

Short-Term Forecasting in the Trading Markets: Suggested Capture Intervals for Bloomberg™ FA-Downloads

Edward J. Lusk

Department of Accounting: School of Business and Economics

State University of New York (SUNY) at Plattsburgh

101 Broad St., Plattsburgh, NY, USA

E-mail: luskej@plattsburgh.edu

or

Department of Statistics: The Wharton School

University of Pennsylvania, 3730 Walnut St. Philadelphia, PA, USA

E-mail: lusk@wharton.penn.edu

Received: January 22, 2019

Accepted: February 8, 2019

Published: February 17, 2019

doi:10.5296/ijافر.v9i1.14354

URL: <https://doi.org/10.5296/ijافر.v9i1.14354>

Abstract

General Context: Forecasting in the market trading world has been recognized as the core decision-making imperative. *Archimedes* remarked: *Give Me a Fulcrum, and I Shall Move the World*; the simile in the market trading world is; *Give Me an Effective Forecasting Model and I can make Bill Gates and Sam Walton look like Paupers*. An essential aspect of creating forecasts that can inform decision-makers are confidence intervals on forecasting projections.

Research Focus: The Bloomberg™ terminals [BBT] provide a treasure-trove of historical data and also, on a selected basis, single-valued forecasting projections. However, these forecasts are single-value estimates, usually one or two periods-ahead only for the Income Statement variables. The focus of this research report is to produce an interval that can be: (1) simply oriented from reported BBT-information, and (2) used to better develop a decision-context for the information accessible through the FA-link of the BBT.

Results: Using forecasting models in the OLS regression-class for: Time-Series, $Y = f(X)$

versions, and a Moving Average Model, seven forecasting capture intervals were developed from randomly selected firms offered by the BBT. These capture intervals were used to produce standardized-capture-scalar-coefficients that were applied to the reported values offered by the BBT to create one-period-ahead capture intervals. This modeling protocol is referred to as the Empirical Interval Capture Model [EICM]. The EICM produces a capture interval with precision of about 30% where the capture rate was on the order of 90%.

Keywords: Random effects, Fixed effects, Moving average

1. Introduction

1.1 Forecasting Context

The essential feature of forecasting is to give likelihood information on specific forecasting projections. Clear is: A single forecast value is of questionable value in informing the decision-making process without likely boundary considerations. Such boundary values are the “GPS” for decision-makers that are essential to making informed decisions. This is succinctly expressed by Hockey & Cormann (2014) who note:

Estimates of economic and fiscal variables over the forward estimates period are subject to inherent uncertainties, which generally tend to increase as the forecast horizon lengthens. Confidence intervals quantifying estimates of uncertainty around the key 2014-15 Mid-Year Economic and Fiscal Outlook forecasts have been constructed using a set of historical forecasting errors based on forecasts made since 1998 (where errors are defined by the difference between the forecasts and actual outcomes). These confidence intervals highlight that there is a range of plausible alternative outcomes around any given point estimate and provide a guide to the degree of uncertainty around these forecasts, typically spanning a wide range of outcomes.

Hockey & Cormann offer *spot-on* conditions for the critical need of using prior empirical information-sets to form useful decision-maker guidance in the forecasting context. [*Pedagogical Note:* The URL of the Hockey & Cormann citation has a wealth of useful information on the critical nature of the developing contextual information for forecasted values. This is a required reading in the Auditing course offered where forecasting platforms are used to create audit evidence in the Analytic Procedure Phase as recommended risk setting through economic analysis. (Note 1)

1.2 Related Literature

Given the obvious importance of an interval context for forecasts and projections, one would expect that there would be a plethora of research reports on: (i) how to construct forecasts from the extensive data reported for the firms listed on the Bloomberg™ Market navigation platform—one of the eminent sources of market information, and (ii) the effectiveness or utility of the Estimates that are provided by Bloomberg for various accounts reported for the Income Statements [GAAP]. However, after conducting comprehensive searches on ABI-Inform™ [ProQuest™] and Business Scours Premier™ [all accessed 5 January 2019 through Lippincott Link of the Wharton School of the University of Pennsylvania], no research

reports were located that specifically addressed these issues. Indeed, this was very surprising!

The two research reports that conducted analyses of BBT information-sets on the financials of the firms therein reported were:

Alister, N.E. (2001) reports in the Abstract of a Master's thesis

Market participants expend considerable resources on forecasting services to improve their decision-making processes. One of the more popular sources of these services is that of Bloomberg. In recent years Bloomberg has regularly polled reputable experts on macroeconomic forecasts in many countries. This study aims to test the accuracy of Bloomberg's forecasts. - - -. This research shows that for the most part Bloomberg's forecasts surveys are unbiased and rational predictors of macro-economic indicators for most countries.

Also there is a Doctoral dissertation level report from Mangee, N.J. (2011) that notes in the Abstract:

The main purpose of this dissertation is to determine which class of models – bubble or Imperfect Knowledge Economics (IKE) – provides the better account of short-term stock price fluctuations – and thus long-swings – on the basis of empirical evidence. - - -. This thesis addresses these difficulties in two main ways. One is to construct a novel dataset based on Bloomberg News' end-of-the-day equity market wrap stories. The textual data provides unambiguous support for IKE models over the bubble models. - - -. The thesis also relies on formal econometric analysis to reexamine the connection between stock prices and fundamental factors. It employs recursive structural change tests and cointegration and out-of-sample fit analyses. The results support those obtained with the Bloomberg data: short-term stock price fluctuations are related to fundamentals but the relationship between prices and fundamentals is temporally unstable at times and in ways that cannot be fully foreseen.

Finally, as expected, there were reports on using downloads of data from the BBTs in the forecasting context. This genre used data from various sources including Bloomberg to develop and test their modeling techniques. For a typical example, the report of: Amendola, A., Candila, V., & Antonio Scognamillo, A. (2017) where they note in the Abstract:

Forecasting oil prices is not straightforward, such that it is convenient to build a confidence interval around the forecasted prices. To this end, the principal ingredient for obtaining a reliable crude oil confidence interval is its volatility. Moreover, accurate crude oil volatility estimation has fundamental implications in terms of risk management, asset pricing and portfolio handling. Generally, current studies consider volatility models based on lagged crude oil price realizations and, at most, one additional macroeconomic variable as crude oil determinant. This paper aims to fill this gap, jointly considering not only traditional crude oil driving forces, such as the aggregate demand and oil supply, but also the monetary policy rate. [- - -] We also find that a negative (positive) variation of industrial production increases (decreases) Crude Oil Forward Price volatility. This means that volatility behaves counter-cyclically, according to the literature.

None of these reports that were retrieved developed specific calibration of the information reported by the BBTs that would enhance the utility of the navigation of the BBTs to create interval guidance for the BBT users.

1.3 Research Focus

This is the point of departure of this research report. Specifically, following is discussed:

1. An overview, with reasonable detail, of the essentials of the components of the Empirical Interval Capture Model [EICM],
2. The information accrued from the Bloomberg Terminals [BBT] used in executing this study,
3. The seven (7) sets of confidence intervals that are employed in developing the capture-interval-scalar-coefficients of the EICM,
4. The details of the EICM protocol,
5. The *Developmental* Dataset [DS] used in forming the EICM & The *Holdback* Dataset [HS] used in testing the acuity of the EICM—specifically, the percentage of time the EICM captured the BBT data tested,
6. A Summary of the results and suggested research directions.

2. The Forecasting Models and Their Confidence Interval Configurations

2.1 Protocol & Caveats Overview

Note should be taken that the model to be presented in this paper is termed: The Empirical Interval Capture Model [EICM]. The term Capture is Key. These Capture Intervals are formed from 95% Confidence Intervals [95% CIs] produced by passing random Panels accrued from the Bloomberg Terminals [BBT] through seven (7) standard forecasting models—i.e., statistical parameterized filters. It is the case that the validity of these 95% CIs is dependent on the myriads assumptions that underlie the standard forecasting models that were used. Question: Were the eight (8) or so assumptions needed to rationalize the “correct” inferential use of the OLS two-parameter linear-regression class of Time Series [TS] and $Y: f[X]$ models verified? No! Rather, their selection and use in this study was rationalized based upon the standard of “precedent”—these model forms are usually the models of choice in the Panel context, as reported almost exclusively in the literature. Specifically, the TS model forms used are: (i) in the select group of formal models use in the numerous time series competitions usually labeled as the Makridakis M-competitions; see the first of these competitions: Makridakis, et. al. (1982); These studies are previewed in the Archives of the *International Institute of Forecasters*™ [<https://forecasters.org/>], (ii) one of the three components of the basic model variants used in the Rule Based Forecasting [RBF] Expert System offered by Collopy & Armstrong(1992), (iii) one of the model forms of the recently reported Complexity Scoring Model of Adya and Lusk (2016). As for the $Y: f[X]$ models in the Panel context these models were used by: Sharpe (1964), Mossin (1966), and Lintner (1965) as a CAPM-volatility screen relative to their Market Panel, usually the S&P500. Also relevant is the work of Blume

(1975, 1979) as a follow-up on the OLSR-CAPM context. Finally, and more recently, the Y:f[X] OLSR context forms the filtering basis of the derivative work of Fama & French (1992, 2012). Therefore, these OLSR models have an established precedent of note. Further, consistent with the above precedent of the OLSR model-class there were no modifications in the precision of the 95% CIs nor of their model parameters, usually β . They were used as fitted. Nor were other model classes tested to determine if there were to be a better conceptual model relative to the assumptions of the inputted dataset. For example, in the Panel context, the ARIMA (0, 2, 2) or Holt model, that is actually one of the RBF component models, is often argued as preferred to the OLSR class of models. Certainly the Holt model mollifies the autocorrelation of the residual series relative to that of the OLSR two parameter linear fit.

This is a protracted way of noting, for transparency, that the assumption-validity of the [95% CIs] of the seven (7) forecasting models that were used in this study relies on their president in use; this is why various model forms were selected and their 95% CIs were uniformly combined to form an aggregate interval. Aggregation and Averaging [A&A] as a protocol is a standard feature in the RBF model as well as other models in the M-competitions as it forms a centroid of values that seems, according the reports of the M-studies, to better align with future realizations so long as there are not biases in the aggregate. This is why no parameter weighing schemes for this study were developed as they would require judgmental re-calibrations using the assumptions of the fitting series for which there are no tested forms for such re-calibration adjustments as demonstrated by the precedent studies all of which eschewed such a research tact.

2.2 The Capture Intervals: En Bref

Using the A&A as the basic construct, the overview of the EICM is:

1. Initially, a random sample of Panel segments (Note 2) is selected from the Income Statement [GAAP] and the Balance Sheet [Std] from a random sample of firms, $n = 26$ that were listed on the ICS:BICS platform of the Bloomberg Terminals. The Ticker Symbols of these firms are reported in the Appendix.
2. Using this Panel, one-period ahead forecasts were created for seven (7) different forecasting models and the 95% confidence intervals [95% CI] were created individually for each of these seven forecasting models.
3. This dataset had seven 95% CI for each firm in the study. For each of these 95% CI, a standardized scalar score was created so as to use their *Aggregate* to form the EICM protocol. Specifically, the Lower and Upper 95% CI were parameterized by the mid-point of the 95% CI. For example, For BBY: [Bed, Bath and BeYond] using reported *Operating Income* [[OI]:the Y-variate] & *Cost of Revenue* [[CoR]: the X-variate] for one of the forecasting models used in this study, the Lower 95% CI value was: \$385.13 and the Upper 95% CI \$2 724.67. In this case, the scalar-standardization for BBY for this forecasting model was:

Lower Scalar Value: 24.77% [$\$385.13 / (\text{AVERAGE}(\$385.13 : \$2\,724.67))$]

Upper Scalar Value: 175.23% [$\$2\,724.67 / (\text{AVERAGE}(\$385.13 : \$2\,724.67))$]

4. Then the Median [the *Averaging* feature] was computed for all of these standardized scores for the accrued firms for the Lower 95% CIs and the Upper 95% CIs. Note these respectively as: $\{LEICM_{Median} \& UEICM_{Median}\}$.

5. Finally, to create the projection, the analyst forms the following information set:

One-Period-Ahead EICM Interval for some given Y_t :

$$\text{LowerLimit} = LEICM_{Median} \times Y_t$$

$$\text{UpperLimit} = UEICM_{Median} \times Y_t$$

For example, assume that overall the $LEICM_{Median} = 83.45\%$, the $UEICM_{Median} = 116.55\%$, and $Y_t = \$10,000$. In this case, the EICM Capture points are: LowerLimit: \$8,345.00 & UpperLimit: \$11,655.00.

6. Finally, the capture rate of the EICM for a Holdback Sample [HS] will be reported. In this Holdback validation, there were random selections of accounts from the HS, $n=13$, from the BBTs. The capture information used the Panel-values at Y_{t-1} so that the EICM interval can be evaluated for the point Y_t , e.g., the last reported actual values or if estimates were provided by the BBT, then these estimates were used. These computations will be illustrated following.

Consider now the forecasting models that will be used in creating the information-set that will be used in forming the final EICM statistical scalars: $\{LEICM_{Median} \& UEICM_{Median}\}$.

2.3 The EICM: Forecasting Components

As mentioned above, the construct guiding the creation of the information is: Aggregation and Averaging [A&A]. To form a stable set of forecasting information using the A&A construction, we used seven (7) forecasting models all of which can be simply programmed in Excel using only the basic functionalities—i.e., non-VBA-code. Following for each of the seven models a concise illustration using the BBBY values for the Y-variate: *Operating Income* [OI] & the X-variate: *Cost of Revenue* [CoR] Panel (Note 3) will be reported.

2.3.1 Moving Average Model [MAM]

For the MAM, the Excel version of the MAM [Period=2, Se] was used for the dataset *Cost of Revenue*: [CoR]: BBBY. In this case, the MAM produces the estimate of the X-variate. This particular MAM is a rolling two-period non-memory indexing model. The Standard Error [Se] option was elected and used to form the 95% CIs in the standard manner. For example, using the datasets for BBBY for the MAM forecast value for Period 5, the CoR is: \$3,952.87: [Average[3,782.027, 4,123.711]] and the standard error at this point using the Excel functionality is: SUMXMY2[] (Note 4) noted as EQ1 and computed as:

$$606.08 = [\text{SQRT}(\text{SUMXMY2}(3,782.027; 4,123.711, \text{AVERAGE}[2,961.377, 3,323.814]; \text{AVERAGE}[3,323.814, 3,782.027])/2)]$$

Using these computations for the full dataset the [CoR], the terminal point $t+1$ had a *Se* of 330.50 which is used to form the 95% CIs from the regression Y [OI]: X [CoR] as follows:

The MAM [CoR] one-point-ahead-projection, i.e., $t+(n+1)$ is: 7 561.49 [AVERAGE[7 483.577; 7 639.407]]. The Se using the Excel Function EQ1 for the CoR is: 330.502. Thus the CoR projection that will be used in the regression for the EICM is:

$$\text{CoR Projection: } 7\,237.60 \text{ [MidPoint}[[(7\,561.49 - (1.96 \times 330.502)) : 7\,561.49]]$$

Thus 7,237.60 is the value that will be used to project the CoR-point to form the 95%CI from the 95%CI of the Excel regression Y [OI]: X[CoR], $t=1, \dots, 13$, \$7 237.60. *Point of information*. As has been demonstrated by Collopy and Armstrong (1992) there seems to be a tendency to “over-project” future values, even in the short-term. For this reason, the MAM:CoR projection was “deflated” by using the value that is the mid-point of: [The projection: The lower-limit of the MAM:CoR 95%CI]. Therefore rather than using the Excel slope multiplier of 7,561.49 or that of the LowerLimit of this estimate: 6,913.706 which may be too low of an indexing projection, the value used was: [6 913.706 + [$\frac{1}{2} \times$ the precision of [1.96×330.502]]] or 7 237.60. This deflates the projection, i.e., moves it more to the regression line. This is important as the estimates from the other six forecasting models were not slope deflated; these values used were the exact Excel values produced from the Regression platform as noted above.

Specifically, for the above information the following is produced:

The Excel functionality for the forecast: Y [OI]: X [MAM [CoR]], $t=1, \dots, 13$ have a 95%CI using the CoR of 7 237.60 of:

Interval: Y: X [MAM[CoR]]:

$$\text{LowerLimit} = [95\% \text{LowerLimitCI Intercept} + [95\% \text{LowerLimitCI Slope} \times 7\,237.60]],$$

$$\text{UpperLimit} = [95\% \text{UpperLimitCI Intercept} + [95\% \text{UpperLimitCI Slope} \times 7\,237.60]].$$

Specifically for the BBY Panel:

$$\text{LowerLimit} = 353.54: [-215.4504 + (0.0786 \times 7\,237.60)]$$

$$\text{UpperLimit} = 2\,622.03: [773.2840 + (0.2554 \times 7\,237.60)]$$

The MAM is one of the standard pre-cursors to most of the various forms of the Exponential Smoothing models. It is data driven and has very few assumptions that form the calculation base. To that extent, it is an independent form of a projection model. The MAM, as scripted above, is used to create an X-variate to be used in the Y [OI]:X[[MAM]CoR] regression of the BBY-dataset. To form the 95% CIs, the Excel:*DataAnalysis:Regression* platform was used. Recall the X-variate projection, as formed above, is a slightly dampened projection following the spirit of Collopy and Armstrong (1992) where they actual have trend-dampening RBF-rules. They argue that failing to have a dampening modality will often create projections that have relatively high Absolute Percentage forecasting errors. So for at least one of the seven models dampened projection were created. As a final note, deflating the X-variate also is reasonable as the MAM is the highest variance form as of the MAM.

2.3.2 Regression Models

There are two frequently used variations on the OLS two-parameter linear-regression model [OLSR] regarding the creation of Confidence Intervals [CI]: The Y:X version and the Time Series model: $Y: X[\text{Time Index}]$. Both of these have three variations: The Excel Projections, The Random Effects CI-modality and the Fixed Effects CI-modality. Consider now these model versions; each will be illustrated using the BBBY datasets.

2.3.2.1 OLSR Inference From the Excel Parameter Range Model

The Excel Regression functionality forms a “wide-covering” confidence interval. These are effectively extreme case CI-scenarios as they are produced from the crisp-end-point parameters of the 95%CI for the intercept and the slope jointly. Here we offer the following notation:

Extreme Left Side [LowerLimit[LL]] 95%Boundary for actual client YE-value: $Y_{[t=(N+1)]}$

$$\hat{Y}_{\text{LowerLimit:}(t=(N+1))} = [\hat{\alpha}_N - [t_{\frac{\alpha}{2}} \times s_{\epsilon(\alpha)}]] + [\hat{\beta}_N - [t_{\frac{\alpha}{2}} \times s_{\epsilon(\beta)}]] \times [V] \quad (2)$$

Extreme Right Side [UpperLimit[UL]] 95%Boundary:

$$\hat{Y}_{\text{UpperLimit}(t=(N+1))} = [\hat{\alpha}_N + [t_{\frac{\alpha}{2}} \times s_{\epsilon(\alpha)}]] + [\hat{\beta}_N + [t_{\frac{\alpha}{2}} \times s_{\epsilon(\beta)}]] \times [V] \quad (3)$$

Where: $t_{\frac{\alpha}{2}}$ is the t-statistic for inference for the two-tailed 95%CI that has $df = [N-2]$; N is the last time index in the data-stream—i.e., The last data point, V is either: $[N+1]$, which is the case for the TS model, or $RW=CoR_{t=N}$, which is the case for the Y [OI]: X [CoR] model. For example,

The TS version of the 95% CIs are:

Lower Limit: 759.99: $[383.3261 + (26.9046 \times (13 + 1))]$

Upper Limit: 2 530.49: $[1\ 023.9226 + (107.6122 \times (13 + 1))]$

Y: X version of the 95% CIs are:

LowerLimit= 385.13: $[-215.4504 + 0.0786 \times 7\ 639.41]$

UpperLimit= 2 724.67: $[773.2840 + 0.2554 \times 7\ 639.41]$

The CoR value used here is NOT the MAM projection as this was the central construct in the MAM model; therefore to avoid a possible redundancy-bias and so to have the most germane variation possible the Random Walk [RW] projection also used in the RBF model was used. In this case the RW is: 7 639.41: CoR[$t=N$].

2.3.2.2 OLSR Inference From the Expected Random Effects Model

The OLSR assumption, in this case, there is random sampling possibility from a well-defined

population of variable data streams of realizations, i.e., the set of:

$$\{Y_t^l; l = \text{Firm Index where } l \text{ is large}\}$$

which constitutes a large blocked or stratified group of “like” firms. In this case, the auditor believes that the sample estimate and the related CIs are formed under the expectation, $E[Y_t^l]$, i.e., *the forecast of the mean of Y_t^l* ; this is termed the Random Effects [RE] assumption. Experience suggests that the RE assumption is possible though not likely in the audit context. Nevertheless, in the service of completeness, the confidence interval for the client value in the RE context for the TS model is:

$$\hat{Y}_{(t=N+1)} \pm t_{\frac{\alpha}{2}} \times s_{\epsilon} \times \sqrt{\frac{1}{N} + \frac{[(N+1)-\bar{N}]^2}{S_{tt}}} \quad (4)$$

Where: $\hat{Y}_{(t=N+1)}$ is the value of the fitted regression projected for the next X-index time-point, $N + 1$, using the parameters produced from the OLSR fit for the N data points; $S_{tt} = [\sum t^2 - [(\sum t)^2/N]$; $s_{\epsilon} = \text{OLS}[N]:MSE^{1/2}$; and \bar{N} is the Mean of the time index for the N data points.

For the Y:X version we have:

$$\hat{Y}_{(t=N+1)} \pm t_{\frac{\alpha}{2}} \times s_{\epsilon} \times \sqrt{\frac{1}{N} + \frac{[(RW)-\overline{CoR}]^2}{S_{xx}}} \quad (5)$$

Where: $\hat{Y}_{(t=N+1)}$ is the value of the fitted regression projected for the next X-index time-point, $N + 1$, using the parameters produced from the OLSR fit for the N data points; $S_{xx} = [\sum CoR^2 - [(\sum CoR)^2/N]$; $s_{\epsilon} = \text{OLS}[OI:CoR]:MSE^{1/2}$; and \overline{CoR} is the Mean of the CoR Panel.

For example, for the TS version:

$$1\,324.94 = 1\,645.24 - [2.200985 \times 247.3449 \times 0.5883]$$

$$1\,965.54 = 1\,645.24 + [2.200985 \times 247.3449 \times 0.5883]$$

Where [For illustration]: S_{tt} : $[819 - [91^2]/13] = 182$; $[(14 - 7)^2] = 49$ Therefore, the third component is: $[\sqrt{(1/13 + 49/182)}] = 0.5883484$; $\sqrt{MSE} = \sqrt{61\,179.48131344} = 247.3448631$; $T.INV.2T(5\%,11) = 2.20098516$; Finally: $1\,645.24$: $[703.624346 + [67.2583791 \times (1+13)]]$

For the Y: X version:

$$1\,309.41 = 1\,554.90 - [2.200985 \times 229.9799 \times 0.48499]$$

$$1\,800.39 = 1\,554.90 + [2.200985 \times 229.9799 \times 0.48499].$$

Where: [For illustration]: S_{xx} : $[406\,477\,309.67 - [69\,699.83^2]/13] = 32\,779\,901.83$; $[[7\,639.407 - 5\,361.53]^2] = 5\,188\,744.65$. Therefore, the third component is:

$$[\sqrt{(1/13 + (5\,188\,744.65/32\,779\,901.83))}] = 0.48499; \sqrt{MSE} = \sqrt{52\,890.7472} = 229.9799;$$

T.INV.2T(5%,11) = 2.20098516; Finally: 1 554.90: [278.9167 + [0.1670261 × 7 639.407]]

2.3.2.3 OLSR Inference From the Fixed Effect Projection

The assumption is that the object of interest is the j^{th} firm with a single set of data stream realizations: i.e., the set of:

$$\{\{Y\}j^l\}; j = a \text{ particular firm Index over all } l \text{ firms}\}$$

Where this j^{th} Firm has longitudinal dummy-variable integrity from all of the other firms in the population and so the projection is not the average of all the firms but only for that j^{th} firm. In this case, and given the usual assumptions rationalizing the OLSR of the time-series, the confidence interval for the extrapolation for the next point in the firm time stream is:

$$\hat{Y}_{(t=N+1)} \pm t_{\frac{\alpha}{2}} \times s_{\epsilon} \times \sqrt{1 + \frac{1}{N} + \frac{[(N+1)-\bar{N}]^2}{S_{tt}}} \quad (6)$$

Where: $\hat{Y}_{(t=N+1)}$ is the value of the fitted regression projected for the next X-index time-point, $N + 1$, using the parameters produced from the OLSR fit for the N data points; $S_{tt} = [\sum t^2 - [(\sum t)^2/N]$; $s_{\epsilon} = \text{OLS}[N]:MSE^{1/2}$; and \bar{N} is the Mean of the time index for the N data points.

For the Y:X version we have:

$$\hat{Y}_{(t=N+1)} \pm t_{\frac{\alpha}{2}} \times s_{\epsilon} \times \sqrt{1 + \frac{1}{N} + \frac{[(RW)-\overline{CoR}]^2}{S_{xx}}} \quad (7)$$

Where: $\hat{Y}_{(t=N+1)}$ is the value of the fitted regression projected for the next X-index time-point, $[N + 1]$, using the parameters produced from the OLSR fit for the N data points; $S_{xx} = [\sum CoR^2 - [(\sum CoR)^2/N]$; $s_{\epsilon} = \text{OLS}[OI:CoR]:MSE^{1/2}$; and \overline{CoR} is the Mean of the CoR Panel.

For example, for the TS version:

$$1\ 013.60 = 1\ 645.24 - [2.200985 \times 247.3449 \times 1.1602]$$

$$2\ 276.88 = 1,645.24 + [2.2009\ 85 \times 247.3449 \times 1.1602].$$

For the Y:X version:

$$992.33 = 1\ 554.90 - [2.200985 \times 229.9799 \times 1.1114]$$

$$2\ 117.47 = 1\ 554.90 + [2.200985 \times 229.9799 \times 1.1114].$$

2.3.3 Forecasting Profile

Therefore, the seven 95% CIs that are produced for BBBY are presented in Table 1:

Table 1. The BBY profile for the EICM set of forecasting models

Models	Lower 95%CI	Upper 95U%CI	Precision
MAM[2,Se]	\$353.54	\$2 622.03	\$1 134.25
OLS:TS: Excel	\$759.99	\$2 530.49	\$885.25
OLS:TS: RE	\$1 324.94	\$1 965.54	\$320.30
OLS:TS: FE	\$1 013.60	\$2 276.88	\$631.64
OLS:Y:X:Excel	\$385.13	\$2 724.67	\$1 169.77
OLS:Y:X: RE	\$1 309.41	\$1 800.39	\$245.49
OLS:Y:X: FE	\$992.33	\$2 117.47	\$562.57
Average BBY	\$876.99	\$2 291.07	\$707.04

In this case, an illustration the concept of creating the EICM coefficients for the Lower and Upper EICM CIs will be useful.

The calibration of the EICM values is presented in Table 2.

Table 2. Points for the lower and the upper EICM-scalars formulation from the BBY download

Models	Lower 95%CI	Upper 95%CI	EICM Lower	EICM Upper
MAM[2,Serror]	\$353.54	\$2 622.03	23.76%	176.24%
OLS:TS: Excel	\$759.99	\$2 530.49	46.19%	153.81%
OLS:TS: RE	\$1 324.94	\$1 965.54	80.53%	119.47%
OLS:TS: FE	\$1 013.60	\$2 276.88	61.61%	138.39%
OLS:Y:X:Excel	\$385.13	\$2 724.67	24.77%	175.23%
OLS:Y:X: RE	\$1 309.41	\$1 800.39	84.21%	115.79%
OLS:Y:X: FE	\$992.33	\$2 117.47	63.82%	136.18%

2.3.4 Illustration

Using the set of information in Table 2, one could create the EICM scalars and then use the Median of the Lower & Upper EICM values for creating the Capture Intervals. Specifically,

here only as an illustration, the first step is to form the scalars that will be applied to the end-point values so as to form the EICM Capture Intervals. For the MAM CIs the scalars [Boded] are:

$$\text{Scalars for the EICM Lower} = \frac{353.54}{\text{Average}[353.54:2\ 622.03]} = 23.76\%$$

$$\text{Scalars for the EICM Upper} = \frac{2622.03}{\text{Average}[353.54:2\ 622.03]} = 176.24\%$$

For these seven end-point scalar percentages the Median was selected to form the EICM Capture Interval scalar-coefficient set. The Medians of the Lower and the Upper are: 61.61% & 138.39% respectively [Shaded]. Finally, assume that the BBY YE[2017] for Gross Profit is: \$4,576.35; thus the EICM Capture Intervals $\{LEICM_{Median} \& UEICM_{Median}\}$ are:

$$LEICM_{Median} = \$2\ 819.49 \ [\$4\ 576.35 \times 61.61\%]$$

$$UEICM_{Median} = \$6\ 333.21 \ [\$4\ 576.35 \times 138.39\%]$$

Incidentally, the Bloomberg estimate for 2018 is: \$4 442.29 and falls into the EICM CIs.

However, to provide the mean-estimates for the EICM Capture Intervals, of course, the ALL of the 95% CIs from the DS-Samples accrued from the BBTs will be used.

3. The EICM Intervals: The Results

Twenty-Six (26) firms were accrued from the BBT: [BICS platform]; they are noted in the Appendix. They were re-sampled into a Development Set [DS, n =13] and a Holdback Set [HS, n =13] for providing an inference structure. Then various accounts were randomly sampled from these firms. For the DS there are about 35 end-point Accounts for the BS and for the IS there are about, 25 end-point Accounts or 60 selection possibilities for the regression platform thus providing 30 Account-Pairs for each firm. This produces a sample frame for the regression-arm of the study of 390 end-point Account Pairs or approximately 2,730. 95% CIs given that each Account Pair produces seven (7) 95% CIs. For the HS there are 60 Accounts or 780 Accounts that could be selected for application of the EICM to evaluate that capture rate as produced by the DS.

3.1 Sample Size Accrual

For the sample size for the evaluation of the capture rate from the HS, which is the inferential test-set of the information produced from the DS, it is assumed that:

1. The EICM capture rate would be 85% - i.e., the binomial success rate - this is reasonable as we computed the 95% CIs for all of the datasets, therefore 85% is certainly conservative,
2. The $[1-\alpha]$ -confidence level for the precision is 95% relative to the capture interval produced by the sample size, and

3. Finally, the Power-test precision is 5% - the difference between Ho[85%] and Ha[90%]. This then will give the sample size for the test of Power of 75% - i.e., a False Negative Error[FNE] of 25%.

In this case, the sample size will be 326 and assure a Power of 75% of detecting an increase in the capture rate from 85% to 90%.

The second sample size computation was for the accrual size of forecast events relative to a reasonable level of precision of the Upper and Lower EICM scalar limits. In this regard, it seemed reasonable to have a precision of 5% for the Mean of the scalar-sets. True, the Mean will not actually be used for the scalars and the sample size, in a random selection context, has no effect on the expectation relative to the population estimate; but, the Means of the scalar confidence intervals will be reported, as they could be of further research interest, thus an accrual calculation would be useful. In this case, as this is a test of samples points and Sample Size/Power for the t-test does not have an unconditional form - i.e., for this context one must use a non-central t-distribution that is conditioned on three variables. One simple way to deal with this is to create a pilot estimate of the sample standard deviation and assume a test precision. This pilot yielded an estimate for the standard deviation of 0.60. Setting the precision of the mean at 5%, the sample size for the t-calculation produced a sample size of 520. In these sampling plans, replicates were permitted so as to conform, insofar as possible, to the infinite population sampling assumption. In this case, there were no replications in the 326 plan and two in the 520 plan. The 520 sample set requires two account selections for each point; in examining this accrual-set there were two datasets that seemed corrupted as there were a few Panel values that were detached from an autocorrelation context. These two points were eliminated from the 520 plan; thus the final sample size for determining the scalars was 518. For the 326 plan there was two sample anomalies where the comparison value was judged to be a recording/download error as their magnitude relative to the previous longitudinal point was a factor higher in one case and in the other it was lower by a factor. This produced a final sample size of 324.

3.2 Developmental Profile

The 518 sampling plan produced the following profile for the scalars.

Table 3. Developmental scalars for the EICM

	Lower Limits	Upper Limits
Mean	0.4906	1.5094
Median	0.6747	1.3253
95% Confidence Interval*	0.4448 to 0.5452	1.4548 to 1.5552

*Note that the Precision is: $0.0502 [(0.5452 - 0.4448) \times 50\%]$ which is approximately the accrual design specification.

In this case, for the reasons of controlling the precision given the non-normal skewness and kurtoses both of which, of course, are identical for the profile of the LLimit and the ULimit which is a feature of the calibration that is used, the Median values, bolded & shaded, are used as the EICM scalars; the other information offered in Table 3 is presented in the case that other versions, such as the Mean or sampled values from the Range, of the EICM scalars would be of use.

Therefore, EICM scalars for creating capture boundaries for ANY value in the Bloomberg Platforms are:

$LEICM_{Median}$: LowerLimit Scalar: $EICM_{LL}$ [67.469900653691%] (Note 5)

$UEICM_{Median}$: UpperLimit Scalar: $EICM_{UL}$ [132.530099346309%]

Finally, the Precision of the EICM intervals is: [32.53%]

The next test is to evaluate this EICM-scalar information set using the Holdback sample of 324 observations.

3.3 Holdback Profile

There are two sets of datasets that are used to test the quality of these EICM-scalars $\{LEICM_{Median}: UEICM_{Median}\}$ one from the BBT-Income Statements the other from the BBT-Balance Sheets selected from the HS-firms accrued.

3.3.1 Income Statement Accruals

For the Income Statements [IS] of the accrued firms: we selected randomly from the following account groupings:

We used for each selected IS-account the 2016 reported value as the base for the EICM interval: $\{LEICM_{Median}: UEICM_{Median}\}$. Then we determined if the 2017 value was in fact in [inclusive] or not in the EICM [2016].

Further, the IS 2017 values where there was a BBT Estimate for 2018 associated with 2017 actual value were selected. And the EICM [2017] interval value was used to determine if the BBT-Estimated value for 2018 was in the EICM [2017].

3.3.2 Balance Sheet Accruals

Samples were taken from the BS [Std] and used to create the EICM intervals $\{LEICM_{Median}: UEICM_{Median}\}$ from the 2016 values so as to judge the 2017 values.

4. The Actual Capture Rates

The EICM-scalars, noted above, were applied to the various random samples selected from the IS & BS of the BBTs. Capture is recorded if the value of interest is in[interior] to the $\{LEICM_{Median}: UEICM_{Median}\}$ intervals. This information is presented in the following Table 4.

Table 4. Evaluation profile for the EICM

	Capture Rate	Percentage
IS[2016v.2017Capture]	187/192	97.40%
IS[2017v.2018[Est]Capture]	83/97	85.57%
BS[2016v.2017Capture]	30/35	85.71%
Overall Weighted Average	300/324	92.59%

The codex is: The numerator is the number of instances that were IN the $\{LEICM_{Median}: UEICM_{Median}\}$ Capture Intervals and the denominator is the number of Panel-Points selected from the BBTs accrued firms. These results are certainly encouraging. The EICM CIs are:

1. Aggregations and Averages over many random accounts panels from the BBT, and so are *Empirically* grounded,
2. *Simple* to use in forming *Capture Intervals* that overall can be expected to provide interval guidance that is accurate to, on-the-order-of, 90%. Specifically, using the 95%CI over the seven standard forecasting models to form the EICM-scalars, the 95%CI for the capture rate of 92.6% is: [87. 9% to 97.3%]; note that 95% is in fact in this interval. Further, considering the False Negative Error [FNE] Concern - there [is/was/may be] a Lower Capture rate assumed as the Null and the statistical analysis fails to detect that the actual capture rate is higher and thus - the analyst commits a FNE by failing to reject the Null of the lower rate. To address this issue assume that the capture rate [is/was/may be] 85%, if the alternative, *Ha*, were to be a 90% capture rate, then using a uniform blend of the standard error and an $(1-\alpha)$ confidence of 90%, the FNE of failing to reject 85% in favor that the capture rate is higher would be 13.95 % which is a Power rate of 86.05%. This means that most of the time the analyst would make the correct decision to believe that the actual capture is higher than 85% and so the testing is certainly reliable and is consistent with the sample size Power calculation offered above.
3. Reasonably *Precise*; the EICM CIs have a precision of, on-the-order-of: 30%.

5. Summary and Outlook

5.1 Summary

This investigation was begun by noting that single forecast projections are of little value in the dynamic economic and market trading world. Additionally and surprisingly, the world standard in market navigation platforms: the *Bloomberg™* set of Market Navigation Platforms [MNP] does not offer interval guidance for their reported data. Given this lack of market guidance, EICM were developed using scalar values for creating one-period-ahead capture intervals that are based upon collection of empirical values of seven forecasting intervals that may be applied

to selected BBT Panel values. The precision and the EICM-capture rates using $\{LEICM_{Median}:UEICM_{Median}\}$ seem to be in the range of practical use in the using the information drawn from the BBTs. This is not surprising as the basis of these EICM capture intervals is formed from the aggregation and averaging of actual forecasting models using empirical data drawn from the BBTs. The EICM are recommended not only to financial analysts but also to those in the Bloomberg organization who may see the need to form Capture Intervals or at least a related modeling perspective that will aid in the de-coding of the market.

5.2 Outlook

Indeed there are numerous studies that flow naturally from the rich datasets of the BBTs, three of which are:

1. There are numerous possibilities afforded by the BBT: Balance Sheets [BS] and Income Statements [IS] to create one period-ahead EICM Information. Recall that there are, on average, 60 Accounts in the BS & IS, take together, that could be used to create EICM intervals. This Capture EICM-reconfiguration of the BS & IS would provide a very informative platform for the BBTs. In this regard, one could evaluate various different account configurations to determine if the EICM Capture rate is sensitive to the nature of the financial statement account groups, such as accounts that impact the Cash Flow from Operation compared to the other account groupings.
2. There are BBT-Estimates presented in the IS for the next reporting period usually the next FY. These began Interval context so as to better inform the users of the BBTs. As noted above, the practical utility of a single valued projection is questionable, bordering on useless, without an interval context. The EICM could offer instructive guidance for these Estimates in this regard. Also in the next reporting period the actual capture rate percentages of the EICM could also be reported.
3. Perhaps there are Event Horizon sensitives for the EICM. For example, consider the following three Event Horizons [EH]:

EH 1: [Circa 1991[the roll-out of the WWW] to 2002[The enactment of the audit requirements of the PCAOB[AS2]]. This period was the “Internet Bubble Era” that effectively led to the creation of the PCAOB due to the pervasive Audit LLP and SEC oversight failures during the 1990s. This Event period could certainly be labeled as one of significant deviations from assurance reporting realities for firms that were traded on exchange markets - *Enron, Quest & HealthSouth* to mention a few of major firms that did all the wrong things and the Accounting LLPs that were complicit - e.g., *Arthur Andersen, LLP*. This period could also serve as an excellent “worst-case” scenario sensitivity benchmark for the other Event periods.

EH 2: [Circa 2002 to September 2008 the month of the Lehman Bros. subprime debacle (Note 6)] This was a time period where the Public Accounting Audit LLPs were learning how to implement the guidelines of AS2 and also developing the audit protocols to justify the scripting of the COSO Opinion on the adequacy of management’s system of Internal Control over Financial Reporting. The information generated during this “pilot-time period” was subsequently used to re-issue a new Audit Standard [PCAOB release: AS5 circa 2007].

Therefore during this learning period perhaps there were market/reporting instabilities that may have been NAICS/SIC sector related, perhaps in the Financial Markets or Banking sectors that could have perturbed the nature of the EICMs; this would be a productive investigation.

EH3: [Circa 2009 to 2018] This has been a period of reasonable Assurance and Economic stability and a relative dearth of financial debacles. This period may offer an excellent benchmark for investigating Event Horizon effects on the EICM projections and capture rates for firms listed in the BBTs.

We look forward with enthusiasm to continuing the investigations of these topics. Collaborations are most welcome.

Acknowledgments Thanks and appreciation are due to: Mr. John Conners, Senior Vice President, Financial Counseling, West Coast Region, AYCO for his generous philanthropy which funded the establishment of the John and Diana Conners Finance Trading Lab at the State University of New York College at Plattsburgh and the Bloomberg Terminals that were instrumental in this research, Prof. Dr. H. Wright, *Boston University*: Department of Mathematics and Statistics, the participants at the SBE Research Workshop at SUNY: Plattsburgh, NY USA, in particular: Prof. Dr. Petrova, and Mr. Manuel Bern, Chief of Internal Audit: *TUI International, GmbH*, Hannover, Germany for their careful reading, helpful comments, and suggestions.

References

- Adya, M., & Lusk, E. (2016). Development and validation of a rule-based time series complexity scoring technique to support design of adaptive forecasting DSS. *Decision Support Systems*, 83, 70-82.
- Alister, N. (2001). *An empirical examination of Bloomberg's macroeconomic forecasts of the G-7 nations*. Concordia University (Canada), ProQuest Masters: Dissertations Publishing.
- Amendola, A., Candila, V., & Antonio Scognamillo, A. (2017). On the influence of US monetary policy on crude oil price volatility. *Empir Econ*, 52, 155-178
- Blume, M. (1975). Betas and their regression tendencies. *Journal of Finance*, 30, 785-795.
- Blume, M. (1979). Betas and their regression tendencies: Some further evidence. *Journal of Finance*, 43, 265-267.
- Collopy, F., & Armstrong, J. S. (1992). Rule-based forecasting: Development and validation of an expert systems approach to combining time series extrapolations. *Management Science*, 38, 1394-1414.
- Fama, E., & French, K. (1992). The Cross-Section of expected stock returns. *The Journal of Finance*, 47, 427.
- Fama, E., & French, K. (2012). Size, value, and momentum in international stock returns. *Journal of Financial Economics*, 105, 457-472.
- Hockey, J. B., & Mathias Cormann, M. (2014). Mid-year economic and fiscal outlook:

2014-15. *Commonwealth of Australia*.

Lintner, J. (1965). The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *Review of Economics and Statistics*, 47, 13-37.

Makridakis, S., Andersen, A., Carbone, R., Fildes, R., Hibon, M., Lewandowski, R., ... Winkler, R. (1982). The accuracy of extrapolation (Time Series) methods: Results of a forecasting competition. *Journal of Forecasting*, 1, 111-153.

Mangee, N. (2011). *The puzzle of long swings in equity markets: Which way forward?*. University of New Hampshire, ProQuest Doctoral: Dissertations Publishing.

Mossin, J. (1966). Equilibrium in a capital asset market. *Econometrica*, 34, 768-783.

Sharpe, F. (1964). Capital asset prices: a theory of market equilibrium under conditions of risk. *Journal of Finance*, 19, 425-442.

Notes

Note 1. https://pcaobus.org/Standards/Pages/05152014_Guidance.aspx

Note 2. The random sample of Account Panels was collected from 5Feb2018 through 18May2018 and the number of time-series-points were $n = 13$ the first value FY: 2005 and last being the reported data point in the BBTs for FY-ending 2017.

Note 3. Here are the two datasets [Millions of US Dollars] that we will use to illustrate the seven forecasting models.

BBBY: Operating Income [OI] Starting FY:2005, $n=13$

792.414	879.171	889.401	838.022	673.896	980.687	1 288.458
1 568.369	1 638.218	1 614.587	1 554.293	1 414.903	1 135.21	

BBBY: Cost of Revenue [CoR] Starting FY:2005, $n=13$

2 961.377	3 323.814	3 782.027	4 123.711	4 335.104	4 620.674	5 135.574
5 568.957	6 525.83	6 938.381	7 261.397	7 483.577	7 639.407	

Note 4. This Excel function is the programming of the following:

$$S_{MAM} = \sqrt{\sum((X - z)^2)/2} \quad (1)$$

Where: X is the independent variable in the OLS $Y=f(X)$ version and z is the Average of the two preceding points back-indexed from the time index of X.

Note 5. There is also the possibility to use the truncated values of: {67.47% & 132.53% } as the EICM-scalars as opposed to the full decimal representation. This is reasonable in that there is, perhaps, a precision anomaly in using the full decimal representation as the BBT

accounts basically are posted in monetary values—i.e., two decimal places. Some argue that the most realistic precision should then be restricted to two decimal places. As a point of information, we re-calculated ALL of the HD capture information and there was no change in the HD-capture profiles in using the two-decimal scalar profile.

Note 6. <https://www.investopedia.com/articles/economics/09/lehman-brothers-collapse.asp>

Appendix

Table A1. Development set of firms selected from the BBT [BICS: Ticker Index]

BBBY	TIF	CLT	AET	DPZ	DECK	IT
ANTM	UPS	RCL	CCL	TXN	MAR	DIS

Table A2. Holdback Set of firms selected from the BBT [BICS: Ticker Index]

XOM	AAL	CMG	WMT	AMZN	LGF	CVS
SKK	PRU	HD	DNKN	AAPL	TAP	IBM

Copyright Disclaimer

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>)