremes: the highest and lowest prices of the day (see Fiess and MacDonald [2002]).

Obviously, any insight or information as to the distribution of the Open, High and Low prices should contribute to a better understanding of the data generation process (DGP). The High and Low price-changes are random variables that depend on the parent distribution of the underlying DGP and the sampling period. Unfortunately, there is neither an economic theory nor a statistical theory to assess the exact form of the parent distribution (of the DGP). Officer [1972] claimed that there is “no natural law that determines which particular function accurately describes the distribution of the variables, if indeed any analytical function accurately describes it.” This view is further reinforced by Los [2003] who states that “the scientific debate – about what kind of distributions best describe financial time series – is not yet settled, and maybe never will”.

Consequently, the “true” distribution of the parent variable is not likely to be precisely known and, therefore, if this distribution is not known, neither is the exact distribution of the extremes[1]. For this reason, the study of extreme asset price changes is carried out using *asymptotic* extreme value theory instead of

[1] Even small errors in the choice of the parent distribution cannot be permitted when making inferences of the extremes.

exact extreme value theory (see Longin [1996]). Asymptotic extreme value theory enables distribution-free results to be obtained. The asymptotic approach, however, is a limiting distribution and is consequently only relevant for large sample sizes.

The contribution of this paper is that an exact distribution function is used to depict the relationship between the observed Close, High and Low price-changes. The underlying distribution is, however, not defined but numerically determined as an empirical cumulative distribution function (ECDF). This approach enables extreme value distributions to be statistically tested using the Kolmogorov-Smirnov (KS) goodness-of-fit (GoF) test.

The daily extreme movements of the DJI30 index and its constituent stocks for the period 1/1/2000 to 1/1/2005 (5 years) are examined. The daily extreme movements are the High log-returns and the Low log-returns. It is shown empirically that the value of the order-parameters for the daily extreme High and Low log-returns fall in the range of 1 and 3 for DJI30 stocks. The DJI30 index log-returns, however, display a value of between 4 and 5. For both stocks and the index, the hypothesis that the distributions of the “extremals” (High and Low log-returns) are exact extreme-valued distributions of the “centrals” (Close log-returns) distributions cannot be significantly rejected for most of the cases considered.

The remainder of the paper is organized as follows: In Section 2 the functional form of the exact extreme-valued distributions is defined and discussed. In Section 3 the empirical distribution function is defined and discussed. The data is described in Section 4. The KS-test is defined and discussed in Section 5. The results are presented in Section 6. Section 7 gives the implications of the findings and suggestions for future research.

2 Exact Distribution Functions

The exact distribution of the extreme values can be written as functions of the initial distribution and the sample size n (see Gumbel [1958]). The probability, that all of n independent observations on a continuous variate are less than x is $F^n(x)$.

This can be interpreted as the probability $F_{n:n}(x)$ that the largest observation amongst n independent observations is less than or equal to x:

(1.1) $$F_{n:n}(x)=F^n(x)$$

Accordingly, given $F(x)$ one can compute $F_{n:n}(x)$. The different functions of $F_{n:n}(x)$ form a system of curves that shifts to the right without intersection for increasing n. All quantiles, the modes, means of the largest value increase with n. Correspondingly, the probability $1 - F_{1:n}(x)$ that the smallest among n independent observations is less than x is obtained from (see Gumbel [1958]):

(1.2) $$F_{1:n}(x) = 1 - (1 - F(x))^n$$

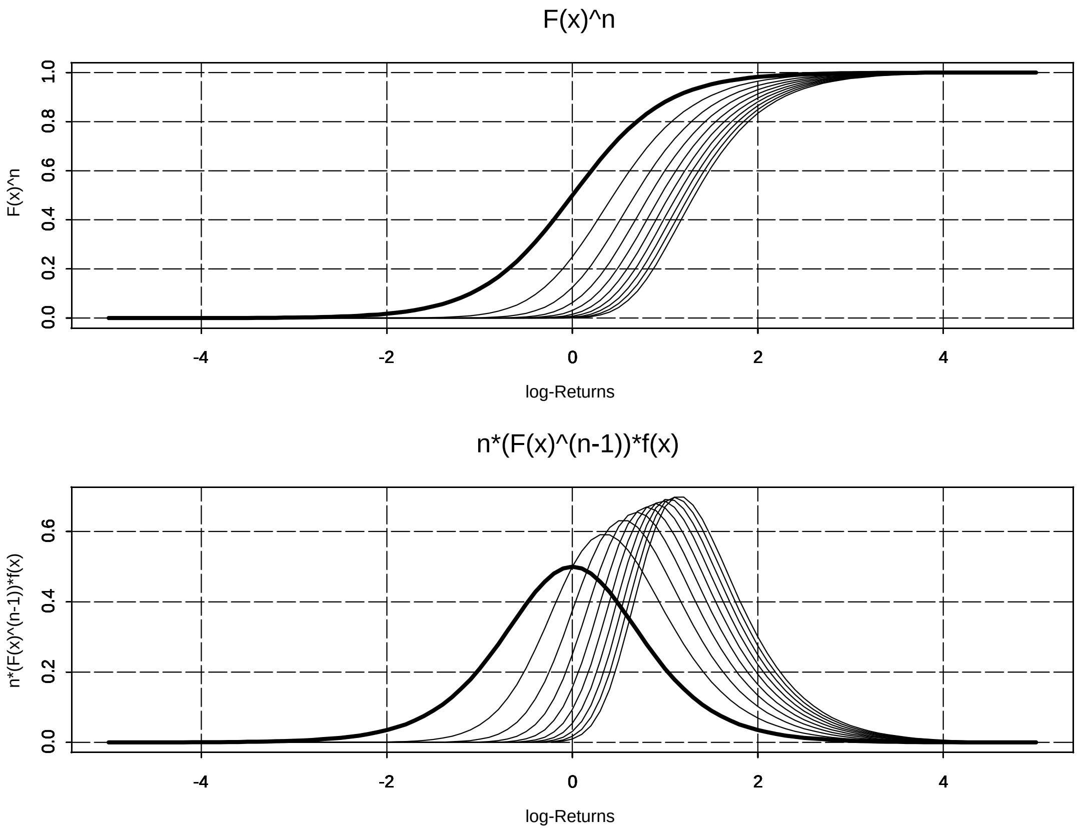

Figure 2-1 Exact Density and Cumulative Distribution Functions
Notes: The top-panel depicts the theoretical CDFs and the bottom-panel depicts the theoretical DDFs

The density functions, $f_{n:n}(x)$ and $f_{1:n}(x)$, of the largest and the smallest values are found by differentiation and are given as (see Gumbel [1958]):

(1.3) $$f_{n:n}(x) = nF(x)^{n-1} f(x)$$

(1.4) $$f_{1:n}(x) = n(1 - F(x))^{n-1} f(x)$$

Figure 2-1 illustrates the theoretical or exact cumulative and density functions depicted by equations (1.1) and (1.3). For increasing sample sizes, the curves representing consecutive distributions of extremes have different shapes. If the initial distribution is symmetrical about median zero, the cumulative extreme-valued distributions are no longer symmetrical since (see Gumbel [1958]):

(1.5) $$1 - F_{n:n}(x) \neq F_{n:n}(-x) \text{ and}$$

(1.6) $$1 - F_{1:n}(x) \neq F_{1:n}(-x)$$

Consequently, the process of taking extremes introduces asymmetry. In contrast, the largest and smallest values of a symmetrical distribution are mutually symmetrical (see Gumbel [1958]):

(1.7) $$F_{n:n}(x) = 1 - F_{1:n}(-x) \text{ and}$$

(1.8) $$f_{n:n}(x) = f_{1:n}(-x)$$

The symmetry principle, depicted by equations (1.7) and (1.8), means that from a given distribution of the largest value the distribution of the smallest value may be obtained by changing the sign of the observed variables. In two mutually symmetrical distributions, the distribution of the largest value of one is the distribution of the smallest of the other, and *vice versa* and can be depicted as:

(1.9) $$min\left\{x_1, x_2, ..., x_n\right\} = -max\left\{-x_1, -x_2, ..., -x_n\right\}$$

As noted by Gumbel [1958] and as we do in this paper, the study of extreme values may be restricted to that of only the largest value or maximum.

3 Empirical Distribution Functions

As already mentioned, in Section 1, if the distribution of the parent variable cannot be precisely defined then neither is the exact distribution of the extremes. We circumvent the problem by using the empirical distribution equivalents. The empirical distribution function (ECDF) of a sample, $F_n(x)$, is a step function defined as:

(1.10) $$F_n(x) = \begin{cases} 0\,; & x < x_{(1)} \\ i/n\,;\ x_{(i)} \le x < x_{(i+1)}; & i = 1,...,n-1 \\ 1\,; & x_{(n)} \le x \end{cases}$$

where $F_n(x)$ is the proportion of observations with a value less or equal to x, with increasing steps of $1/n$ at each observation. Figure 3-1 depicts the ECDFs and the empirical density distribution functions (EDDFs).

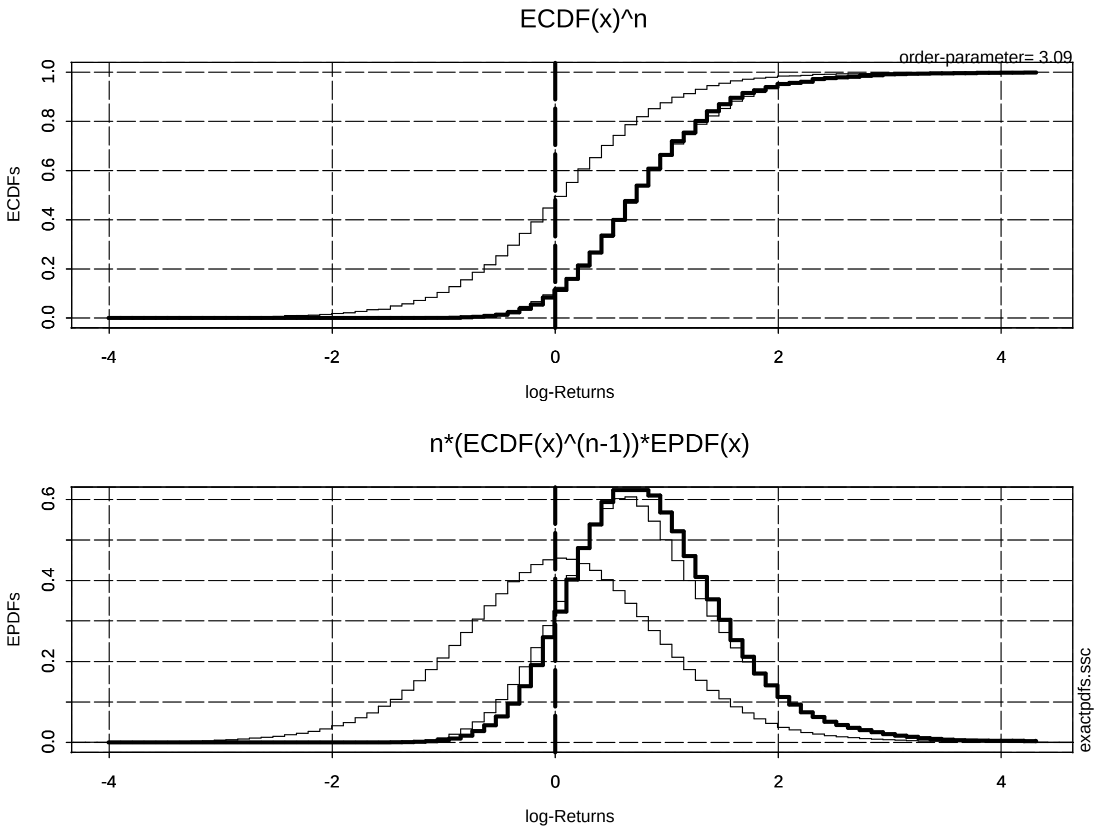

Figure 3-1 Empirical Density and Cumulative Distribution Functions
Notes: The top-panel depicts the ECDFs and the bottom-panel depicts the EPDFs.

The goodness-of-fit (GoF) test used in this paper compares two empirical distribution functions. The GoF test is based on the differences between the two fitted ECDFs and rejects the null when the differences are too large. Although the KS-test can be used for any distribution, critical values for these tests depend upon the null distribution. Fortunately there are modified KS-tests which are distribution free[2] (see Gibbons and Chakraborti [1992]). Figure 3-2 is a plot of the ECDF for the normalized trade-returns. The dark line is the symmetrical parent ECDF and the grey line is the asymmetrical extremal ECDF. Note that the two ECDFs are quite distinct from each other, with the ECDF for the extremes deviating quite significantly from the ECDF for the parent variates.

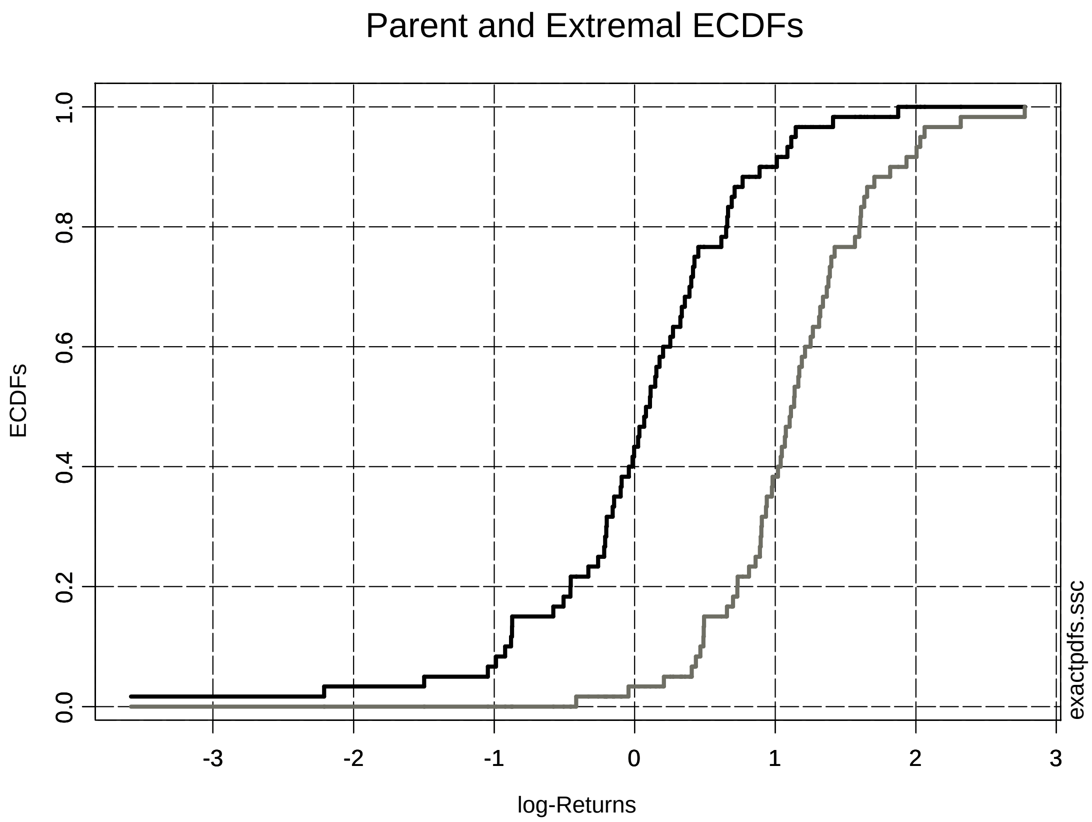

Figure 3-2: Two Empirical Cumulative Distribution Functions
Notes: The dark line is the symmetrical parent ECDF whereas the grey line is the asymmetrical extremal ECDF.

[2] The Wilcoxon ranksum test is appropriate to detect differences in location and the Siegel-Turkey test is especially appropriate to detect differences in dispersion, both of which are not of concern here.

The statistic measuring the difference between $F_n(x)$ and $F(x)$ are called the EDF statistics[3], D^+ and D^-, which are respectively the largest vertical difference when $F_n(x) > F(x)$ and the largest vertical difference when $F_n(x) < F(x)$. Formally,

(1.11) $$D^+ = \sup_x \{F_n(x) - F(x)\} \text{ and}$$

(1.12) $$D^- = \sup_x \{F(x) - F_n(x)\}$$

The more commonly used EDF statistic by Kolmogorov [1933] however is:

(1.13) $$D = \sup_x |F_n(x) - F(x)| = \max(D^+, D^-)$$

D is also known as the Kolmogorov-Smirnov statistic (see Chakravart, Laha, et al. [1967]). The Kolmogorov-Smirnov (KS) test is used in this paper to determine if there is a direct relationship between the distributions of the Close, High and Low log-returns.

4 Data

The DJI30 index and the DJI30 stocks dataset each consist of 1255 daily logarithmic returns that cover a 5-year sample period from 1/1/2000 to 1/1/2005. The daily logarithmic returns are computed as Close-to-Close, High-to-Close and Low-to-Close log-returns.

---

[3] Note $F_n(x)$ is used to represent the empirical distribution and $F_{n:n}(x)$ is used to represent the exact nth-order extremal distribution.

The DJI30 Index logarithmic returns are shown in Figure 4-1. The top-panel depicts the High and Low log-returns, whereas the bottom-panel depicts the Close log-returns. The asymmetry in the High Low log-returns timeseries can be seen from the top-panel plot. In both panels the heteroskedasticity is similar across the different log-returns. Whatever the DGP that is generating the Close-to-Close log-returns also has the same volatility influence on the High and Low log-returns.

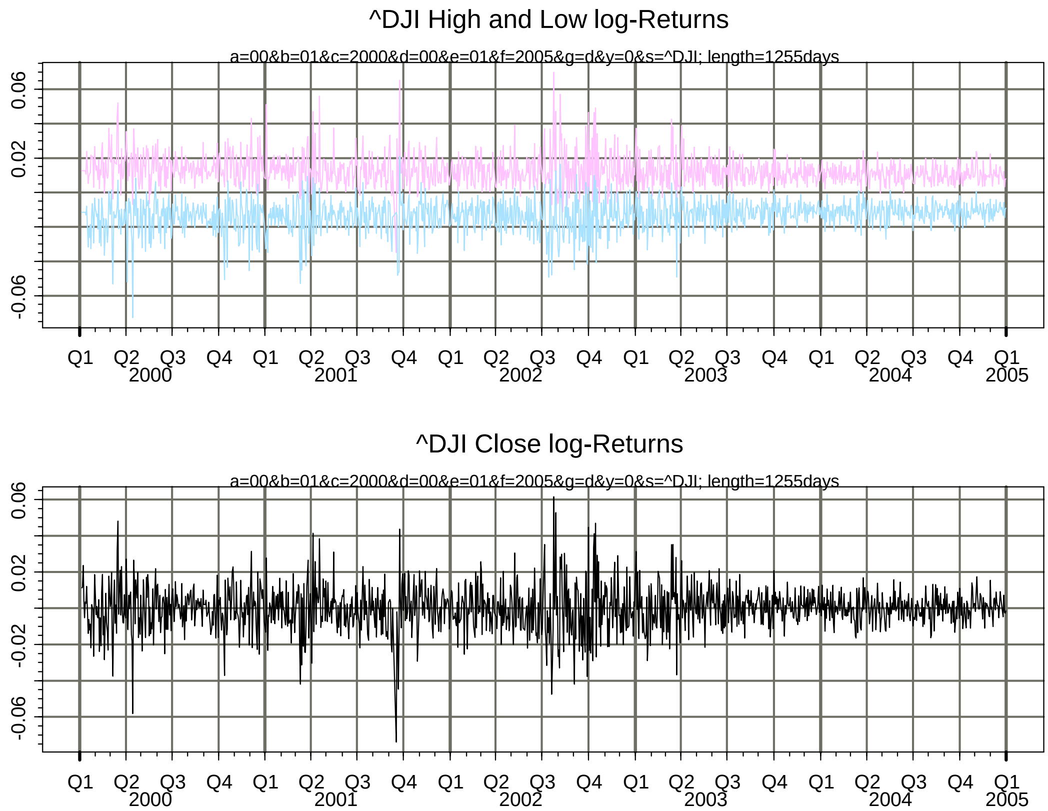

Figure 4-1 DJI30 High, Low and Close log-returns
Notes: The top panel depicts the High and Low log-returns. The bottom panel depicts the Close log-returns.

The summary statistics for the DJI30 index is listed in Table 4-1. The non-symmetry in the High and Low log-returns is affirmed by the positive mean and skewness values across all the stocks and the ^DJI index (the Low log-returns have been multiplied by a factor of -1.0).

Symbol	High.mean	High.stdev	High.skew	High.kurt	Low.mean	Low.stdev	Low.skew	Low.kurt
AA	0.0157	0.0185	1.0198	6.388	0.0168	0.0193	0.9285	6.4388
AXP	0.0091	0.0173	1.0686	5.4215	0.0085	0.0186	0.9225	6.045
BA	0.0111	0.0173	0.7265	7.987	0.0146	0.0177	1.2414	7.1911
C	0.014	0.016	1.2745	7.8481	0.0095	0.0187	0.9429	7.29
CAT	0.0103	0.0167	1.0592	6.0721	0.0136	0.016	1.1595	6.9862
DD	0.0095	0.0158	1.1188	7.5758	0.0132	0.0145	0.8377	5.8461
DIS	0.0104	0.0199	0.9365	10.8747	0.0169	0.0198	0.3538	7.4579
EK	0.0087	0.018	0.977	9.2804	0.0139	0.016	1.0743	8.7591
GE	0.0133	0.0159	1.4112	8.2334	0.0142	0.0168	1.1696	9.0923
GM	0.0102	0.0169	0.7492	7.0761	0.0136	0.0168	1.1047	6.884
HD	0.0102	0.0196	0.9072	7.6701	0.0116	0.0195	0.7943	6.1487
HON	0.0116	0.0195	0.8653	7.0993	0.0167	0.021	1.2357	9.3602
HPQ	0.0187	0.0259	0.7976	8.019	0.0199	0.0249	1.1704	7.5383
IBM	0.0128	0.0167	0.9761	8.121	0.0088	0.0169	0.4658	7.081
INTC	0.0212	0.0244	0.6694	6.4131	0.0194	0.0247	0.7028	6.5853
IP	0.0101	0.0168	1.4373	8.1248	0.0146	0.0161	1.0351	6.3262
JNJ	0.0076	0.0124	1.0345	9.1224	0.0102	0.0119	0.8069	5.6445
JPM	0.0158	0.0197	1.2037	7.4615	0.0103	0.0212	0.9804	6.9648
KO	0.0066	0.0138	1.0689	7.2536	0.0094	0.0144	1.4844	10.4074
MCD	0.0127	0.0151	0.9576	8.3085	0.0138	0.0157	1.1847	7.418
MMM	0.0079	0.0136	1.6677	8.3985	0.0106	0.0131	1.2419	9.0512
MO	0.009	0.0168	1.6792	10.1571	0.0075	0.0181	1.2319	9.7345
MRK	0.0116	0.0143	0.8911	7.6115	0.013	0.0153	1.536	8.8304
MSFT	0.0109	0.0185	0.7323	7.5007	0.0147	0.0184	0.762	6.4343
PG	0.0057	0.0125	1.9277	10.9469	0.0075	0.0123	1.3346	10.6521
SBC	0.0104	0.0173	0.8055	7.7778	0.011	0.0186	1.685	10.078
T	0.0102	0.0211	0.849	8.5899	0.0181	0.0203	1.1682	7.1529
UTX	0.0103	0.0155	1.1748	6.2333	0.009	0.0156	1.0359	5.5246
WMT	0.0097	0.0165	1.3738	8.3027	0.0101	0.017	1.0227	6.439
XOM	0.0103	0.0112	0.8232	5.0765	0.0103	0.0129	1.2664	7.2275
^DJI	0.0128	0.0091	1.093	7.4939	0.0134	0.0092	0.8975	6.1705

Table 4-1 Summary Statistics for DJI stocks and index.

Notes: High.mean=mean of High log-returns; High.stdev=standard deviation of High log-returns; High.skew=skewness of High log-returns; High.kurt=kurtosis of High log-returns; Low.mean=mean of Low log-returns; Low.stdev=standard deviation of Low log-returns; Low.skew=skewness of Low log-returns; Low.kurt=kurtosis of Low log-returns.

The non-normality of the High and Low log-returns is affirmed by the significant skewness and kurtosis values observed. The skewness values range from 0.7265 to 1.9277 and the kurtosis values range from 5.0765 to 10.9469. The standard deviation for the DJI30 index log-returns is smaller than the standard deviations of the stocks, as one would expect.

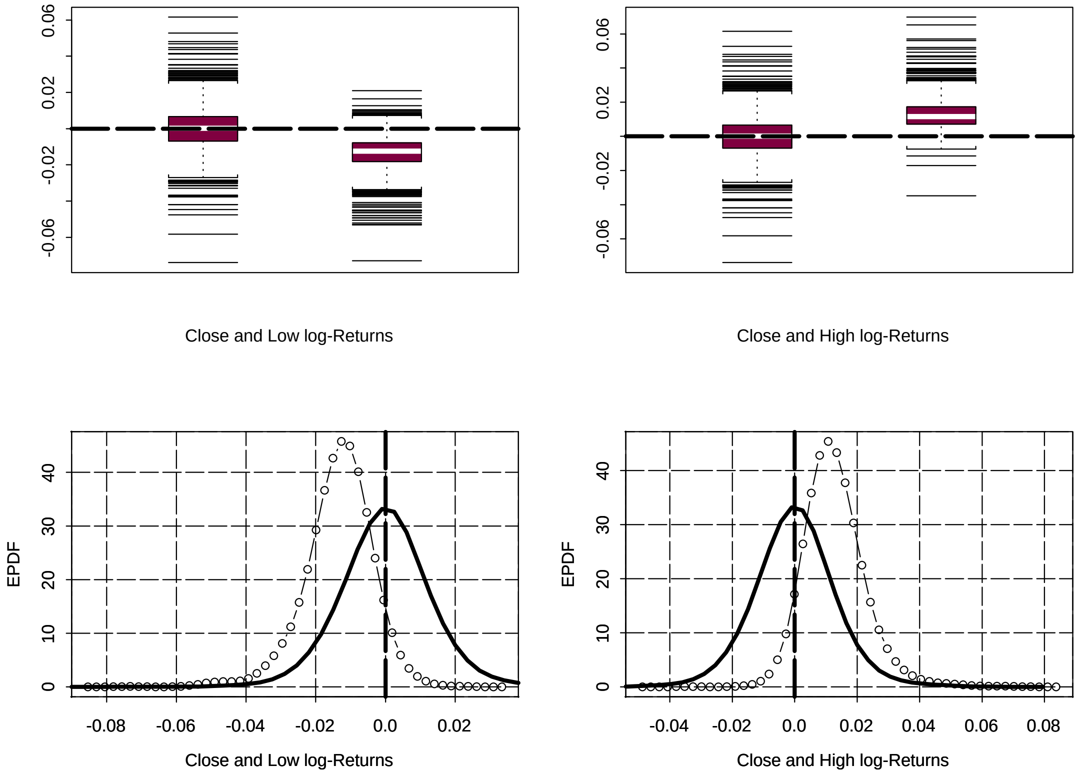

Figure 4-2: Boxplots and PDFs for DJI30 High, Low and Close log-returns
Notes: The top-panel depicts boxplots of log-returns. The bottom-panel depicts the probability density functions of log-returns.

Figure 4-2 displays the boxplots and empirical density functions for the High, Low and Close log-returns. The empirical density functions are fitted using kernel estimates (see Wegman [1972]). The right-panel shows the positive skewness inherent High log-returns and the left-panel the negative skewness in Low log-returns. The bottom-panel depicts symmetrical distributions for Close log-returns and asymmetrical distributions for the High and Low log-returns. The Low log-returns are left-skewed and the High log-returns are right-skewed. Note that the Low log-returns are to the left of the Close log-returns and the High log-returns are to the right of the Close log-returns.

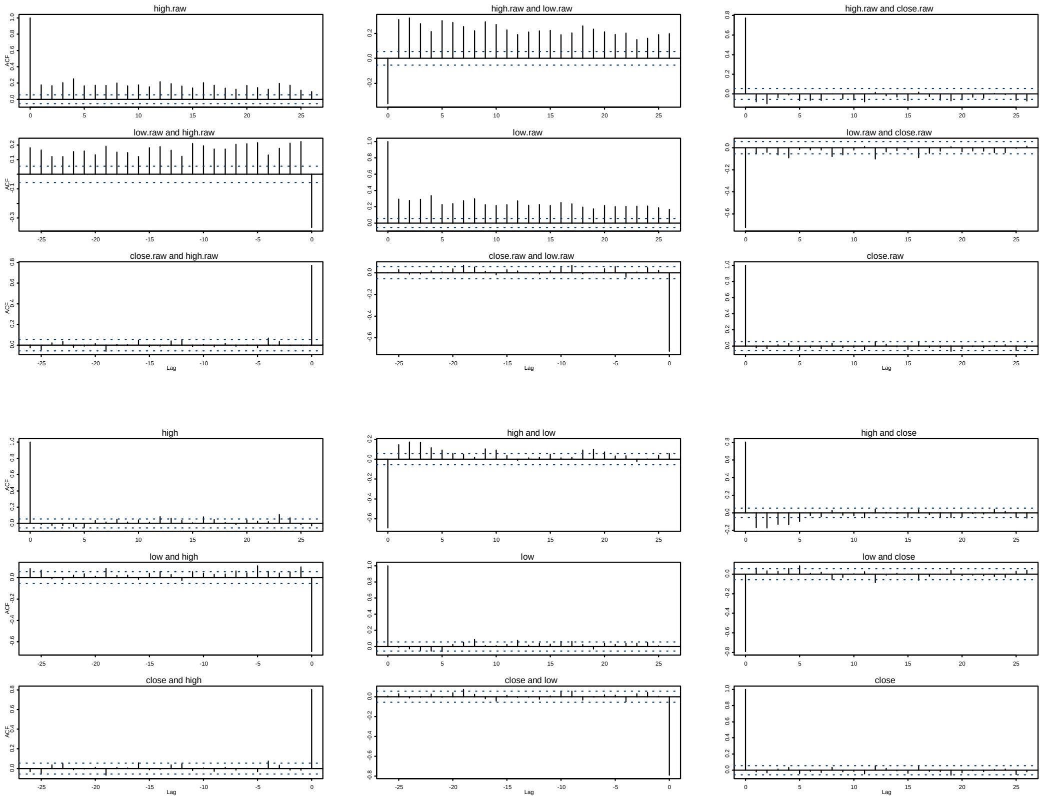

Figure 4-3 ACFs before and after Cochrane-Orcutt corrections
Notes; Top panels depict the ACFs before applying the Cochrane-Orcutt Correction. The bottom panels depict the ACFs after applying the Cochrane-Orcutt Correction.

In Figure 2-1 the ACF outcomes of Cochrane-Orcutt corrections are illustrated (see Cochrane and Orcutt [1949]). The ACF plots indicate that all the auto-correlations at various lags in the Close, High and Low log-returns have been effectively removed. Consequently, reliable statistical inferences can be made from the statistical results obtained from testing the High, Low and Close log-returns.

5 Methodology

The Kolmogorov-Smirnov (KS) test is used in this paper to determine if there is a direct relationship between the distributions of the Close, High and Low log-returns (see Chakravart, Laha, et al. [1967]). The test is based on the empirical cumulative distribution function (ECDF). Given n ordered data points, the test statistic is defined as:

$$
\begin{aligned}
D^{+} &= \max_{i=1,\ldots,n}[i/n - p_{(i)}] \\
D^{-} &= \max_{i=1,\ldots,n}[p_{(i)} - (i-1)/n] \qquad (1.14) \\
D &= \max\ [D^{+}, D^{-}]
\end{aligned}
$$

where i/n is the number of points less than $p_{(i)}$ and the $p_{(i)}$s are ordered from smallest to largest value. This is a step function that increases by $1/n$ at the value of each ordered data point.

An attractive feature of this test is the distribution of the KS test-statistic itself does not depend on the underlying cumulative distribution function being tested. Another advantage is that it is an exact test (the Chi-Squared test depends on an adequate sample size for the approximations to be valid). The KS-test if used with estimated parameters, however, tends to be conservative. This conservatism means the actual significance level for the test is smaller than the stated significance level. A conservative test may incorrectly fail to reject the null hypothesis, therefore decreasing its power[4] (Type II error).

---

[4] This is only a problem for sample sizes of less than 50. In this paper, the samples sizes are over 1000.

The KS test-statistic is used to fit and to compare the empirical cumulative distribution functions of the Close log-returns with the empirical cumulative distribution functions High and Low log-returns. Figure 5-1 illustrates the fitting process used. The fitting process used is based on a quasi-Newton method using the double dogleg step with the BFGS secant update to the Hessian (see Dennis and Mei [1979] and Dennis, Gay, et al. [1981]). The process finds a local minimum of the nonlinear EDCF function by varying the order parameter n.

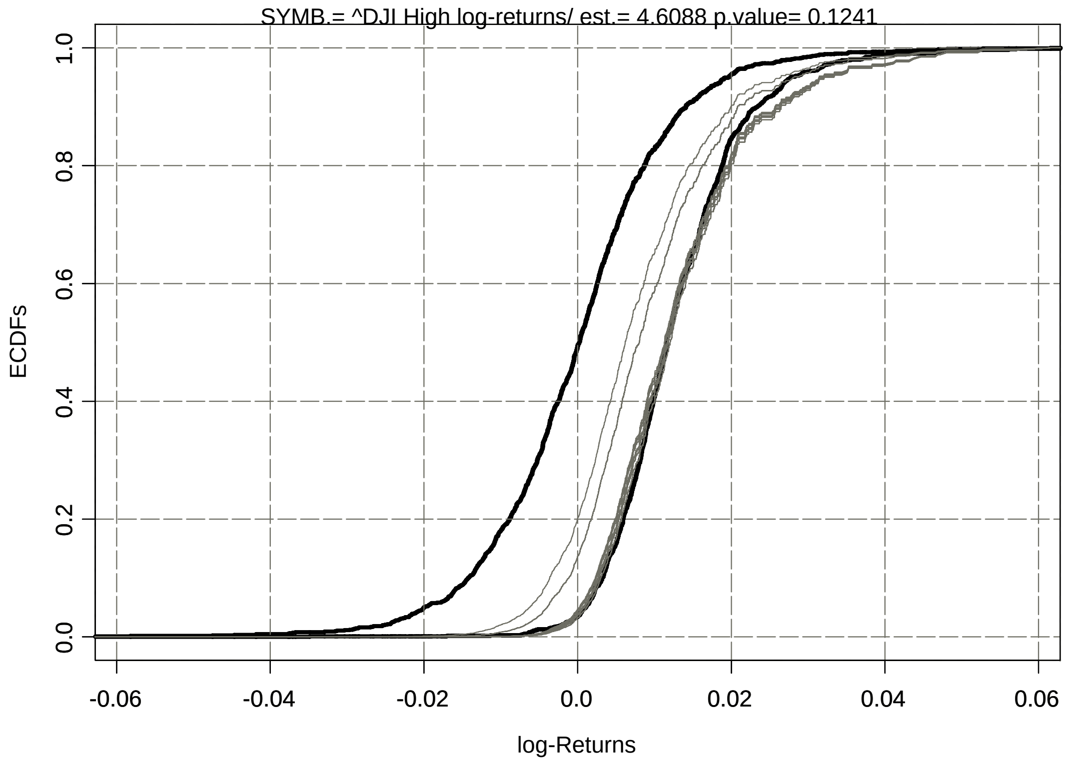

Figure 5-1 Observed and Fitted ECDFs
Notes: The dark lines depict the parent ECDF and the extremal ECDF. The multiple grey lines depict the ECDFs for the various values of order-parameter attempted by the fitting process.

The value of the KS statistic for two samples is based on the procedure given by Hollander and Wolfe [1999]. The p-value of the KS-statistic is determined using the algorithm given by Kim and Jennrich [1970].

In this paper, the following four hypotheses are tested.

Hypothesis 1, *H1* (High versus Close log-returns):

$$(15)\qquad \begin{array}{l} H1_N : ECDF_{n:n}(High\ returns) = ECDF^n(Close\ returns) \\ H1_A : ECDF_{n:n}(High\ returns) \neq ECDF^n(Close\ returns) \end{array}$$

Hypothesis 2, $H2$, (Low versus Close log-returns):

$$(16)\qquad \begin{array}{l} H2_N : ECDF_{n:n}(-Low\ returns) = ECDF^n(Close\ returns) \\ H2_A : ECDF_{n:n}(-Low\ returns) \neq ECDF^n(Close\ returns) \end{array}$$

Hypothesis 3, $H3$, (High versus Low log-returns):

$$(17)\qquad \begin{array}{l} H3_N : ECDF_{n:n}(High\ returns) = ECDF^n(-Low\ returns) \\ H3_A : ECDF_{n:n}(High\ returns) \neq ECDF^n(-Low\ returns) \end{array}$$

Hypothesis 4, $H4$, (Low versus High log-returns):

$$(18)\qquad \begin{array}{l} H4_N : ECDF_{n:n}(-Low\ returns) = ECDF^n(High\ returns) \\ H4_A : ECDF_{n:n}(-Low\ returns) \neq ECDF^n(High\ returns) \end{array}$$

The best order parameter n that minimizes KS-statistic is determined initially and subsequently the minimal KS-statistic obtained is tested for statistical significance. The hypotheses are tested against p-values of 1% (“***”), 5% (“**”) and 10% (“*”).

6 Results

Figure 6-1 depicts the “original” ECDFs (bold dark lines) and the “fitted” ECDFs (thin grey lines) for the DJI30 index logarithmic returns. The top-panel shows all the DJI30 log-prices for the period 1/1/2000 to 1/1/2005. The High, Low, Open and Close log-prices are illustrated. The middle-panel depicts the High log-returns versus the Close log-returns; whereas the bottom-panel depicts the Low log-returns versus the Close log-returns.

The fitted order-parameters are 4.6088 and 5.0953 for the High and Low log-returns respectively. Both estimates cannot be rejected even at the 10% level

("*"). The fitted ECDFs appear to progress towards the extremal ECDFs in the plots. In Figure 6-2 the ECDFs for two of the DJI30 stocks are illustrated: Hewlett-Packard and International Business Machines. The ECDF fits for HPQ cannot be rejected at the 10% level. The ECDF fit for IBM High log-returns can be rejected at the 5% level and the ECDF fit for Low IBM log-returns can be only be rejected at the 10% level. These findings are consistent across all the other DJI30 stocks.

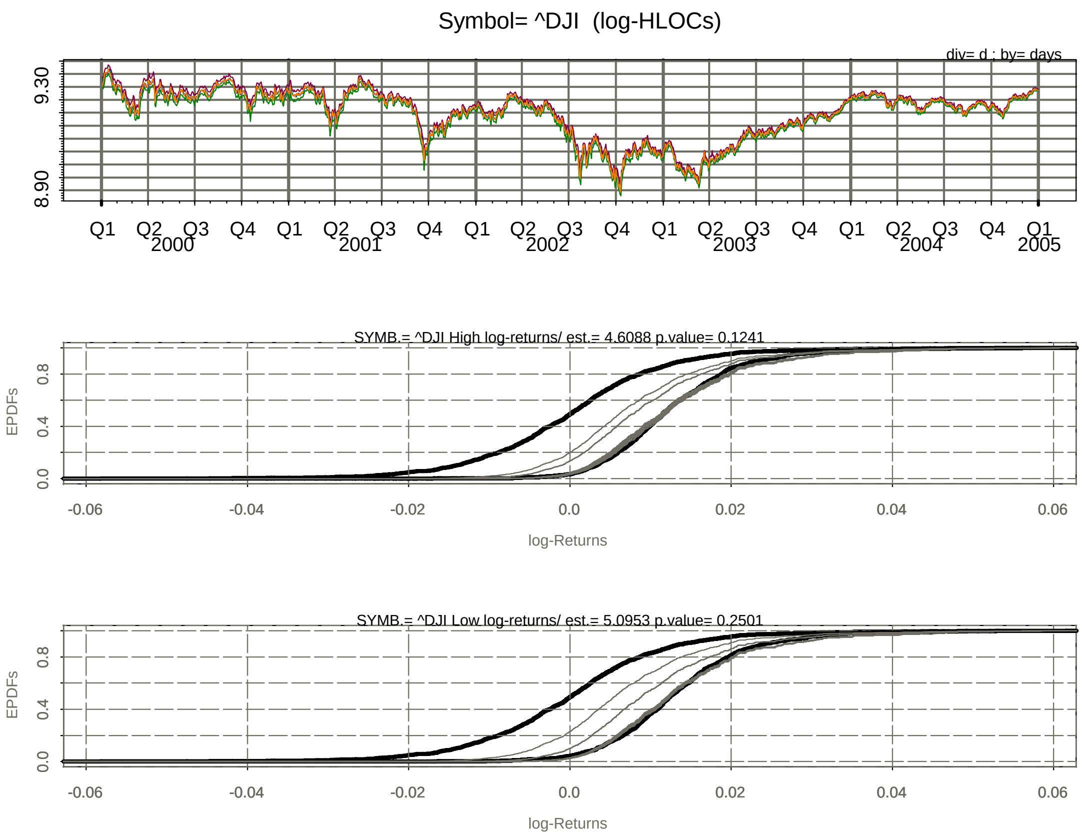

Figure 6-1 ECDFs of DJI30 Index Close, High and Low log-returns

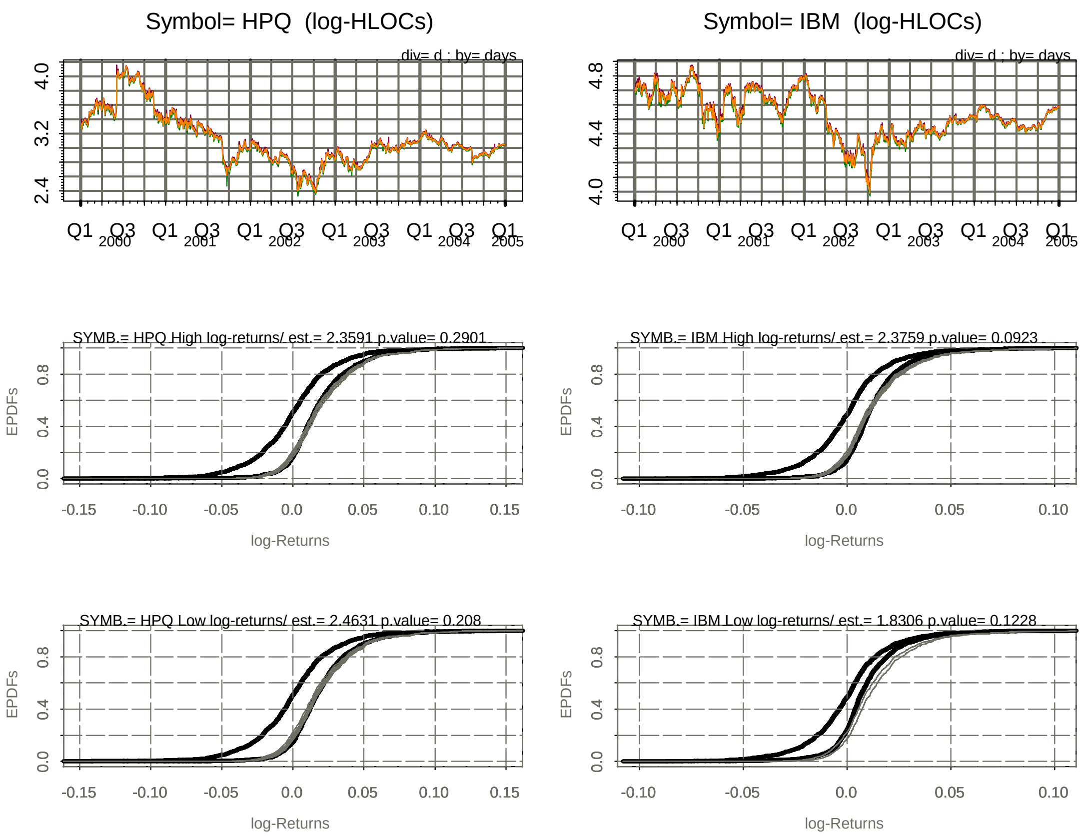

Figure 6-2 Close-High and Close-Low GoFs for HPQ and IBM log-returns

Symbol	High.est	High.ses	High.ks	High.p	Low.est	Low.ses	Low.ks	Low.p
AA	2.2777	0.0417	0.0752	***	2.4144	0.0468	0.0583	**
AXP	1.6443	0.0292	0.0499	*	1.5786	0.0305	0.0483	
BA	1.8333	0.0328	0.0599	**	2.3375	0.0399	0.0349	
C	2.4778	0.047	0.0641	**	1.6933	0.0309	0.0407	
CAT	1.8627	0.0336	0.0602	**	2.2627	0.0451	0.0559	**
DD	2.0355	0.0363	0.0456		2.7978	0.0529	0.0528	*
DIS	1.8161	0.0322	0.0628	**	2.5859	0.0454	0.0302	
EK	1.7412	0.0334	0.0514	*	2.5287	0.0507	0.0645	***
GE	2.464	0.0471	0.0583	**	2.6177	0.0487	0.051	*
GM	1.9135	0.0349	0.0556	**	2.3769	0.0466	0.0466	
HD	1.7553	0.031	0.064	**	1.9132	0.0359	0.043	
HON	1.864	0.0348	0.0591	**	2.508	0.0476	0.0553	**
HPQ	2.3582	0.0465	0.0402		2.4709	0.0481	0.0422	
IBM	2.3759	0.0434	0.049	*	1.8306	0.0344	0.0466	
INTC	2.4344	0.0487	0.0611	**	2.2937	0.0438	0.0594	**
IP	1.947	0.0373	0.0576	**	2.6862	0.0496	0.0386	
JNJ	1.7996	0.034	0.0474		2.2994	0.0427	0.0514	*
JPM	2.5092	0.0458	0.0483		1.6903	0.0293	0.0359	
KO	1.6876	0.0304	0.0499	*	2.1475	0.0389	0.0415	
MCD	2.5837	0.0538	0.0466		2.6687	0.0464	0.0422	
MMM	1.7916	0.0339	0.0582	**	2.4056	0.0432	0.0627	**
MO	1.6887	0.0311	0.0452		1.444	0.025	0.0582	**
MRK	2.3549	0.0458	0.0502	*	2.4572	0.0475	0.0472	
MSFT	1.9089	0.0346	0.061	**	2.4692	0.0425	0.0548	**
PG	1.4963	0.0282	0.0686	***	1.8651	0.0365	0.0589	**
SBC	1.9351	0.0343	0.0564	**	1.8632	0.0342	0.0454	
T	1.7711	0.0317	0.0786	***	2.7409	0.0521	0.0496	*
UTX	1.8064	0.0323	0.0612	**	1.6092	0.0307	0.0581	**
WMT	1.9409	0.0381	0.0625	**	1.9575	0.0364	0.0406	
XOM	2.4135	0.0424	0.0306		2.2692	0.0469	0.0252	
^DJI	4.6088	0.0881	0.0465		5.0953	0.105	0.0401	

Table 6-1: Close-High and Close-Low GoFs

The Close-High Close-Low order-parameters for the DJI30 stocks are found to fall in the range from *1.4963* to *2.5837* for the High log-returns and from *1.444* to *2.7409* for the Low log-returns. The index order-parameter is *4.6088* for the High log-returns and *5.0953* for the Low log-returns. Of High log-returns only *3* out of the *30* stocks considered could be rejected at the *1%* significance level. Of the Low log-returns only *1* out of the *30* stocks could be rejected at the *1%* significance level. These results can be found in Table 6-1 and is graphically summarized in Figure 6-3.

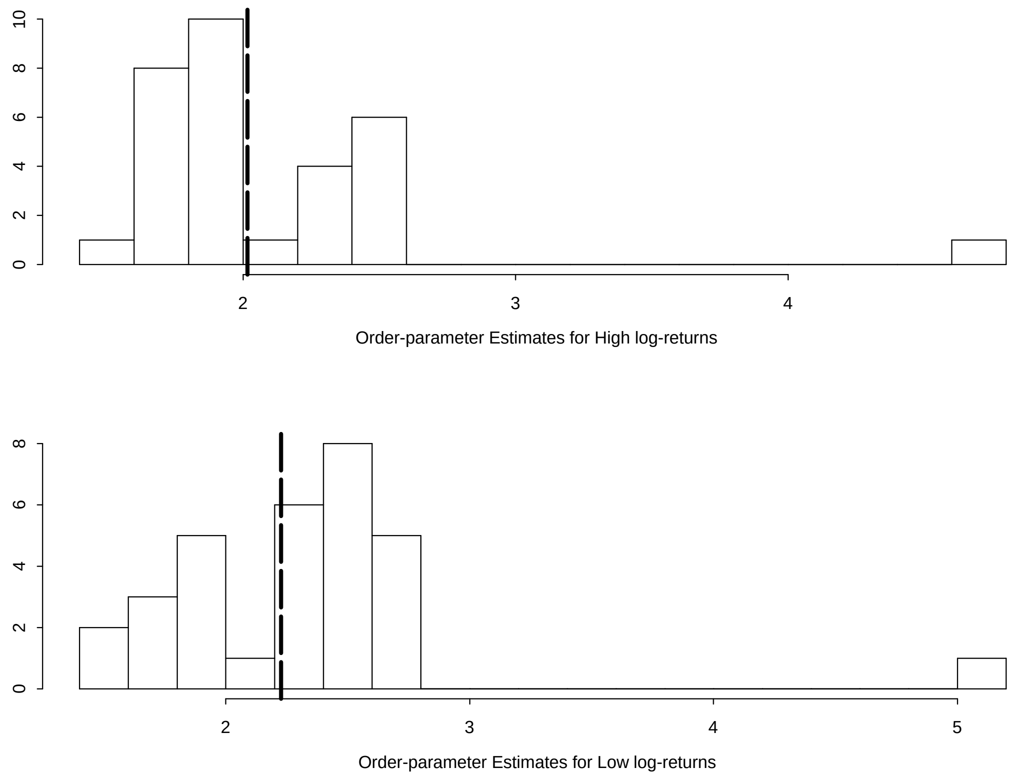

Figure 6-3 Histograms of Order-Parameters for Close-High and Close-Low Pairs

Figure 6-3 displays the histogram plots of the order-parameters for the set of DJI30 stocks and the DJI30 index considered. The order-parameters for the stocks are clustered together, whereas the order-parameters for the corresponding index log-returns tend to be “outliers”. Consequently, cross-sectional aggregation of stocks must be non-stable.

Next we compare the High and Low log-returns with each other to see if they also can be expressed a functions of each other. From Figure 6-4 we can deduce that the High and Low log-returns are mutually symmetric to each other, i.e. the order-parameters are very nearly identical can cannot be rejected at the 10% level. The estimated values of the index order-parameters are *0.9167* for the Low-High pair and *1.1008* for High-Low pair.

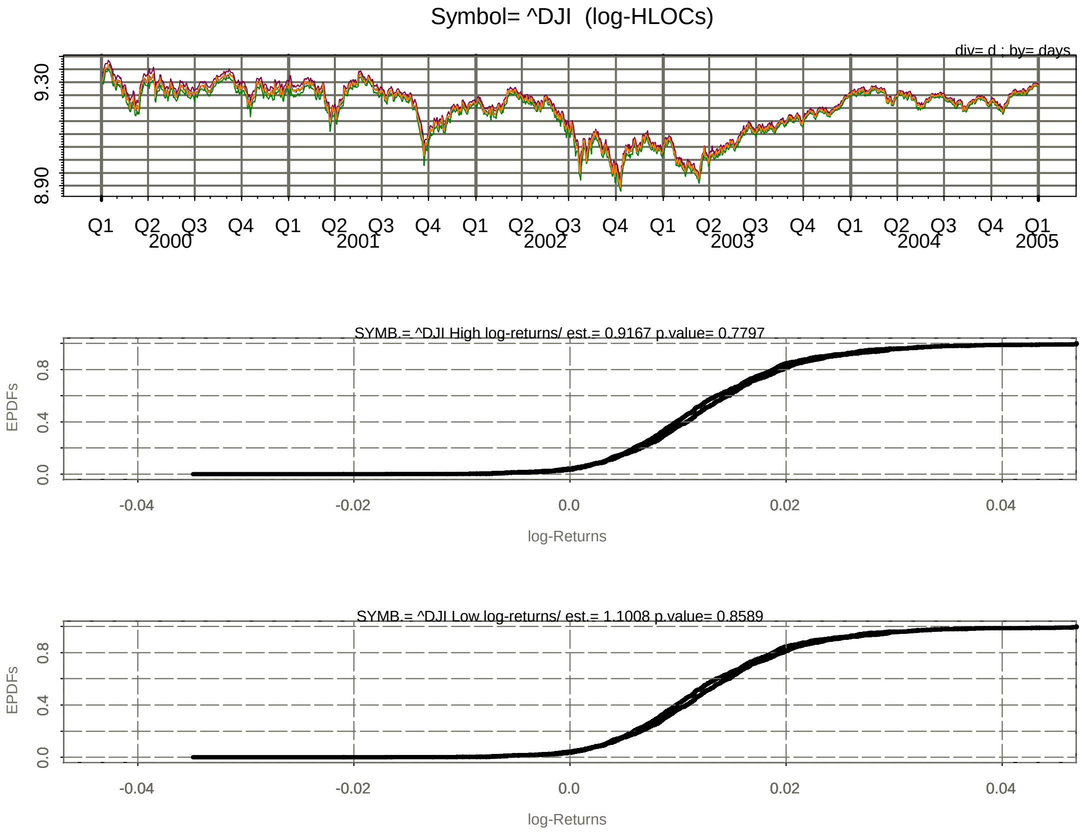

Figure 6-4 Low-High and High-Low GoFs for DJI Index log-returns

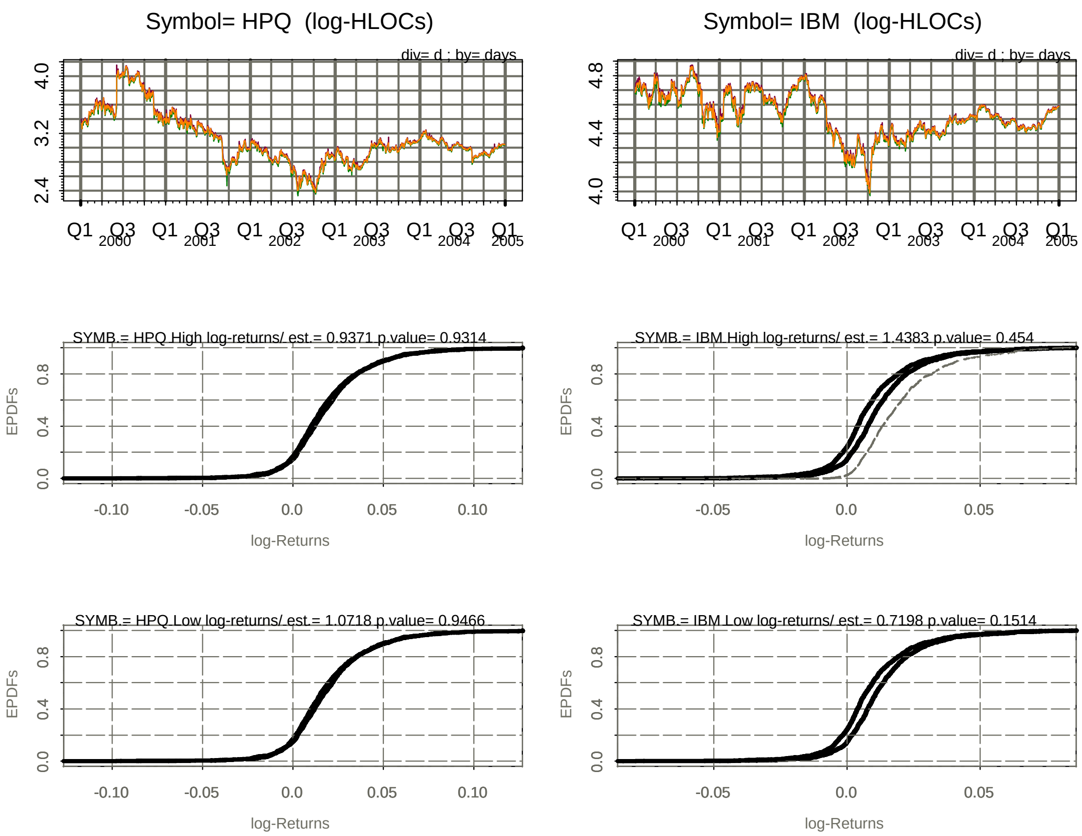

Figure 6-5 Low-High and High-Low GoFs for HPQ and IBM log-returns

Figure 6-5 shows the original ECDFs and the fitted ECDFs of the High and Lows for the two DJI30 stocks. In Table 6-2 the Low-High and High-Low order-parameters for the DJI30 stocks are found to fall in the range from *0.6191* to *1.5152* for the Low-High pairs and from *0.6459* to *1.7701* for the High-Low pairs. Of Low-High pairs only *2* out of the *30* stocks considered could be rejected at the *1%* significance level. Of the High-Low pairs none out of the *30* stocks could be rejected at the *1%* significance level.

Symbol	High.est	High.ses	High.ks	High.p	Low.est	Low.ses	Low.ks	Low.p
AA	0.9154	0.0163	0.0358		1.0964	0.0186	0.0358	
AXP	1.0457	0.0187	0.0246		0.9541	0.0174	0.0247	
BA	0.7811	0.0138	0.0457		1.3169	0.0262	0.036	
C	1.4617	0.0306	0.0481		0.6459	0.0141	0.0383	
CAT	0.7547	0.0139	0.04		1.3774	0.0278	0.0283	
DD	0.6889	0.0119	0.0471		1.4923	0.0303	0.0333	
DIS	0.6719	0.012	0.0573	**	1.6513	0.036	0.0429	
EK	0.6191	0.0114	0.0555	**	1.6663	0.0325	0.0531	*
GE	0.9192	0.0171	0.0331		1.0985	0.0218	0.0308	
GM	0.7759	0.0163	0.0325		1.3445	0.0279	0.0228	
HD	0.8963	0.016	0.0392		1.1286	0.0217	0.036	
HON	0.7099	0.0129	0.0713	***	1.5355	0.0316	0.0492	*
HPQ	0.9371	0.0176	0.0211		1.0718	0.0209	0.0204	
IBM	1.4383	0.0286	0.0337		0.7198	0.0126	0.0448	
INTC	1.0878	0.0195	0.0221		0.9177	0.0168	0.0213	
IP	0.6866	0.0127	0.0485	*	1.5415	0.0308	0.0322	
JNJ	0.7516	0.0152	0.0386		1.4205	0.033	0.0237	
JPM	1.5152	0.0294	0.0437		0.6595	0.0117	0.0451	
KO	0.7655	0.0144	0.0443		1.3548	0.0271	0.0319	
MCD	0.9401	0.0166	0.0326		1.0732	0.0246	0.0324	
MMM	0.7112	0.0135	0.0458		1.4665	0.0274	0.0329	
MO	1.1903	0.0254	0.0235		0.8453	0.015	0.0235	
MRK	0.9183	0.0156	0.0209		1.0912	0.0188	0.0206	
MSFT	0.7645	0.0146	0.036		1.3579	0.0283	0.0244	
PG	0.7666	0.015	0.0456		1.3373	0.0283	0.045	
SBC	1.0126	0.0179	0.0277		0.9871	0.0177	0.0279	
T	0.6127	0.0108	0.0707	***	1.7701	0.0362	0.0424	
UTX	1.105	0.0192	0.0207		0.9026	0.0156	0.0216	
WMT	0.9655	0.0181	0.0416		1.0408	0.0209	0.0411	
XOM	1.0614	0.0202	0.0299		0.9386	0.0185	0.0292	
^DJI	0.9167	0.017	0.0257		1.1008	0.0281	0.0236	

Table 6-2: Low-High and High-Low GoFs

It can also be noticed from Table 6-2 that the order estimates for the High to Low and the Low to High log-returns are inverses of each other.

This is expected from the relationship between the pairs defined as:

$$F_{n:n}(x)=F^{n}(x);\ and \quad F_{n:n}^{1/n}(x)=F(x) \tag{19}$$

The results indicate that the High and Low log-returns are similar but not identical. The "hidden" drivers of the Highs are similar to the Lows, but opposite in nature. The Low-High pairs are generally of a higher order than the High-Low pairs. This is depicted by the greater number of order values that are larger than 1.0 in the "Low.est" columns in Table 6-2.

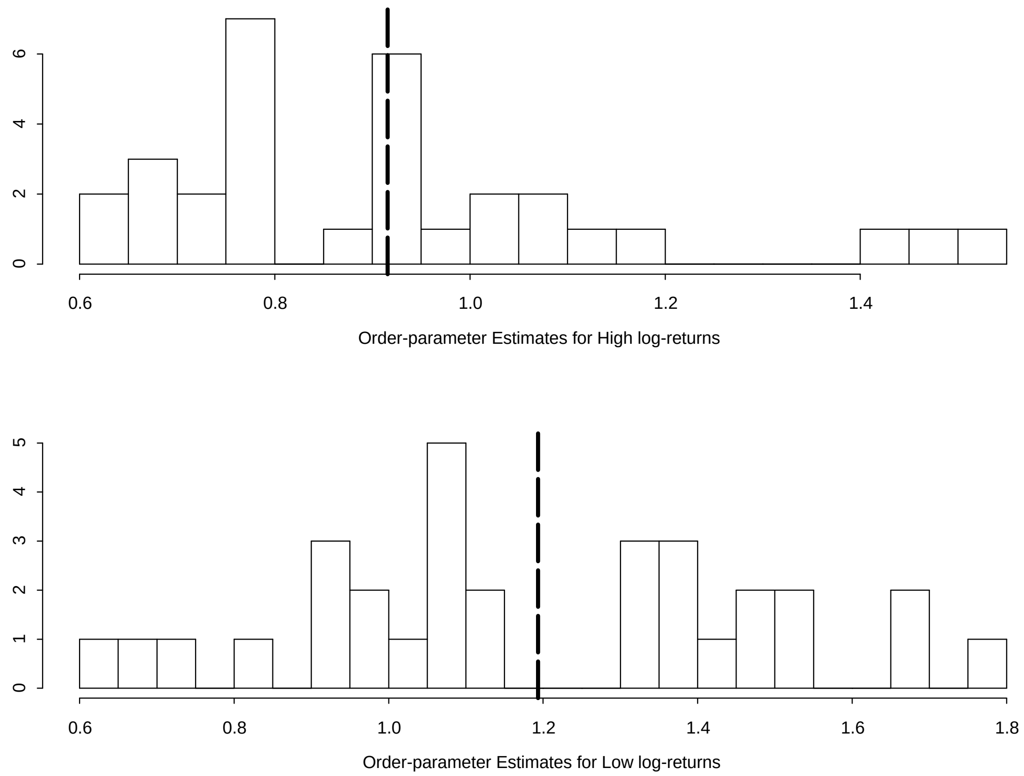

Figure 6-6 Histograms of Order-Parameters Low-High and High-Low Pairs

In Figure 6-6 the mean order-parameter for the Low-High pairs is less than 1.0, whereas the mean order-parameter for the High-Low pairs is more than 1.0. This indicates that the Low log-returns are more fat-tailed than the High log-returns. The "bust" effect is larger than the "boom" on average.

7 Conclusion

Exact extreme value theory (EVT) states that there is an exact relationship between the parent distribution and the extreme value distribution of random variates. This study investigates this possibility for the distributional relationship between the Close, High and Low price changes or returns. The empirical cumulative distribution function (ECDF) is used to depict the parent Close-to-Close distributions and the extreme-valued High-to-Close and Low-to-Close distributions.

The findings indicate that:

(i) there is an *exact* relationship between distributions of High and Low logarithmic returns and Close logarithmic returns;

(ii) the magnitude of the order-parameters are fairly consistent across the DJI30 stocks;

(iii) the magnitude of the order-parameters are fairly similar between High and Low logarithmic returns;

(iv) the cross-sectional aggregation of the extreme logarithmic returns are not stable as indicated by the significant difference in the magnitudes of the outlying order-parameters of the DJI30 index log-returns.

Implicit in the findings is the notion that the n-power of the transformed empirical distribution reflects the n-order of the trading horizon for daily returns. Given the parent empirical distribution, then the trading horizon implied by the results is greater than 1 and less than 3 days for the DJI30 stocks; i.e. if one desires to obtain the highest or the lowest daily price then one must be willing to has to trade over a 1-3 day horizon. The DJI30 index, however, has an order of 4-5 days, indicating that the aggregation is not stable across cross-sections.

Much research on exact distributions has been carried out to investigate the distribution of Close-to-Close returns. Analogously, it might be time now to investigate exact distributions for High-to-Close and Low-to-Close returns. Possibly, the exact distribution chosen for the *extremals* might help determine the appropriate exact model for the *centrals* (Close-to-Close returns) from amongst the numerous distributions investigated in finance to-date (see Bachelier [1900], Roberts [1959] ,Mandelbrot [1963], Cootner [1964], Brada, Ernst, et al. [1966], Officer [1972], Barnea and Downes [1973], Blattberg and Gonedes [1974], Upton and Shannon [1979], Smith [1981], Bookstaber and McDonald [1987], Akgiray and Booth [1988], Hall, Brorsen, et al. [1989], Gray and French [1990], Madan and Seneta [1990], Gribbin, Harris, et al. [1992], Aggarwal and Aggarwal [1993], Lux [1996], McDonald [1996], McCulloch [1996], Mauleon and Perote [1998] , Bingham and Kiesel [2001], Theodossiou [2001], Andreou, Pittis, et al. [2001], Aparicio and Estrada [2001], Harris and Kucukozmen [2001], Yu [2001], Knight and Satchell [2001], Bibby and Sorensen [2002], Gabaix, Gopikrishnan, et al. [2003], Shang and Tadikamalla [2004] and many others) by investigating whether the extreme value transforms of the proposed parent distributions are able to fit the extremes. This additional constraint on the extremals or the extreme-value distributions should narrow the range of distributional fits for the parent or Close-to-Close distributions of asset log-returns.

The findings shed new light on the statistical behaviour of financial asset returns, especially the High and Low log-returns. The multivariate relationships identified in this paper could be used to improve our understanding of the returns generation process of financial assets.

References

Aggarwal, R, and R Aggarwal, 1993. Security return distributions and market structure: Evidence from the NYSE/AMEX and the NASDAQ markets, *Journal of Financial Research*, 16, 209-23.

Akgiray, V, and G G Booth, 1988. The stable-law model of stock returns, *Journal of Business and Economic Statistics*, 6, 51-57.

Andreou, E, N Pittis, and A Spanos, 2001. On modelling speculative prices: The empirical literature, *Journal of Economic Surveys*, 15, 187-220.

Aparicio, F M, and J Estrada, 2001. Empirical distributions of stock returns: European securities markets,1990-95, *European Journal of Finance*, 7, 1-21.

Bachelier, L, 1900. Theory of speculation, in P H Cootner, ed., *The Random Character of Stock Market Prices* (Risk Publications, London).

Barnea, A, and D H Downes, 1973. A re-examination of the empirical distribution of stock price changes, *Journal of the American Statistical Association*, 168, 348-50.

Bibby, B M, and M Sorensen, 2002. Hyperbolic processes in finance, *Working Paper*, The Royal Veterinary and Agricultural University, Copenhagen.

Bingham, N H, and R Kiesel, 2001. Modelling asset returns with hyperbolic distributions, in J Knight, and S E Satchell, eds., *Return Distributions in Finance* (Butterworth-Heinemann, Oxford).

Blattberg, R C, and N J Gonedes, 1974. A comparison of stable and Student distributions as statistical models for stock prices, *Journal of Business*, 47, 244-80.

Bookstaber, R M, and J B McDonald, 1987. A general distribution for describing security returns, *Journal of Business*, 60, 401-24.

Brada, J, H Ernst, and J van Tassel, 1966. The distribution of stock price differences: Gaussian after all?, *Operations Research*, 14, 334-40.

Chakravart, Laha, and Roy, 1967. *Handbook of Methods of Applied Statistics: Volume I* (John Wiley).

Cochrane, D, and G H Orcutt, 1949. Application of Least Squares Regressions to Relationships Containing Austocorreclated Error Terms, *Journal of the American Statistical Association*, 44, 32-61.

Cootner, P H, 1964. *The Random Character of Stock Market Prices* (Risk Books, a division of Risk Publications, London).

Dennis, J E, D M Gay, and R E Welsch, 1981. An adaptive nonlinear least-squares algorithm, *ACM Transactions on Mathematical Software*, 7, 348-83.

Dennis, J E, and H H W Mei, 1979. Two new unconstrained optimization algorithms which use function and gradient values, *Journal of Optimization Theory and Applications*, 28, 453-83.

Fiess, N M, and R MacDonald, 2002. Towards the fundamentals of technical analysis: Analysing the information content of High, Low and Close prices, *Economic Modelling*, 19, 353-74.

Gabaix, X, P Gopikrishnan, V Plerou, and H E Stanley, 2003. A theory of power-law distributions in financial market fluctuations, *Nature*, 423, 267.

Gibbons, J D, and S Chakraborti, 1992. *Nonparametric Statistical Inference* (Marcel Dekker, New York).

Gray, J B, and D W French, 1990. Empirical comparisons of distributional models for stock index returns, *Journal of Business Finance and Accounting*, 17, 451-59.

Gribbin, D W, R W Harris, and H S Lau, 1992. Futures Prices Are Not Stable-Paretian Distributed, *Journal of Futures Markets*, 12, 475-88.

Gumbel, E J, 1958. *Statistics of Extremes* (Columbia University Press, New York).

Hall, J A, B W Brorsen, and S H Irwin, 1989. The distribution of futures prices: A test of stable Paretian and mixture of normals hypotheses, *Journal of Financial and Quantitative Analysis*, 24, 105-16.

Harris, D F R, and C C Kucukozmen, 2001. The empirical distribution of UK and US stock returns, *Journal of Business Finance and Accounting*, 28, 715-40.

Hollander, M, and D A Wolfe, 1999. *Nonparametric Statistical Methods* (Wiley, New York).

Kim, P J, and R I Jennrich, 1970. Tables of the exact distribution of the two sample Kolmogorov-Smirnov criterion, in H L Harter, and D B Owen, eds., *Selected tables in Mathematical Statistics* (American Mathematical Society, Providence, Rhode Island).

Knight, J, and S E Satchell, 2001. *Return Distributions in Finance* (Butterworth-Heinemann, Oxford).

Kolmogorov, A N, 1933. Sulla Determinazione Empirica di una Legge di Distribuzione, *Giornale del Instituto Italiano degli Attuari*, 4, 83-91.

Longin, F, 1996. The asymptotic distribution of extreme stockmarket returns, *Journal of Business*, 69, 383-408.

Los, C A, 2003. *Financial Market Risk: Measurement and Analysis* (Routledge, London).

Lux, T, 1996. The stable Paretian hypothesis and the frequency of large returns: An examination of major German stocks, *Applied Financial Economics*, 6, 463-73.

Madan, D B, and E Seneta, 1990. The variance gamma model for share market returns, *Journal of Business*, 63, 511-24.

Mandelbrot, B, 1963. The variation of certain speculative prices, *Journal of Business*, 36, 394-419.

Mauleon, I, and J Perote, 1998. Testing densities with financial data: An empirical comparison of the Edgeworth-Sargan density to the Student-t, *Nouveaux Instrumentset Marches Financiers Emergents*, Paris (Organised by AEA).

McCulloch, J H, 1996. Financial applications of stable distributions, in G S Maddala, ed., *Handbook of Statistics, 14* (Elsevier Science).

McDonald, J B, 1996. Probability distributions for financial models, in G S Maddala, and C R Rao, eds., *Handbook of Statistics, 14* (Elsevier Science).

Officer, R R, 1972. Distribution of stock prices, *Journal of the American Statistical Association*, 67, 807-12.

Roberts, H V, 1959. Stock market patterns and financial analysis: Methodological suggestions, *Journal of Finance*, 14, 1-10.

Shang, J S, and P R Tadikamalla, 2004. Modelling financial series distributions: A versatile data fitting approach, *International Journal of Theoretical and Applied Finance*, 7, 231-51.

Smith, J, 1981. *The Probability Distribution of Market Returns: A Logistic Hypothesis,* Unpublished PhD Thesis (University of Utah).

Theodossiou, P, 2001. Skewed generalised error distribution of financial assets and option pricing, *Working Paper*, School of Business, Rutgers University, New Jersey.

Upton, D, and D Shannon, 1979. The stable Paretian distribution, subordinated stochastic processes and asymptotic lognormality: An empirical investigation, *Journal of Finance*, 1031-39.

Wegman, E J, 1972. Nonparametric probability density estimation, *Technometrics*, 14, 533-46.

Yu, J, 2001. Testing for a finite variance in stock return distributions, in J Knight, and S E Satchell, eds., *Return Distributions in Finance* (Butterworth-Heinemann, Oxford).

Economics Discussion Papers

2005-01 Investment and Savings Cycles and Tests for Capital Market Integration, **Arusha Cooray** and **Bruce Felmingham**

2005-02 The Efficiency of Emerging Stock Markets: Empirical Evidence from the South Asian Region, **Arusha Cooray** and **Guneratne Wickremasinghe**

2005-03 Error-Correction Relationships Between High, Low and Consensus Prices, **Nagaratnam Jeyasreedharan**

2005-04 Tests for RIP Among the G7 When Structural Breaks are Accommodated, **Bruce Felmingham** and **Arusha Cooray**

2005-05 Alternative Approaches to Measuring Temporal Changes in Poverty with Application to India, **Dipankor Coondoo, Amita Majumder, Geoffrey Lancaster** and **Ranjan Ray**

2005-06 Intertemporal Household Demographic Models for Cross Sectional Data, **Paul Blacklow**

2005-07 Some Recent Evidences about the Global Integration of Chinese Share Markets, **Yong Hong Yan** and **Bruce Felmingham**

2005-08 Managerial Objectives and Use Limits on Resource-Based Recreation, **Hugh Sibly**

2005-09 The Feldstein-Horioka Model Re-Visited for African Countries, **Arusha Cooray** and **Dipendra Sinha**

2005-10 Analysis of Changes in Food Consumption and Their Implications for Food Security and Undernourishment: The Indian Experience in the 1990s, **Ranjan Ray**

2005-11 An Extended Feldstein-Horioka Test for the Degree of Capital Mobility, **Alexis Wadsley, Bruce Felmingham** and **Arusha Cooray**

2005-12 Extreme-Valued Distributional Relationships in Asset Prices, **Nagaratnam Jeyasreedharan**

2004-01 Parametric and Non Parametric Tests for RIP Among the G7 Nations, **Bruce Felmingham** and **Arusha Cooray**

2004-02 Population Neutralism: A Test for Australia and its Regions, **Bruce Felmingham, Natalie Jackson** and **Kate Weidmann.**

2004-03 Child Labour in Asia: A Survey of the Principal Empirical Evidence in Selected Asian Countries with a Focus on Policy, **Ranjan Ray –** published in *Asian-Pacific Economic Literature*, 18(2), 1018, November 2004.

2004-04 Endogenous Intra Household Balance of Power and its Impact on Expenditure Patterns: Evidence from India, **Geoffrey Lancaster, Pushkar Maitra** and **Ranjan Ray**

2004-05 Gender Bias in Nutrient Intake: Evidence From Selected Indian States, **Geoffrey Lancaster, Pushkar Maitra** and **Ranjan Ray**

2004-06 A Vector Error Correction Model (VECM) of Stockmarket Returns, **Nagaratnam J Sreedharan**

2004-07 Ramsey Prices and Qualities, **Hugh Sibly**

2004-08 First and Second Order Instability of the Shanghai and Shenzhen Share Price Indices**, Yan, Yong Hong**

2004-09 On Setting the Poverty Line Based on Estimated Nutrient Prices With Application to the Socially Disadvantaged Groups in India During the Reforms Period, **Ranjan Ray** and **Geoffrey Lancaster –** published in *Economic and Political Weekly, XL(1), 46-56, January 2005.*

2004-10 Derivation of Nutrient Prices from Household level Food Expenditure Data: Methodology and Applications, **Dipankor Coondoo**, **Amita Majumder**, **Geoffrey Lancaster** and **Ranjan Ray**

Copies of the above mentioned papers and a list of previous years' papers are available on request from the Discussion Paper Coordinator, School of Economics, University of Tasmania, Private Bag 85, Hobart, Tasmania 7001, Australia. Alternatively they can be downloaded from our home site at http://www.utas.edu.au/ecofin