

Forecasting Financial Market Annual Performance Measures: Further Evidence⁺

¹Edward J. Lusk, ²Michael Halperin, ³Atanas Tetikov and ³Niya Stefanova

¹Department of Statistics, The Wharton School, University of Pennsylvania, Philadelphia, PA, USA

²Lippincott Library of the Wharton School, University of Pennsylvania, Philadelphia, PA, 19014, USA

³Faculty of Economics and Management, University of Magdeburg, 39106 Magdeburg, Germany

Abstract: Problem statement: Forecasting is simple; producing accurate forecasts is the essential task. Experience suggests that financial managers often assume that because models used in forecasting are appropriate that they are effective. This study addresses this assumption. Effective is taken to mean forecasts where the Absolute Percentage Error (APE) is equal to or less than 10%. It has been reported that forecasts of the CAPM- β using the Bloomberg heuristic did not provide effective forecasts. We were interested to determine if the lack of forecasting accuracy is peculiar to β or is more pervasive. **Approach:** We expanded the analysis to include three measures of Excess Market Return: Jensen's α ($J\alpha$), the Sharpe Performance Index (SPI) and the Treynor Performance Index (TPI) and two measures of market risk: we once again consider β and also a measure of non-market risk called idiosyncratic Risk (iR). We used information on 58 firms continuously traded on the NYSE or the NASDAQ from 1980 to and including 2008 to evaluate the effectiveness of forecasts of: $J\alpha$, SPI, TPI, β and iR. **Results:** Using Exponential Smoothing or (1,0,0) ARIMA models, we found no evidence that effective forecasts of these five market measures can be derived from such forecasting models. **Conclusion/Recommendations:** The important implication is: Financial Managers should be aware that even though they are they are using appropriate models to generate forecasts of $J\alpha$, SPI, TPI, β and iR that is no guarantee that such forecasts are effective. Finally, the authors' results are posted on Scholarly Commons.

Key words: Jensen's α , SPI, TPI, β , Ben-Horim/Levy non-market risk, data modification

INTRODUCTION

Resources devoted to the collection, verification and dissemination of firm market performance data are considerable. Data sources such as:

- BloombergTM [www.bloomberg.com]
- Dow-JonesTM [www.dowjones.com]
- Interactive DataTM [www.interactivedata.com]
- MorningstarTM [www.Morningstar.com]
- StandardandPoorsTM [www.standardandpoors.com]
- Thomson-ReutersTM [www.thomsonreuters.com]
- Wharton Research Data Service WRDSTM [wrds.wharton.upenn.edu]

Are costly but have survived, even in the recently challenging economic environment, because they provide reliable sources of firm performance data. Financial managers who provide fee for service investment advice depend on such data services to better understand the past so as to forecast, where possible, performance in the future. We use the

conditional: "where possible" because there are limits to forecasting. For example, consider forecasting stock prices or return; despite the existence of reliable historical data and the obvious economic advantage of being able to forecast the next day's price or return, research over almost a century leads one to the conclusion that the random component of daily stock seems to dominate this data generation thus rendering attempts to forecast daily stock activity inefficient and so an unwise use of resources because the forecasts have been shown to be ineffective. (For an excellent reading compendium on this subject area see (Cootner, 1964) which has the first research on the forecasting of stock prices. Also (Pelaez, 2003) is an excellent source and more recently (Abbondante, 2010; Lusk *et al.*, 2010).

Regarding forecasting in general, experience suggests that often, individuals believe that (1) if the assumptions of the forecasting model are likely to be satisfied i.e., they have reasonable assurance that the forecasting model is appropriate, or (2) if research shows that a particular forecasting model outperforms

Corresponding Author: Edward J. Lusk, Department of Accounting, The State University of New York (SUNY), College at Plattsburgh, School of Business and Economics, Plattsburgh, NY, USA

other reasonable models choices based upon one of the usual evaluation measures of error, then they believe that it follows that the forecasts will be useful. This confuses utility in use, which is an effectiveness issue, with the “best” forecasts in a statistical sense which is a model appropriateness issue. It is the utility in use or the effectiveness of forecasting information that financial managers must first consider so as to rationalize the cost of generating the forecasts and so calibrate the efficiency dimension. This is another way of saying that because one has the “correct/appropriate” forecasting model it does not follow that the forecasts will be useful/effective.

This is the point of departure of our study where the Bloomberg forecasting heuristic was tested and found to be ineffective. Lusk and Koulayan (2007) report:

We investigated the performance of the Bloomberg forecasting heuristic: $1/3+2/3 \times \beta$, as a one-period-ahead forecast of the one-factor CAPM β . We tested this Bloomberg heuristic using data from 131 companies that were on the S&P 500 continuously for more than 15 years. We found that the Bloomberg forecasts of β were more than five times higher in Absolute Percentage Error (APE) than the APEs produced by Collopy and Armstrong (1992) using Rule Based Forecasts of general time series of economic data

We have now, in this research report, expanded the scope of the inquiry to examining the effectiveness of forecasting the usual three measures of excess market return: (i) Jensen’s α ($J\alpha$), (ii) the Sharpe Performance Index (SPI) and (iii) the Treynor Performance Index (TPI). In addition, we have revisited the possibility of effectively forecasting β -i.e., market indexed or systematic risk. This is a “validity” check on the current study; finally, to complete the forecasting of firm risk, we have added to our study the evaluation of forecasting non-systematic risk, called idiosyncratic Risk (iR). This then is a rich rendering of the Return/Risk profile of the firm.

In summary, we are interested to ascertain the possibility of effectively forecasting these characteristic Return and Risk measures of the firm’s market profile. This study is called for, because surprisingly, extensive searches of the business literature have not identified research reports that speak to the possibility of effectively forecasting these five critical market measures Poon and Granger (2003) and Hooper *et al.* (2005). To be clear, there are a large number of research reports that test the forecasting of beta or other related

market measures, using various different forecasting models. All of the research reports that we reviewed did not report the APE relative to actual as a recommended by (Armstrong and Collopy, 1992). We offer that the information reported on relative forecasting performance of various models is of little interest to financial managers who should be interested in the effectiveness of the forecasts of important financial parameters not in which models give the best ineffective forecasts. Therefore, in this research report, we are ONLY interested in the possibility of producing Effective Forecasts which we define as Forecasts which have median APE of not greater than 10% relative to actual.

MATERIALS AND METHODS

Forecasting of market performance: As noted there are five measures that usually form the basic profile of the market performance of the firm: $J\alpha$, the SPI, the TPI, β and iR. Following we will give brief definitions of these measures. The reader is referred to (Lusk *et al.*, 2008) and also to (Brealey *et al.*, 2008) for more details on these measures.

Return measures: Jensen’s Alpha ($J\alpha$) is the difference between the average return of the firm, (R_f) and the β -conditioned CAPM projection of the market premium: [$\beta \times (R_m - R_{rf})$] added to the risk free rate: (R_{rf}). The risk free rate is usually surrogated by the return paid on the US Department of the Treasury short term investment certificates often call the T-Bills. $J\alpha$ is formed as the following:

$$\text{Identity: } J\alpha = R_f - [R_{rf} + \beta \times [R_m - R_{rf}]]$$

Thus, Jensen’s α is a measure of the excess of the firm’s average period returns over what is expected by the CAPM Projection: [$R_{rf} + \beta \times [R_m - R_{rf}]$] (Nielsen and Vassalou, 2004).

The SPI and the TPI are benchmarked measures of average firm period returns over and above the average of the period risk-free-rate. In this way they are both measure of excess return as is $J\alpha$. The SPI is benchmarked relative to total risk where total risk is surrogated by the standard deviation of the returns of the firm: (sd_f); the TPI has the same numerator as does the SPI but is benchmarked by systematic risk surrogated by the period β . The formulae are:

$$\begin{aligned} \text{SPI} &\equiv [R_f - R_{rf}] / sd_f \\ \text{TPI} &\equiv [R_f - R_{rf}] / \beta \end{aligned}$$

Risk measures: The CAPM- β is the slope of the OLS-two parameter, one stage, linear regression of the firm

returns as the response variable given the matched market returns. We will use, as the market surrogate, the S&P500. All of the firm and market return data was downloaded as an EDT from the CRSP data source through WRDS™: A data service of the Wharton School of the University of Pennsylvania.

Non-market risk is the ANOVA “variation” left over after the OLS filter has been applied to the data series. This was first proposed by (Sharpe, 1970) and has since been refined by (Ben-Horim and Levy, 1980) whose measure we will use as it is unbiased compared to the Sharpe measure. Non-market risk is often referred to as idiosyncratic Risk (iR). It is operationalized as:

$$iR \equiv sd_r - [\beta \times sd_m]$$

The forecasting study design: These five market measures: [Excess Return: {J α , SPI and TPI} and Risk: { β and iR}] were tested as to the effectiveness of their one-period-ahead predictability. These are usually annual measures of performance and so we have used as the forecasting period: One-year-ahead. We used, as does (Ibbotson, 2010), a rolling contiguous window-i.e., non-overlapping 5 year time series segments to measure points in the times series for each of the five market variables. This gives then a time series of 25 measures for each variable starting in (1980-1984) until (2004-2008). So as to not bias the study to a particular sector, we accrued from firms often classified as part of: the New Economy (NE), the Old Economy (OE), a group from the Vice-sector-i.e., Drugs, Alcohol and Weapons (Vice) and another group. The condition for inclusion was that the firm had to be continuously traded on the NASDAQ or NYSE exchanges for the 29 year accrual period. Two firms were eliminated due to lack of reported data or activity for some time segments, thus giving the final study set of n = 58.

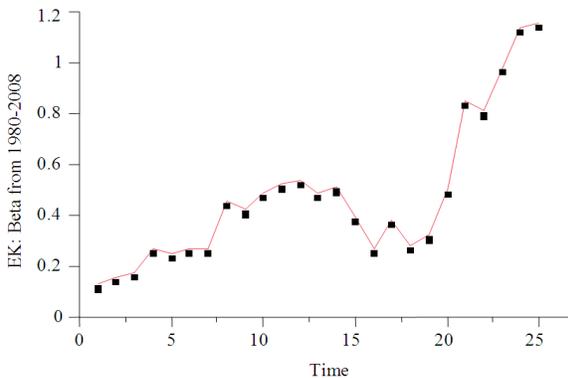


Fig. 1: Time series plot of EK of β , n = 25

We have included for review purposes information on all the firm tickers and the sector classifications as well as the data details. This information may be found in the Review Excel Worksheet labeled: Study_Variable_Information; tab: APE Data here appended. This information is uploaded to Scholarly Commons [http://repository.upenn.edu/].

A further aspect of the study was to conduct the forecasting analysis using screened data as well as unscreened data. Lusk *et al.* (2009) show that often there are firm performance differences as between screened data and its un-screened counterpart. This is of course due to the existence of outliers and to some extent to non-central-fat-tailed distributions (Filzmoser *et al.*, 2005; Gelper *et al.*, 2010). For our study all five variables were measured twice: once applying three screens: A trimming window with width {Mean \pm 2 \times Standard deviation}, the Box-Plot screen due to Tukey which is a window of width {Median \pm 1.5 \times IQR} and a relational screen due to (Mahalanobis, 1936) which uses the Mahalanobis-D measure to screen correlation outliers and is set at the 95% CI level (Sall *et al.*, 2008; Mitchell and Niederhausen, 2010). These three screens were applied only once in the order noted and eliminated, on average, approximately 15% of the data. We are mentioning this as a robustness issue but are not going to give more details as the inferences from the screened data and the non-screened data were identical for our study. Thus, we will report the screened data results as they are conservative respecting rejecting the Null. For further information on the screening procedures used (Lusk *et al.*, 2009).

In summary, we had 58 firms, five measures of financial performance each of which was a time series of 25 points. Each such time series point was developed from a dataset of five years of market and firm trading daily data from (1980-1984) to and including (2004-2008). For example, in Fig. 1 is a graphic of the time series of β for the Eastman Kodak Corporation EK, NYSE see the DemoDataSeries tab in the Excel file called: Study_Variable_Information.

Each point on this graphic was developed from about 1,200 daily returns of Eastman Kodak (EK) and the S&P500 matched returns; the entire graph then used about 30,000 daily returns for each the firm and the market.

For the purposes of the forecasting, we used the first 24 time series points to forecast the 25th point. Therefore for the EK data set, as the data exhibits strong Durbin-Watson and Fisher’s Kappa autocorrelation, as did more than 90% of the data series for the study, we used the Holt model [two-parameter

exponential smoothing model as programmed in SAS: JMP Version8]; this was the default model for our study. The forecast produced for EK was: $H(25) = 1.27703082$ and the computed value of β for EK for the five years 2004-2008 was 1.1568276. This produces the summary test statistic Absolute Percentage Error (APE relative to actual) of:

$$APE = \text{Abs}[1.27703082 - 1.1568276] / 1.1568276 = 10.4\%$$

There will then be 58 such APE calculations-one for each of the accrued firms; we will use this information to determine if the 24 yearly time series measures of β lead to an effective forecast of β . Our a-priori test of effectiveness will be the RHS one-tailed test of the Research Hypothesis that the median APE population value is $\leq 10\%$; we will use for inference the Wilcoxon Signed Rank Test. We have relaxed our measure of forecast effectiveness from the median APE for general time series presented by (Collopy and Armstrong, 1992) of 6.3% used in (Lusk and Koulayan, 2007) to 10% due to the fact that in numerous studies we found that 10% seems to be an acceptable practical APE limit in most forecasting studies (Adya *et al.*, 2009).

Therefore, median APE values greater than 10% that have one-tailed p-values less than 0, 01 will suggest that the actual realization from the sample would unlikely be drawn from an APE population centered at 10% in favor of the alternative that the underlying APE is greater than 10%. In this case, our inference thus will be that the forecast would not be effective in providing actionable decision information.

RESULTS

Factor study: To give a “credibility” check on the accrual of the firms used in the study, we will first report two validity pre-hypothesis test results. The first are the Factor Results of our study. For the factor study, we used the five study variables each measured under two screening protocols: Screening, noted as “Mod Data” and Non-screened data-i.e., directly downloaded and used without modification, noted as “Non-Mod”. Our expectation for our five measures each measured twice, from extensive reporting of such information relative to the firm market performance measures, is that:

- There will be three factors which will have eigenvalues greater than 1.0
- iRisk for the screened and non-screened data will load in a dominate fashion-i.e., the Varimax rotated projections for these two variables will be greater

than $\sqrt{0.5}$ and have the same sign and so define a factor; the same will be true for the two measures of β

- the six measures of Excess Return will load together on a factor

This is to say that if we see results at variance from these expectation this will call into question the generalizability of our results. This reason for this is that iR, β and Excess Returns are usually independent constructs and so for a general datasets should produce a three factor set as indicated above. The results found in the Excel file called Study_Variable_Information under the tab: Factors and the Factors data is included in the tab: FactorData. These results are reproduced in Table 1.

As is clear these factor results strongly support the validity expectations; the only slight exception is that the TPI Non-Mod data factor loading was not greater than $\sqrt{0.5}$ essentially due to a few outliers in that dataset. Also, the variable Beta Download are the values of β that were downloaded from the CRSPTM database. In this instance, we were interested to determine if the β that we computed from the daily dataset and the β -Downloaded were correlated. As one can see, the β s downloaded and computed are highly correlated as so group together. In summary, Factor 1 is the Excess return factor; Factor 2 is the iRisk factor were the two measures of iR, screened and non-screened load together and finally Factor 3 is the β -loaded factor. The dominate variable loading are highlighted in bold. These results simply suggest that the data is a reasonable sample and offers confidence in generalizing the results of the study.

Based upon these factor results we have added a regression test to the forecasting of $J\alpha$. We will use the SPI and the TPI as the X factor and $J\alpha$ as the Y or response variable. We will then test if using the SPI or TPI variable aids in developing an effective forecast of $J\alpha$. This is demonstrated in the Excel file: Study_Variable_Information the DemoDataSeries tab for the BNI $J\alpha$ ModData(Y) and BNI: SPI ModData (X).

Table 1: Rotated factor patterns

Variables	CAPM excess returns	Non-market risk	Market risk
BetaDownLoad	0.014	-0.030	0.90
Beta Mod Data	-0.110	-0.080	0.94
JAlpha non-mod	0.830	0.340	0.03
JAlpha Mod data	0.820	-0.350	-0.14
BH-L iR non-mod	-0.110	0.960	-0.10
BH-L iR mod data	-0.120	0.950	-0.06
SPI non-mod	0.830	-0.018	0.07
SPI mod data	0.830	-0.290	0.01
TPI non-mod	0.670	0.100	-0.43
TPI mod data	0.800	-0.310	-0.11

Table 2: The test of the effectiveness of the return/risk forecasts

	$J\alpha$ (OLS)SPI	$J\alpha$: ES	SPI:ES	TPI:ES	β :ES	iR:ES
Median APE	40.9%	35.6%	72.3%	78.1%	25.6%	41.1%
p-value	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
IQR	[22 -149%]	[11-81%]	[30-175%]	[33-250%]	[14-42%]	[22-64%]

Using these two date series, $n = 24$ and the OLS regression, we find that using the SPI value at point 25 of 0.09824738 one generates the Y-response of 0.00089233 as the forecast of $J\alpha$. Note this is a bias forecast in favor of the 10% hypothesis as the SPI value was known and not forecasted as would be the case in an actual organizational setting. We consider this then as the best case forecast for using the SPI to forecast $J\alpha$. As the results were the same for the SPI and the TPI we will report only the SPI results.

The second validity check is that in the previous study reported by (Lusk and Koulayan, 2007) the median APE for the forecasts using the Bloomberg forecast of β was 20.5% and for the Holt model 20%. The median APE for β for this study was 25.6%. It is of interest that the median for β of 25.6% is “on the order” or close to the previously reported values of 20.2% or 20%. This gives another generalizability verification point:

Results for [$J\alpha$, SPI, TPI, β and iR]: The APE results are reported in Table 2.

DISCUSSION

We can see that for the OLS regression: $J\alpha \leftarrow$ (OLSReg[SPI]) and for the Exponential Smoothing (ES) times series forecasts of $J\alpha$, note as $J\alpha$: ES, as well as for the ES time series forecasts of the TPI, β and iR, that the median APE from the sample strongly supports the rejection that the median APE is 10% or less which is our maximum acceptable value from an information decision perspective. For example, consider the median of the 58 ES model forecasts of iR. The median APE of the sample was 41.1%. The chance that a median of 41.1% or greater could have come from an APE population centered at 10.0% by random sampling chance would happen less than 1 time in 10,000. This probability value suggests rejecting the proposition that the APE is 10% or less in favor of the alternative that the APE is in fact greater than 10%. As this was the case for all of the variables so tested the conclusion is that for the times series of Return: { $J\alpha$, SPI, TPI} and Risk: { β and iR} there is no evidence to support the likelihood of producing effective forecasts.

As an informational note, in all but a few cases there was strong evidence of autocorrelation using the

Fisher’s Kappa test. As indicated above, our initial model of choice for forecasting was the Holt model given its performance in (Makridakis *et al.*, 1982). Note that the Holt model is also the ARIMA (0,2,2) model. However, for the Holt model, when there were Hessian, Stability or Invertability problems identified by the SAS:JMP system, we used an alternative model from the Exponential Class: {Simple: ARIMA (0,1,1), or Double Exponential Smoothing due to Brown (1963); the Brown model is also the ARIMA: (0,1,1) \times (0,1,1)} or the ARIMA (1,0,0) model (Sall *et al.*, 2008; Box *et al.*, 1994). After application of one of these models, which essentially exhausts the indicated times series modeling possibilities, the Fisher’s Kappa test suggested that there was no remaining residual structure. In this case then we used the forecast from this model as it was considered the appropriate model as there was no significant structure remaining after the application of the selected forecasting model. There were however two cases where there were anomalous results. This happened only in two instances: For the firms: CMTL and NP. These data points were eliminated. See Study_Variable_Information; APE Data.

Summary: The results of the study may be summarized as:

- Even given the strong autocorrelation of the 25 measured values for the five Return and Risk measures, there is no evidence that Exponential Smoothing/ARIMA models which are the time series model recommended in the presence of autocorrelation produce effective forecasts
- These results are robust compared to using Screened Data or Data downloaded directly from the EDT source
- These results are consistent with the information reported by (Lusk and Koulayan, 2007) regarding the forecasting of β
- In addition, similar results were found for the Y/X: regression of $J\alpha$ as the response variable given the SPI a relationship that was taken from the factor study results
- Finally, solely to complete the one-stage time series modeling possibilities, even though there was strong autocorrelation, we also used the OLS time series regression model to develop the

forecasts-i.e., in this instance the OLS time series regression is not the indicated model. Here, for this reason, there was no Fisher's Kappa checking. These time series regression results were no different than the results from the ES model results in that the median APEs tested higher than 10%

The strong implication from this study where we have extended the previous results reported by (Lusk and Koulayan, 2007) to consider the effectiveness of forecasting Returns: $\{J\alpha, \text{SPI}, \text{TPI}\}$ and Risk: $\{\beta \text{ and } iR\}$ is that there is no evidence that forecasting the Return and Risk measures will be effective; therefore, this result calls into question the wisdom of developing forecasts of these measures if one needs to have median APE errors of 10% or less. If financial managers can accept an APE greater than 10%, we have provided the IQR for the Median estimates from our study; these ranges may provide useful expectations.

CONCLUSION

This is a critically important result in that (1) no others have reported testing the effectiveness of forecasting these measures and (2) in our experience, it is often the case that financial analysts and managers are too busy to track or post-audit the effectiveness of their forecasts. In this environment one makes forecasts using the recommended models and "assumes" that because these are the forecasts from appropriate forecasting models-i.e., the modeling assumptions are satisfied, that the forecasts are therefore effective. Our study has shown this NOT necessarily to be the case.

Perhaps then it rests as a challenge to search for forecasting models that can develop effective forecasts. One could examine other Transfer Functions of the ARIMA class although many of these models were also tested in (Makridakis *et al.*, 1982) and found not to provide better results than did the ES class of models in particular the Holt or Brown's Double Exponential Smoothing model. Also a possibility is Rule Based Forecasting models proposed by (Collopy and Armstrong, 1992) that adds a judgment component into the analysis. Testing of such RBF models has one difficulty in that the data needs to 100% current to avoid the judgmental bias of forecasting into a known future; so there can be no "holdback" to calibrate the effectiveness of the forecasts. So these studies are very challenging from an experimental design perspective (Lusk *et al.*, 2010).

Considering the likely inability of the other Transfer Function models and the experimental design issues of judgment models such as RBF, we offer that we seem to

be in the same forecasting conundrum as one finds in trying to forecast daily stock prices/return. Perhaps, one-period-ahead Return and Risk forecasting is just not an area where effective forecasting works despite our motivation to forecast such Return and Risk information.

ACKNOWLEDGEMENT

We wish to thank Professors Neuhauser of the Department of Finance, Lamar University, Beaumont Texas, USA and Lee, Chuo-Hsuan, The Department of Accounting SUNY: Plattsburgh, Plattsburgh, NY, USA for their detailed and constructive comments.

REFERENCES

- Abbondante, P., 2010. Trading volume and stock indices: A test of technical analysis. *Am. J. Econ. Bus. Admin.*, 2: 287-292. <http://www.scipub.org/fulltext/ajeba/ajeba23287-292.pdf>
- Adya, M., E. Lusk and M. Balhadjali, 2009. Decomposition as a complex-skill acquisition strategy in management education: A case study in business forecasting. *Dec. Sci. J. Innov. Educ.*, 7: 9-36. DOI: 10.1111/j.1540-4609.2008.00199.x
- Armstrong, J.S. and F. Collopy, 1992. Error measures for generalizing about forecasting methods: Empirical comparisons. *Int. J. Forecast.*, 8: 69-80. <http://ideas.repec.org/a/eee/intfor/v8y1992i1p69-80.html>
- Ben-Horim, M. and H. Levy, 1980. Total risk, diversifiable risk and nondiversifiable risk: A pedagogic note. *J. Financ. Quant. Anal.*, 15: 289-297. <http://www.jstor.org/pss/2330346>
- Box, G., G.M. Jenkins and G. Reinsel, 1994. *Time Series Analysis: Forecasting and Control*. 3rd Edn., Prentice Hall, New Jersey, ISBN: 10: 0130607746 pp: 592.
- Brealey, R., S. Myers and F. Allen, 2008. *Corporate Finance*. 9th Edn., McGraw Hill, New York, pp: 515.
- Brown, R.G., 1963. *Smoothing, Forecasting and Prediction of Discrete Time Series*. 1st Edn., Prentice-Hall, New Jersey, ISBN: 10: 0138153086, pp: 468.
- Collopy, F and J.S. Armstrong, 1992. Rule-based forecasting: Development and validation of an expert systems approach to combining time series extrapolations. *Manage. Sci.*, 38: 1394-1414. <http://www.jstor.org/stable/2632670>
- Cootner, P.H., 1964. *Random Character of Stock Market Prices*. 1st Edn., The MIT Press, Mass, ISBN: 10: 0262030098, pp: 510.

- Filzmoser, P., R.G. Garrett and C. Reimann, 2005. Multivariate outlier detection in exploration geochemistry. *Comput. Geosci.*, 31: 579-587. DOI: 10.1016/j.cageo.2004.11.013
- Gelper, S., R. Fried and C. Croux, 2010. Robust forecasting with exponential and holt-winters smoothing. *J. Forecast.*, 29: 285-300. <http://ideas.repec.org/a/jof/jforec/v29y2010i3p285-300.html>
- Hooper, V.J., K. Ng and J.J. Reeves, 2005. Beta forecasting: A two decade evaluation. University of New South Wales. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=868867
- Ibbotson, S.B.B.I., 2010. Valuation Yearbook. Market Results for Stocks, Bonds, Bills and Inflation. 1st Edn., Morningstar Press, Chicago, pp: 214.
- Lusk, E.J. and H. Koulayan, 2007. Forecasting: An evaluation of the Bloomberg heuristic. *Int. J. Invest. Manage. Financ. Innov.*, 4: 56-60.
- Lusk, E.J., M. Halperin and F. Heilig, 2009. Market and financial performance as related to idiosyncratic risk and the effect of outlier screening in market studies. *J. Financ. Manage. Anal.*, 22: 59-69.
- Lusk, E., M. Halperin and V. Mitrev, 2010. Forecasting daily returns on the NYSE: A rule based forecasting test. SBE.
- Lusk, E., M. Halperin, M. Bern and M. Idiosyncratic, 2008. Risk and social responsibility: Adding to the R^2 meta-analysis. *J. Financ. Manage. Anal.*, 21: 1-38.
- Mahalanobis, P.C., 1936. On the generalized distance in statistics. *Proc. Nat. Inst. Sci. Ind.*, 2: 49-55. <http://www.citeulike.org/user/ashleygeorge/article/4155812>
- Makridakis, S., A. Andersen, R. Carbone, R. Fildes and M. Hibon *et al.*, 1982. The accuracy of extrapolation (time series) methods: Results of a forecasting competition. *J. Forecast.*, 1: 111-153. DOI: 10.1002/for.3980010202
- Mitchell, G. and M. Niederhausen, 2010. On replenishing items with seasonal intermittent demand. *Am. J. Econ. Bus. Admin.*, 2: 90-102. <http://www.scipub.org/fulltext/ajeba/ajeba2190-102.pdf>
- Nielsen, L.T. and M. Vassalou, 2004. Sharpe ratios and alphas in continuous time. *J. Financ. Quant. Anal.*, 1: 103-114. <http://www.jstor.org/pss/30031896>
- Pelaez, R.F., 2003. Ten-year forecasts of real stock price changes. *Empir. Econ.*, 28: 417-429. DOI: 10.1007/s001810200139
- Poon, S.H. and C.W.J. Granger, 2003. Forecasting volatility in financial markets: A review. *J. Econ. Literat.*, 41: 478-539. <http://ideas.repec.org/a/aea/jeclit/v41y2003i2p478-539.html>
- Sall, J., L. Creighton and A. Lehman, 2008. JMP™ Start Statistic-Version 8. 4th Edn., Duxbury Press, California, pp: 434.
- Sharpe, W.F., 1970. *Portfolio Theory and Capital Markets*. 1st Edn., McGraw-Hill, New York, pp: 316.