

Daily Stock Returns, Non-Normality and Hypothesis Testing

GEORGE S. FORD

Daily stock returns typically have non-normal and asymmetric distributions, potentially leading to problems with hypothesis testing based on reported probability statistics from regression analysis (Fama 1976; Brooks 2002). While daily stock return data for many years is readily available, recent studies indicate that the non-normality problems may persist even in large samples (Peiro 2002; Brooks 2002; Bai et al 2002; Corrado and Zivney 1992). Given the frequent use of daily return data with standard estimation methods such as least squares, the practical importance of non-normality on traditional hypothesis testing is of some significance.

One statistical application employing daily return data is the dummy variable version of the event study, which is commonly estimated by least squares (MacKinlay 1997). In such studies, statistical significance of abnormal returns is measured frequently using t-tests and F-tests based on the estimated standard errors of the regression (MacKinlay 1997; Fisman 2001, Simpson 2001, and many others). If the error term is non-normal, however, appealing to the reported test-statistics may be problematic.

The magnitude of this problem is assessed in this paper by comparing the empirical critical statistics – estimated by bootstrap – to their asymptotic analogues. These empirically determined critical values also are compared those estimated using a Monte Carlo simulation under the assumption that the error term is normally distributed, but otherwise identical to the actual error term of the regression.¹

An analysis of a year's trading activity for thirty stocks from the S&P 500 Index suggests that asymptotic critical values are not reliable for hypothesis testing from regression analysis using daily stock returns. Statistical significance of events should be based on alternative criteria or by non-parametric methods.

I. Methodology

The empirical analysis begins with a sample of 30 publicly traded stocks from the S&P 500 index. Prices for all trading days of year 2002 are employed, rendering 251 daily returns². Two events are defined based on the largest positive (*DP*) and largest negative (*DN*) abnormal returns for the year, and each event is assigned a dummy variable.³

For each stock, the disturbances (i.e., abnormal returns) are estimated using the market model including the event dummy variables:

$$(1) \quad R_{i,t} = \alpha_i + \beta_i M_t + \gamma_i DP_i + \phi_i DN_i + \varepsilon_{i,t}$$

where R_t is the return of daily return of stock i at time t , M_t is the daily return on the market index at time t (S&P 500), α , β , γ and ϕ are the estimated coefficients for stock i , and ε_t is the econometric disturbance term with mean zero and variance s^2 . The Shapiro-Wilk, Kolmogorov (Lilliefors) and Jarque-Bera test reject normality of ε for all stocks in the sample. The disturbances are homoscedastic (via White's test, which is asymptotically valid for non-normal errors) and exhibit no serial correlation (via Durbin Watson test; $d_U = 1.82$). Thus, deviations of the empirical test statistics from the asymptotic values derive solely from the non-normality and asymmetry of the disturbance, rather than these other problems (Salinger 1992; MacKinlay 2002).

The empirical distributions of the test statistics are estimated using the following model (suppressing the t and i subscripts):

² Daily return data are from finance.yahoo.com.

³ Some simulation studies create events in stock data by adding (say) 1% to the abnormal return for a defined period. An interest in the symmetry properties of the distribution makes this approach undesirable.

¹ That is, both have a common mean and variance.

$$(2) \quad \hat{R} + \varepsilon^b = \alpha^b + \beta^b M + \gamma^b DP + \phi^b DN + \mu^b$$

where \hat{R} is the predicted dependent variable of (1) and ε^b is bootstrapped from ε (a random draw, with replacement, from ε). Each disturbance ε is multiplied by $[n/(n-k)]^{0.5}$ (MacKinnon 2002). This same approach is used for the Monte Carlo simulations, although the dependent variable of (2) is computed using random draws, with replacement, from a normally distributed random series with the same mean and variance as ε .

A new dependent variable is created and Equation (2) is estimated 999 times for each stock in the sample. This process creates a distribution of 999 statistics, each one computed as

$$\frac{\gamma^b - \gamma}{se_{\gamma}^b} \quad \text{and} \quad \frac{\phi^b - \phi}{se_{\phi}^b},$$

where se^b is the standard error of the estimated coefficient from (2). These 999 statistics are then sorted, so that the 5% critical values can be selected as the $(0.05/2)(999 + 1)$ and $(1 - .05/2)(999 + 1)$ values of the series (the 25th and 975th values of the series). The same calculations are made for the Monte Carlo simulation.

Because the boots trap and Monte Carlo procedures are empirical, both will generate unique critical values for each of the iterations. Thus, average critical values are computed from 100 repetitions of 999 iterations producing a distribution of the critical values. The average critical value, therefore, is based on 99,900 regressions. Descriptive statistics for the stock returns and regression disturbances are provided in Table 1.

II. Results

Table 2 summarizes the bootstrap and Monte Carlo test statistics based on the average of the 100 computed test statistics. The asymptotic critical value with 247 degrees of freedom is 1.97 and Monte Carlo simulation produces an average critical value of 1.978 across all 120 simulated critical t-statistics, so the simulated critical values based on the normal distribution are compatible with the asymptotic values.

Unlike the Monte Carlo statistics, the critical values from the bootstrap are highly variable and typically depart non-trivially from the asymptotic values⁴. For the negative critical statistics, the average from the bootstrap is about -2.11. For the positive critical values, the average is 2.02, which is much closer to the asymptotic value. On average, the null-hypothesis will be too frequently rejected when using daily stock returns and asymptotic critical values, but over rejection is not always the problem.

Looking at the average critical value across the individual stocks masks a number of interesting facts. First, 41 of the 120 critical values (about one-third) fall below the asymptotic critical value, indicating that relying on asymptotics may, in many cases, lead to an under-rejection of the null for individual stocks.

Second, the largest (in absolute terms) bootstrapped critical value is about -2.75, which represents a non-trivial departure from the asymptotic analogue. Nevertheless, a researcher can be reasonably certain that the non-normality issue will not implicate statistical significance for t-statistics exceeding (in absolute terms) about 2.75. The smallest critical value is about 1.65, also a meaningful departure from the asymptotic counterpart. If statistical significance is absent using standard asymptotic interpretations of results, the bootstrap may be a preferred method for assessing statistical significance.

Third, the bootstrapped critical values are about twice as variable as the Monte Carlo values. The volatility of bootstrapped critical values may render quite different conclusions about statistical significance depending on the unique outcome of a particular bootstrap. A potential danger of bootstrapping critical values is a "critical value search," where a researcher estimates multiple critical values by bootstrap until a statistically significant result is achieved. The averaging approach employed here may be a suitable response to this potential.⁵

⁴ Means different tests reject equality (5% level) of the bootstrap and Monte Carlo test statistics for 94% of the critical values (113 of 120).

⁵ Performing 99,900 regressions is computationally expensive (the bootstrap took about 30 minutes per stock on a 750Mhz machine), but the specification of the event study

Fourth, While the Monte Carlo critical statistics are generally symmetric (-1.977: 1.978), the bootstrap statistics are not. With the Monte Carlo statistics, the largest deviation from symmetry, measured as the ratio of the positive and negative critical values, is just under 1.02. The mean of the ratio is 1.007. With the bootstrapped values, however, the largest discrepancy is 1.56. On average, the ratio is 1.14. Thus, non-normality has a meaningful impact on the symmetry of the critical values.

III. Conclusion

A common form of the event study in financial analysis is the dummy-variable regression model, where statistical significance of events is determined using the asymptotic critical values. Daily stock returns, however, generally have non-normal distributions even in large samples, suggesting an appeal to asymptotics may be inappropriate. Comparing bootstrapped critical values to their asymptotic counterparts for thirty stocks reveals that asymptotic values are frequently unreliable determinants of statistical significance (or confidence intervals using reported standard errors). When combining daily stock returns and regression analysis, statistical significance of events should be determined by alternative methods, perhaps by non-parametric methods such as the bootstrap (Brock et al 1992).

REFERENCES

- Bai, Xuezheng, et al., *Beyond Merton's Utopia (I)*, (Univ. of Chi. Working Paper, 2002).
- Brock, W., J. Lakonishock and B. LeBaron. *Simple Technical Trading Rules and the Stochastic Properties of Stock Returns*, 47 *Journal of Finance* 1731 (1992).
- Brooks, C.. *Introductory Econometrics for Finance* (2002).
- Corrado, C. J. and Terry L. Zivney. *The Specification and Power of the Sign Test in Event Study Hypothesis Tests Using Daily Stock Returns*, 27 *The Journal of Financial and Quantitative Analysis* 465 (1992).
- Fama, E. F., *Foundations of Finance* 38 (1976).
- Fisman, R. *Estimating the Value of Political Connections*, 91 *American Economic Review* 1095 (2001).
- MacKinnon, J. G. *Bootstrap Inference in Econometrics*, Paper Presented as the Presidential Address at the 2002 Annual Meeting of the Canadian Econometrics Association (Revised June 2002).
- MacKinlay, A. G. *Event Studies in Economics and Finance*, XXXV *Journal of Economic Literature* 13-39 (1997).
- Peiro, Amado, *Skewness in Individual Stocks at Different Investment Horizons*, 2 *Quantitative Finance* 139 (2002).
- Simpson, J. D.. *Did May Company's Acquisition of Associated Dry Goods Corporation Reduce Competition? An Event Study Analysis*, 19 *Review of Industrial Organization* 351 (1991).

is well defined so that specification search is not generally an issue.

Table 1. Descriptive Statistics of e_t

Stock Ticker	St. Dev. (s)	Skewness	Kurtosis	Jarque-Bera Stat	Lilliefors Stat	Shapiro Wilk	White Test	DW Test
PFE	0.018	-0.990	7.630	265.2	0.241	5.314	0.960	2.120
EP	0.063	-0.324	10.74	630.4	1.620	7.548	0.800	1.871
HON	0.024	-1.533	14.27	1426.6	0.814	7.183	2.300	2.248
RSH	0.023	-1.829	16.27	1981.0	0.431	7.096	1.170	1.887
GCI	0.014	0.407	6.079	106.1	0.157	4.115	2.082	1.995
WIN	0.025	-0.009	4.333	18.60	0.296	3.205	4.140	2.040
ADM	0.015	-0.892	7.062	205.8	0.246	5.003	1.300	2.097
BMC	0.028	-0.147	4.450	22.90	0.222	3.132	2.460	1.975
SBL	0.046	-1.160	10.14	589.7	0.940	6.914	1.240	1.909
WAG	0.017	-0.684	6.450	144.0	0.298	5.034	0.640	2.248
VZ	0.021	0.429	6.055	105.3	0.292	4.517	3.810	2.124
BA	0.020	-0.013	3.970	9.900	0.154	2.013	0.890	2.029
MCD	0.021	-1.114	8.729	395.2	0.797	6.430	0.730	2.108
RJR	0.022	-1.937	11.49	910.5	1.293	7.76	3.240	2.095
GPS	0.032	-0.599	6.729	160.4	0.664	5.396	1.290	2.057
NFB	0.012	-0.505	5.201	61.30	0.396	4.692	2.830	2.049
Q	0.089	0.163	14.68	1426.8	1.747	7.929	2.710	2.027
AN	0.027	-0.035	6.136	102.9	0.270	4.424	1.930	2.419
AVP	0.015	0.920	6.590	170.2	0.276	5.178	0.470	1.994
CA	0.040	-0.405	9.694	475.5	0.651	6.641	0.440	2.009
LTR	0.014	-0.192	4.436	23.10	0.296	3.749	4.450	2.093
TXN	0.028	-0.528	7.550	228.2	0.156	4.832	4.480	1.887
AMGN	0.027	0.840	7.387	230.8	0.426	5.697	0.590	2.167
ERTS	0.023	-0.368	4.396	26.00	0.140	3.100	2.980	2.321
SPLS	0.025	0.504	6.717	155.2	0.243	4.607	0.150	2.141
FISV	0.022	-0.829	7.367	228.2	0.267	5.291	0.710	2.013
ACE	0.025	-0.108	4.585	26.80	0.215	2.935	3.550	2.160
S	0.032	-5.114	55.67	30106.4	1.905	9.458	1.800	1.945
PCS	0.067	0.020	4.942	39.50	0.311	4.039	1.290	1.834
FBF	0.019	-0.490	8.188	291.5	0.511	6.072	2.120	2.010

Critical Values: Jarque-Bera = ; Kolmogorov (Lilliefors) = ; Shapiro-Wilk = ; White $\chi^2 = 4.63$; DW $d_U = 1.82$.

Table 2. Summary of Results for Monte Carlo Simulation

Stock Ticker	Positive Return (γ)				Negative Return (ϕ)			
	Lower Mean	Lower St. Dev	Upper Mean	Upper St. Dev.	Lower Mean	Lower St. Dev	Upper Mean	Upper St. Dev.
PFE	-1.975	0.084	1.989	0.081	-1.982	0.097	1.962	0.067
EP	-1.970	0.087	1.986	0.080	-1.973	0.096	1.997	0.076
HON	-1.974	0.079	1.985	0.093	-1.972	0.094	1.975	0.095
RSH	-1.974	0.094	1.997	0.092	-1.962	0.094	1.968	0.075
GCI	-1.967	0.087	1.969	0.080	-1.963	0.085	1.982	0.076
WIN	-1.972	0.093	1.983	0.089	-1.977	0.087	1.976	0.083
ADM	-1.960	0.072	1.991	0.090	-1.966	0.084	1.986	0.097
BMC	-1.985	0.089	1.962	0.094	-1.968	0.085	1.973	0.092
SBL	-1.991	0.089	1.971	0.084	-1.969	0.078	1.984	0.069
WAG	-1.977	0.102	1.956	0.080	-1.992	0.099	1.964	0.090
VZ	-1.978	0.080	1.980	0.074	-1.979	0.088	1.969	0.089
BA	-1.981	0.088	1.976	0.087	-1.980	0.088	1.983	0.096
MCD	-1.971	0.084	2.000	0.081	-1.982	0.084	1.986	0.079
RJR	-1.972	0.083	1.980	0.085	-1.991	0.090	1.972	0.078
GPS	-1.971	0.089	1.993	0.086	-1.980	0.092	1.958	0.089
NFB	-1.973	0.090	1.975	0.092	-1.982	0.087	1.986	0.074
Q	-1.956	0.087	1.974	0.086	-1.973	0.085	1.982	0.086
AN	-1.980	0.090	1.977	0.082	-1.975	0.086	1.963	0.086
AVP	-1.998	0.088	1.972	0.080	-1.983	0.093	1.986	0.085
CA	-1.983	0.087	1.995	0.090	-1.959	0.086	1.976	0.073
LTR	-1.957	0.079	1.989	0.078	-1.986	0.082	1.980	0.092
TXN	-1.985	0.086	1.977	0.086	-1.986	0.077	1.978	0.083
AMGN	-1.976	0.090	1.980	0.086	-1.992	0.080	1.971	0.085
ERTS	-1.971	0.091	1.977	0.089	-1.982	0.075	1.975	0.082
SPLS	-1.968	0.086	1.980	0.106	-1.975	0.075	1.976	0.094
FISV	-1.976	0.091	1.977	0.089	-1.991	0.078	1.961	0.080
ACE	-1.967	0.085	1.972	0.083	-2.005	0.090	1.978	0.080
S	-1.967	0.079	1.993	0.087	-1.978	0.089	1.976	0.080
PCS	-1.966	0.071	1.994	0.079	-1.990	0.093	1.979	0.086
FBF	-1.966	0.082	1.973	0.109	-1.988	0.088	1.971	0.086
Average	-1.974	0.086	1.981	0.087	-1.979	0.087	1.976	0.083
Maximum	-1.956		2.000		-1.959		1.997	
Minimum	-1.998		1.956		-2.005		1.958	
St. Dev.	0.009		0.010		0.010		0.009	

Table 3. Summary of Results for Bootstrap

Stock Ticker	Positive Return (γ)				Negative Return (ϕ)			
	Lower Mean	Lower St. Dev	Upper Mean	Upper St. Dev.	Lower Mean	Lower St. Dev	Upper Mean	Upper St. Dev.
PFE	-2.266	0.25	1.814	0.056	-2.300	0.263	1.818	0.057
EP	-2.492	0.222	2.005	0.141	-2.462	0.211	2.022	0.154
HON	-1.934	0.137	2.142	0.127	-1.919	0.148	2.139	0.132
RSH	-1.930	0.155	1.998	0.141	-1.932	0.170	1.985	0.142
GCI	-2.040	0.127	1.850	0.100	-2.026	0.104	1.871	0.099
WIN	-2.090	0.133	2.297	0.070	-2.090	0.129	2.288	0.070
ADM	-2.247	0.205	1.869	0.063	-2.194	0.255	1.859	0.054
BMC	-2.257	0.182	1.958	0.120	-2.241	0.219	1.986	0.128
SBL	-2.120	0.364	1.976	0.126	-2.056	0.302	1.984	0.131
WAG	-1.953	0.127	2.020	0.055	-1.965	0.145	2.017	0.050
VZ	-1.906	0.104	2.050	0.116	-1.901	0.088	2.026	0.119
BA	-2.075	0.139	2.219	0.122	-2.075	0.145	2.223	0.133
MCD	-2.320	0.229	1.877	0.143	-2.350	0.240	1.873	0.141
RJR	-2.751	0.455	1.762	0.163	-2.699	0.403	1.796	0.146
GPS	-2.300	0.206	1.988	0.095	-2.298	0.192	1.982	0.068
NFB	-2.276	0.234	2.168	0.160	-2.261	0.222	2.171	0.185
Q	-1.697	0.116	2.307	0.270	-1.693	0.087	2.390	0.308
AN	-2.041	0.162	2.056	0.093	-2.029	0.174	2.057	0.094
AVP	-1.855	0.104	2.320	0.245	-1.864	0.089	2.245	0.237
CA	-2.003	0.253	1.905	0.162	-1.987	0.274	1.895	0.161
LTR	-2.363	0.198	2.121	0.188	-2.356	0.233	2.123	0.184
TXN	-1.904	0.134	1.897	0.080	-1.900	0.161	1.894	0.070
AMGN	-1.837	0.101	2.364	0.176	-1.855	0.101	2.368	0.190
ERTS	-2.173	0.173	1.706	0.071	-2.165	0.178	1.685	0.047
SPLS	-2.135	0.122	2.075	0.095	-2.141	0.114	2.074	0.110
FISV	-2.090	0.176	1.972	0.137	-2.114	0.169	1.988	0.131
ACE	-2.049	0.180	2.068	0.154	-2.032	0.180	2.078	0.145
S	-1.762	0.156	1.788	0.133	-1.764	0.187	1.781	0.137
PCS	-2.098	0.176	2.359	0.124	-2.084	0.168	2.388	0.130
FBF	-2.377	0.266	1.762	0.106	-2.381	0.274	1.762	0.124
Average	-2.111	0.186	2.023	0.128	-2.104	0.187	2.026	0.129
Maximum	-1.697		2.364		-1.693		2.390	
Minimum	-2.751		1.706		-2.699		1.685	
St. Dev.	0.227		0.187		0.221		0.19	