European Journal of Science, Innovation and Technology
ISSN: 2786-4936
www.ejsit-journal.com
Volume 5 | Number 3 | 2025

The RAE FM[Triage] Protocol: Selection Among Forecasting Model Candidates

Moncef Belhadjali[1] & Edward J. Lusk[2*]
[1]School of Business, Norfolk State University: Norfolk, Virginia, USA
ORCID ID: https://orcid.org/0009-0000-8219-8222
[2]Emeritus: The Wharton School, [Statistics], The University of Pennsylvania, USA & School of Business and Economics [Accounting], SUNY: Plattsburgh, USA & Chair: International School of Management: Otto-von-Guericke, Magdeburg, Germany
ORCID ID: https://orcid.org/0000-0002-6847-5956

ABSTRACT

Context. The Relative Absolute Error [RAE], proposed by Theil (1958) and calibrated by Armstrong & Collopy (1992), for a forecasting context, is a benchmarking-measure used by many forecasters. *Experientially*, we find that, due to its composition as: The ratio of two Absolute Percentage Error [APE]-profiles, the RAE is often confusing to Decision-Makers wishing to use the RAE to select among Forecasting Model [FM]-Candidates—i.e., FM-Triage. ***Focus***. In consultations, we often advise forecasters to *Keep It Simple* and initialize their FM-triage by using RAE-screening for: The two most basic and effective FMs: {The OLS-R Time Series [OLSR] & The Random Walk [RW]}. Our tacit assumption is: As there is NO FM that is the best overall for Panels that require forecasts for multiple extrapolations, ***initially*** examine ***which*** of these two Basic Models emerges as *Relatively the Best* as measured by their RAE-Profiles. ***Rationale.*** Over the years, both in consultations and for academic presentations, it has been clear that the Best FMs depend upon (i): The Nature of the Data, and (ii) The Nature of the Extrapolations. ***Focus.*** We offer, for the first time, an *instructive* RAE-triage-analysis of: {The OLSR & The RW} FMs for: (i) A Set of Panels randomly sampled from the S&P$_{500}$, then (ii) Grouped by their Coefficient of Determination [CoD], for (iii) Forecasts over three Extrapolations. ***Results.*** Indeed, the nature of the presentation of the RAE-FM-profiles can be a source of confusion. ***Thus***, we offer a multi-Stage *RAE FM*[*Triage*] *Protocol* to simplify the FM-selection procedure. We found that: (i) The Nature of the CoD-Dataset & Extrapolations likely impact the nature of the FM that has the best RAE-profile, and (ii) not infrequently, both the OLSR & RW need to be used to refine the forecasting acuity.

Keywords: Forecasting Model Evaluation, The Makridakis et al. (1982) M-Study

1. FORECASTING MODELS: THE OLS REGRESSION AND THE RANDOM WALK

1.1 Overview Our extensive experiential interactions in consulting & academia have consistently indicated that in *Time Series Forecasting*,

(i) The relatively: effective, efficient, simple and transparent FMs are:

{The OLSR Time-Series FM [OLSR]} ***and/or*** {The Random Walk FM [RW]}, and

(ii) The exact FM: configuration that seems to be the best depends upon: The Nature of: The Data & The Extrapolations use to generate the forecasts.

1.2 The FM-Specifics Even though the OLSR FM is extensively detailed in every forecasting text and in: Excel™ [Data[Descriptive Statistics[Regression]]], this is not the case for the RW FM. Thus, we offer the following to detail these exact FMs:

* Corresponding Author

I. The Random Walk FM, introduced by Makridakis *et al.* (1982: Initially Called: *The Naive 1*) is defined as:
The Y-Response Variate-***forecast*** is: The ***Actual Panel value at*** the Time Index, t_n where *n* is the last Y-Variate used to fit the OLSR TS FM. This will be the RW-forecast for ***all the forecasts made***, and

II. The OLSR TS FM,] **is:**

$$\hat{Y}_p = [\hat{\alpha} + [\hat{\beta} \times t_p]] \qquad \text{Eq[1]}$$

where: $\hat{\alpha}$ is the estimate of the Population-Intercept; $\hat{\beta}$ is the estimate of the Population-Slope of the Two-Parameter ***Linear*** forecasting model; these Population-estimates are determined by minimizing the **O**rdinary **L**east **S**quared Errors [*Actual v. Linear: Forecast*], *p* is a forecasting projection. Usually, *p* is in: Range [1, - - -, *n*, - - -, *m*.] where *m* is the practical Extrapolation-limit. For p_s > n, these are called Extrapolations.

1.3 Rationale for Electing these: "Keep It Simple-FMs*"* The reasons we prefer to ***initialize*** *The RAE FM[Triage] Protocol* with The {OLSR & The RW} FMs are:

1.3.1 Their Math/Stat Gearing Differentials

There are profound Math/Stat-definitional differences between: The {OLSR & The RW} FMs. The OLSR is a ***dynamic*** **FM** that is driven by its Coefficient of Determination [CoD] defined as:

CoD ≡ [The Pearson Product Moment Correlation [PPMC] of: [The Y-Variate-Panel *w.* The Time-Index]]^2

Thus,

I. The **CoD** of the Panel is identically the same as [≡] the $\boldsymbol{R^2}$ of the OLSR FM,

II. ***For The OLSR FM***: The $\boldsymbol{R^2}$ of the OLSR FM is the proportion of the variation in the Y-Variate that is accounted for by the OLS Regression on the X-Variate [The OLSR-Time-Index] See[Tamhane &Dunlop (2000, p. 354)][T&D]. Thus, the CoD is a ***Strength*** indication of the Nature of the ***dynamic association*** of the Y-Variate Panel with the Time-Index; where: We define ***Strength*** as: FM-prediction ***Accuracy/Acuity*** such as: Mean Average Percentage Error [MAPE], Mean Square Error [MSE], Average Ranking [AR], Medians of Absolute Percentage Error [Md], or Percentage Better [PB>]—i.e., these were the FM-performance measures used in: Makridakis *et al.* (1982). It is the case, that **if** the CoD tests to be inferentially different between two Panels of equal-size, **then** the Panel with the higher CoD usually will outperform the other Panel on most of the above forecasting **Acuity**-measures,

III. ***For The RW FM***: It generates a ***static-forecast*** that is *uniformly* the Y-variate at: t_n, noted as: [$\hat{Y}_n$].

1.3.2 Summary Rationalizations of the KIS-Election

These two FMs are categorically differentiable and thus are devoid of systemic: Math/Stat-***overlap***. This facilitates their conceptualization and ***rationalizes their performance Expectations*** as FMs. As the CoD is the R^2 of the OLSR FM, we will use the CoD-***Strength*** measure to group the panels to be forecasted. In this regard, we will use as our calibration and labeling of the CoD-Panels, the following Investopedia™ web-link* intel:

*https://www.investopedia.com/articles/economics/09/lehman-brothers-collapse.asp

> *"In investing, a high R^2, from 85% to 100%, indicates that the stock's or fund's performance moves relatively in line with the index. A fund with a low R-squared, at 70% or less, indicates that the fund does not generally follow the movements of the index."*

1.4 Historical Evidence of the Expected Extrapolation Profiles of: The RW FM v. The OLSR FM
We evaluated the Makridakis *et al.* (1982) dataset [hereafter noted as the: M-Study] and identified a *Veritable Cornucopia* of comparative "*Population*" data on the performance profile of 21 of the "*en vogue*" Forecasting Models *circa* the early 1980s. The M-Study has been used ever since to develop and calibrate many FMs in use today. In the M-Study, *Table 6(a)* is: a *Population* Profile of *The Percentage of Time that the* **RW FM** *was Better* [>] *than the Other Methods* ***one of which was the OLSR TS FM***! This ***abridged***-Table is presented following:

Table 1 ***abridged*** M-Study [**6(a)**: n=1,001] Percentage of times the **RW FM** is Better [>] than **the OLSR FM**

Extrapolations	t_1	t_2	t_3	t_4	t_5	t_6	t_7	t_8	t_9	t_{10}	t_{11}	t_{12}
RW FM[%]	62.6%	54.6%	52.0%	50.1%	44.7%	44.9%	48.0%	45.9%	50.6%	54.1%	56.2%	57.7%
50%/50% Marker	50%	50%	50%	50%	50%	50%	50%	50%	50%	50%	50%	50%
RW[%] > OLSR[%]	*Yes*	*Yes*	*Yes*	*Yes*	*No*	*No*	*No*	*No*	*Yes*	*Yes*	*Yes*	*Yes*

We label this as: ***A RW V-Bound***—i.e., the "*Yes*" indications in Row[4] correspond to the Profile where the RW[APE] was [***better***] than the OLSR TS[APE]; the "*No*" indications in Row[4] correspond to the Profile where the RW[APE] was [***not better***] than the OLSR TS[APE]. ***RW V-Bound*** seems reasonable given the differences in their Math/Stat-generating processes. We offer the ***RW V-Bound*** as a possible Population Structure/Feature for the interaction of: The OLSR & The RW FMs.

2. THE RAE FM PROTOCOL

2.1 Overview Following, we offer a simple and thus, readily understandable form, for encoding the RAE-profiles of: {*The OLSR TS FM* & *The Random Walk FM*}. To motivate and rationalize the need for a simplification, we offer an example of the *all-too-common* difficulty with scoring the forecasting performance of FM-Candidates by their RAE-values.

2.2 Scoring and Encoding the RAE The usual [or Standard Form] configuration for the RAE is:

RAE ≡ The Ratio of The Absolute Percentage Errors [APEs] of two FMs.

where: the ***Numerator*** is the More Complicated FM [FM_c]; and, the ***Denominator*** is the Simpler FM [FM_s]**.** Thus, the value of the RAE in this usual format is:

$$\text{RAE} \equiv [\text{APE}[FM_c]] \,/\, [\text{APE}[FM_s]] \qquad [\text{R1}]$$

For a **Particular Panel** under evaluation, **IF** the RAE is >1.0, this indicates the ***fact*** that:

The APE[FM_c] was > the APE[FM_s].

Thus, this rationalizes that the FM_s out-performed the FM_c on their APE-measures and so a logical Triage-choice would be: Use the FM_s. ***Simply*** A RAE > 1.0 indicates that The Logical choice is the FM_s ***for this Panel***!

Caveat Alert: ***What is the Fact for a Single Panel is usually NOT true for a Statistical Profile of n-Panels where the inferential measure is the RAE-value***. To illustrate this, consider the information offered in: {*Appendix B The inferential profile of the Standard Format of the RAE-configuration of Nine Panels*}. Given that the RAE test-measure is:

$$\text{RAE} \equiv [\text{APE}[FMc] \,/\, \text{APE}[FMs]]$$

and the Mean [RAE] of the RAEs for the 9-Panels tested is: ***2.22***, ***and*** the directional-FPE[Null[p-value]-test[$H_o[\mu_{RAE} = 1.0]$]], is: ***0.212***, this profile suggests that the H_a: is: $\mu_{RAE} > 1$. **This inferential profile** would, in turn, suggest that the FM of Choice would likely be The ***FMs*** as the Mean of the RAEs was ***2.22*** which is > RAE[$\mu_{RAE} = 1.0$].

Reality Check According to the ***actual***-profile for these nine panels, there are four panels where the RAE is > 1.0 indicating, on a *Panel Basis*, the Preferred FM is: **FMs, as the RAE, on a**

Panel Basis is >1.0. However, for the Last four Panels, the Preferred FM is: **FMc, as the RAE, on a panel Basis, is < 1.0**. Thus, the inferential reality is that **50%** is the Preference for each of the two FMs, indicating that the RAE does not inferentially suggest a preference between these two FMs!

2.3 Summary This inferential disconnect-**indicates and -demonstrates** that there is **usually no** reliable Triage FM selection information in the ***parametric*** profile of the RAEs-values tested over multiple Panels. For this reason, in this research report, we will suggest a **Category RAE-Protocol** that is simple and intuitive for selecting between two Forecasting Models in the RAE-context. In addition, following are three issues that need to be addressed in creating RAE-intel that will be used in developing our FM-Triage Protocol:

I. The Benchmark [*FMs*[APE]], not infrequently, is a very small fraction. This results in Box-Plot Whisker-Outliers that render the RAE-inferential results very difficult to understand in most decision-contexts. For this reason, The RAEs are usually Trimmed or Winsorized [Winsor (1895–1951)]. For a trimmed estimator, the extreme values are censored [eliminated from the computations]*;* for a Winsorizing protocol, the extreme values are instead *replaced* by certain percentiles. For example, Armstrong & Collopy (1992) [C&A] use a Winsorizing-Window where *all* the RAEs are ***IF-transformed*** so they are in the Interval [0.1 through 10.0] and so the sample-size is preserved; whereas, in the trimming/censoring case, the sample size is reduced. Both affect the inference-profiles in different ways. We recommend using the following *Excel*™ [C&A]-Winsorizing Screen:

=IF(RAE<0.01,0.01, IF(RAE>10,10,RAE))

In this research report, we will Winsorize all RAEs; however, they still will be noted as: **RAEs.** ***Point of Information*** For *The RAE FM[Triage] Protocol*, Winsorizing will not affect the inferential results. Trimming may, as there are many Trimming Protocols. For consistency, we will Winsorize all of the RAEs, on a Panel-Basis, as this seems to be the standard in the Forecasting Milieu,

II. ***In the forecasting context***, the *Nature of the Data* is critical in forming relevant APEs that are used to create the RAEs. In the usual TS Forecasting-context, the ***Y-Response Variables*** can be:

i. Generated by ***Multiplicative Process(es)***: ***If*** there is evidence that the panel-generating process*(es*) creates longitudinal panel-point-associations over the panel that have ***a high Coefficient of Determination*** **[*CoD*]**, ***then*** the Y-Response-transformation called for is: The log_{10}- or log_e[*ln*]-transform: noted as: [Y^{ln}]. ***Rationale*** A high CoD indicates that {The Y-Panel & The Time Index} are highly Pearson Product Moment Correlated. This indicates that the likely generating process of the Panel was ***Multiplicative*** in Nature and so the *ln*-transformation is often recommended to better linearize the data for forecasting. See: [Box & Cox (1964)] and [Collopy & Armstrong (1992: App B Rule2[p.1409)]. ***if this is not the case, then:***

ii. Y-Response Variate is assumed to be generated by ***Non-multiplicative Process*(*es*)**]. In this case, ***if*** the CoD is sufficiently low—i.e., no "interesting" evidence of PPMC-association for the {Panel-Points with the Time Index}, ***then*** the Y-Response transformation called for is: ***No Transformation***. The analyst just uses the data as downloaded; this is the, so called, Arithmetic Mean [***AM***]-option.

III. Finally, in the TS-domain **Interpolations** are almost never used to create forecasts. **Extrapolations** are the common FM-fare. However, for **Extrapolations**, T&D[p. 363] offer the following "Forecasting Bump in the Road" **T&D Conjecture**:

"- - -extrapolation beyond the range of the data is a ***risky business*** *should be avoided."*

If T&D are correct, *actually they never develop data to rationalize this causation*, that: ***Extreme*** *Forecasting Extrapolations somehow compromise the linkages of the Projection of the Past into a Relevant Future,* then indeed, *Extrapolations* should be the subject of scrutiny and, to be sure, testing.

3. THE RAE FM[TRIAGE] PROTOCOL: THE COMPONENTS

3.1 Creating the RAE-Inferential Testing Frame Following are the essential details of our *RAE FM[Triage] Protocol*. We will present and discuss them and then detail our test results for *The RAE FM[Triage] Protocol.*

The Selected Components of the Testing Frame are:

I. **FMs Tested** As noted and discussed above, the FM-Candidates for ***initial*** triage-testing with The RAE FM[Triage] Protocol will be: { *The OLSR TS FM & The RW FM* },

II. **Extrapolations** Referencing the *Tamhane & Dunlop -Alert* and after reviewing the population profile of the M-Study, we have selected *three Extrapolations:* $\{t_{n+1}, t_{n+2}$ & $t_{n+7}\}$ where: *n* is the Last Data Point used to fit the OLSR FM. ***Rationale*** We selected these Forecasting Projections as we were interested to determine if the ***RW V-Bound*** would follow the M-Study profile for the CoD-organized Panels where the expectation would be*:

$$\{t_{n+1}[\text{RW[RAE]}] > t_{n+2}[\text{RW[RAE]}] < t_{n+7}[\text{RW[RAE]}]\}$$

*Recall that the RW[MeanRAE%] ≡ [1-OLSR[MeanRAE%]

We have no *a priori* expectation; this is an *Investigative-Arm* of the design. As this is, an uncharted research domain, we suggest two Inferential Screening Conditionals for the support of the likelihood that there was likely a RW V-Bound:

Strong support for the RW V-Bound Screen**:** The Welch[ANOVA] test FPE[Null[p-value[RW[RAE]]]] over the three RAE-extrapolations must be **<0.1**. Further, the Tukey-Kramer HSD-Pairwise directional p-values for the two exhaustive tests:

$$[\{t_{n+1} > t_{n+2}\} \ \& \ \{t_{n+7} > t_{n+2}\}\}$$

must ***all be*** **<0.1,** or

Suggestive support for the RW V-Bound Screen**:** The Welch p-value <**0.25** & The above T-K[p-values] are **<0.25**.

III. **Simplified RAE Category Scoring** Given the analytic disconnect with using the parametric Statistical profiles of the RAEs as demonstrated in Appendix B, we will be using ***only*** Category Scoring for *The RAE FM*[*Triage*] *Protocol*. For example, using the *Standard Format,* ***on a Panel Basis***, the binary-scoring of the Two-Category RAE-Variables will be:

i. {***Prefer: The OLSR***} will be scored as: ***1.0, otherwise scored as: 0***} AND

ii. {***Prefer: The RW***} will be scored as: ***1.0, otherwise scored as: 0***}

Given this binary scoring protocol, the Forecasting Analyst will compute the Mean: RAE[Frequency[%]], **for *simplicity* noted hereafter as**: Mean[RAE[%]] for these ***two*** RAE: Category Variables. For example, assume that there were **50** TS Panels accrued and evaluated using *The RAE FM*[*Triage*] *Protocol*, and their Frequencies were scored as:

Mean[RAE[%]] [***Prefer: OLSR***] = 40% [20/**50**] and
thus Mean[RAE[%]] [***Prefer : RW***] = 60%[30/**50**]

This intel is simple to understand in a statistical context. In Summary:

There were 20 APEs of the OLSR that were < than the APEs of the RW, and

There were 30 APEs of the RW that were < than the APEs of the OLSR. This binary scoring will enable *a vetting check* as: The SUM of the binary Mean[RAE[%]] frequencies MUST = 100%. This category-scoring method is ***MUCH*** preferred to using the actual Mean[Values] of the RAEs. The reason for this is that usually the RAE[Mean]s are BOTH > 1.0 as we saw for the dataset in Appendix B. Our uniform experience is that even experienced Forecasting Analysts have a difficult time to de-code the implications of the two RAE[Mean]s if both are >1.0.

IV. **Dataset Accruals** We used the *Market Screener*™ Market Profiler to select our test firms. Specifically, we selected from the *Bloomberg*™ Market Trading Platform: The following three Groups of Firms Ranked by the *Market Screener*: {The **Top** 20[Firms: [1through 20], The **Middle** 20: [We selected a random number from 100 to 400 as the accrual-starting-point. Thus, we selected Firms [141 through 160]] and finally, The **Bottom** 20: Firms [481 through 500]]. The reason for this Selection Blocking was to offer variation in the testing context. However, will not be Blocking the RAE: analyses by S&P500 Rankings. As for the Variables to be used to parameterize the OLSR-Model, we selected the following variable sets for the forecasting model Eq[1]. For the Y-Response Variables, we selected five Income Statement [BBT[GAAP-version]] as listed on the BBTs:

*https://www.marketscreener.com/quote/index/S-P-500-4985/components/

Typically, most of these variables have Data-Generation Processes that produce ***Multiplicative***-longitudinal Y-Variate Panels:

EBITDA [BBT[line[**53**]]],
SALES_REV_TURN[line[**6**]],
GROSS_MARGIN[line[**57**],
OPER_MARGIN[line[**58**]], and
PROF_MARGIN[line[**59**]].

where: the BBT-**line** designation is where the variable is listed in the IS[GAAP]-version, and the BBT-definitions for these variables may be accessed by scrolling over the account and Right-clicking to pull-up the BBT-definition.

For these five Y-Variates. we selected ***only the first 14 Quarters for each firm***, usually starting with 3rd Quarter 2014 were selected. In addition, we selected three Forecast holdback actual Y-values [HBs] to create the APEs and so the RAEs. We selected the 3rd Quarter 2014 as the starting time index as it was more than five years after the *Lehman Bros*™ Sub-Prime financial debacle that almost crashed the world trading markets. We judged that this was a sufficient time-lag for the Markets to re-adjust after the 2008 Lehman-Event. See *LinkedIn*™ for a discussion of issues re: The Recovery from the Lehman-Debacle*. We selected 14-quarters as this was a Panel of sufficient length for over-lapping four calendar-years for fitting an OLSR FM. For the OLSR FM-fitting, if there were **any** instances of missing Panel-values, we eliminated such Panels. We did not use Regression- nor Near-Neighbor-fill for these missing-values: rather, we ***eliminated*** that Y-Variate Panel-set. Further, any Y-Variate Panel with a value ≤ 0, will not be possible to use in the TS-forecast protocol for the OLSR or the RW as these Data Panels are usually multiplicative and thus *may need* to be *ln*-transformed. These Panels were also ***eliminated***. The number of qualifying panels, after eliminations, was $n = 202$. Thus, as each Panel will have Three FM-projections there will be 606 evaluations indications.

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

V. **Forming the Panels to be Tested** For the Panels to be used in The RAE Triage Protocol, we recommend organizing the Y-Variate Panels into the following CoD ordered groups:

From: ***Low***[Mean[CoD]] *To:* ***High***[Mean[CoD]]

The reason for this grouping is that the Math/Stat generating Process(es) difference between: {The OLSR FMs *v.* The RW} is bi-categorical and non-overlapping: The OLSR Forecasts are ***dynamic*** and driven by the CoD and the CoD is ≡ The OLSR: R^2, whereas the RW Forecasts are ***static*** "*smoothing" sort of FM*. The implication of this is that it is possible use these CoD-drivers to create reasonable expectations of relative FM-Acuity. For example, for the Low[Mean[CoD]] where the CoD's strength is Low there is likely to be a Balance between: The OLSR & The RW FMs re: The RAEs where the Default FM to be selected would be the RW FM as it is the ***simplest version***. Whereas, for the High[Mean[CoD]], the OLSR FM should outperform the RW FM. Thus, our re-organization of the S&P500 Panels into **High-Low** ordered CoD: groups re: their Mean[CoD]-Profiles is expected to create the following Binary Split: **High** CoD-Panels will be dominated by the OLSR-FM; and the Panels in the ***70%*** and lower CoD-range are expected to engage **The OLSR, as well as, The RW FMs**.

3.2 Profiling and Testing of The RAE FM[Triage] Protocol For probing the CoD-screening-sets for their RAE-Triage-profiles, we offer the following intel, the intention of which is to clarify the nature of the computations and discuss their logic in profiling the Forecasting-Acuity of The OLSR FMs & The RWs:

I. ***CoD Titles*** In the Title of the CoD-Profile for the re-organized 202 S&P*500* Panels, we will note: The Mean[CoD] & The 95%CoD:CIs & Our ***expectation*** of the relative forecasting performance of The: {OLSR FM} *v.* {RW FM},

II. V***etting Check*** We will verify that the Mean[RAE[%]s] of: The RAE[OLSR/RW] & of: The RAE[RW/OLSR] **SUM to 100%**. If this is not the case*, then the inferential analysis will* ***STOP*** as this will signal an error. If founded, then

III. **Profiling Robustness** Recall, we have selected ***three Extrapolations:*** $\{t_{15}, t_{16}$ ***&*** $t_{21}\}$ **for inferential testing**. For each Extrapolation in the CoD-Table, we will compute the Non-Directional FPE[Null[p-value]] created from the test of:

$$[\ \bar{x}_n^{RAE} - \ \mu_{50\%}^{RAE}]$$

where: $\bar{x}_n^{RAE}$ is: The Actual-Sampled Mean[RAE[%]], and $\mu_{50\%}^{RAE}$ is: The Population *Test-Against of 50%.*

using ***two*** *different Standard Errors. Specifically,*

The N[0,1]-PDF-Version: $[\boldsymbol{z_{Cal}}]$*

$$\boldsymbol{z_{Cal}} = \textbf{[ABS[}\bar{x}_n^{RAE} - 50\%\textbf{] / [}\boldsymbol{\sigma_{N[0,1]}}\textbf{]] : } \boldsymbol{\sigma_{N[0,1]}} = \textbf{[[\{50\% × (1-50\%]] / n]\^{}.5,}$$

Assuming that the CoD-dataset is: BHH-qualified*, and

The t_{df}^{RAE}-PDF-Excel: Version: $[\boldsymbol{t_{Cal}}]$

$$\boldsymbol{t_{Cal}} = \textbf{[ABS[}\bar{x}_n^{RAE} - 50\%\textbf{] / [}\overline{\sigma}\textbf{]].}$$

where: The Vector of the scored RAE-activity is $[[n{\times}1]^{RAE}]$, using the *Excel* the Mean is: $\bar{x}_n^{RAE}$ & the $\overline{\sigma}$ is: [Variance$[[n{\times}1]^{RAE}]]$ / *n*]^.5

**Technical Note:* For the Test of Proportions [ToP] to be used as the inference tool where the Nature of the test is a Bernoulli Process, one uses the screen offered by Box, Hunter & Hunter [BHH] (1978) is to check IF [n × p is >5] & [n × (1 – p) is >5] where p = is the Population Frequency for the NULL-test. If this is the case, then the ToP-test is a reliable inference test that uses the N[0,1] PDF and context rather than the exact Binomial PDF. All of the datasets that were used in the above tables for the *overall results* passed the BHH-Screen.

Discussion The reason for assuming that are two Inference-Versions is to offer analytic-***robustness for each of the CoD-Tables***. Specifically, we will have The N[0,1]-inference version in Cells [3,4 : 4,4 & 5,4]; and, in Cells [3,5 : 4,5 & 5,5] we will have the t_{df}^{RAE}-inference version from the Excel-Platform. ***Inference Note*** The ***robustness*** difference is that the Standard Errors *as well as* the

FPE[Null[p-values] are not the same. Thus, the resulting FPE[Null[p-values]] are different for the two versions of the testing assumptions if: [The $\bar{x}_n^{RAE}$ ≠ The Test Against].

IV. ***95%RAECIs for the Three Extrapolations:*** $\{\boldsymbol{t_{15}}, \boldsymbol{t_{16}}\ \&\ \boldsymbol{t_{21}}\}$. For each Extrapolation, we will compute the 95%RAECIs using the *Excel*[Data[Descriptive Statistics] *Platform*. This uses the t_{df}^{RAE}-PDF for inference. For these computations, the 95%RAE:CI-results will be in Cells [3,3 : 4,3 & 5,3] & Cells: [3,6 : 4,6 & 5,6],

V. ***Overall Extrapolation Test*** For the set of three Extrapolations, the overall **Welch[ANOVA[Mean]** FPE[Null[p-value]] will be reported in Cells [6,2] & Cells [6,5]. **Note:** The Welch[ANOVA] assumes that the Group-Variances are not equal for the three Test-profiles. For each CoD Table, the Welsh p-values are identical for the RW- & OLSR-sections,

VI. We will offer an *"en bref"* **Narrative** of the ***overall*** performance of the OLSR- & RW-FMs in the CoD-Table, and

VII. Finally, we ***will score these CoD profile-results*** so as to facilitate forming an opinion re: The Inferential Question of Interest:

What is the Forecasting Model among the FM-Candidates that likely would be the best for the Dataset under examination?

4. OUR CoD-SCORING PROTOCOL

4.1 Overview There are a number of profiles posted in each of the CoD-Tables. We have selected ***eleven*** **P**oints of **I**nteresting **I**nferential **I**ntel [**PIIIs**] that are relatively independent and so are ideal in constructing a Practical Gestalt of each CoD-Table that will facilitate abstracting the essential details that will aid in gleaning the Forecasting Triage-intel of the study.

4.2 Scoring Codex for the [*PIIIs*] If the Non-directional FPE[Null[p-value]] for ***a*** PIII is < 0.1, then an asterisk [*****] will be affixed to that intel in the CoD-Table. For ***illustrative purposes***, we will use the following **CoD[*91.6%*]-Table**, for scoring of the **PIII**s.

Table 2A *Coefficient of Determination* {*CoD*[*Mean* [***91.6%***]] *& 95%CI* [*89.2% : 94.1%*]} ***A-Priori Suggested Expectation*****:** ***Strong*** *OLSR-Association of the Panel w. the Time-Index. Y-Variate Panel Transformation [**ln**].* We have Shaded the OLSR-Partition as it overall dominated the RW FM.

Triage Codex of the ***WRAE*** Category	*If the* **WRAE[OLSR/RW]** is **>1.0,** *then the preferred FM indicated is* ***The RW:*** [*t-Test*]		**Two-tailed z**[p-value test] Mean% ***v.*** RAE50%	*If the* **WRAE[RW/OLSR]** is **>1.0,** *then the preferred FM indicated is* ***The OLSR****:* [*t-Test*]	
Profiles n=14	*Mean*[*RW%*]	*95%CIs*	*Means v.50%*	*Mean*[*OLSR%*]	*95%CIs*
$t_{p=15}$[n=14]	42.9.%	13.2% : 72.5,	p=0.59	57.1%[p=0.61]	27.5%:86.8%
$t_{p=16}$[n=14]	21.4.%	0.0% : 46.0%	**p=0.03***	**78.6%***[p=0.03]	**54.0%:**100%*
$t_{p=21}$[n=14]	14.3%	0.0% : 35.3%	**p<0.008***	**85.7***[p=<0.003]	**64.6%:**100%*
Welch p-values	0.260	N/A	Both Means	0.260	N/A
Actual Counts	**RW**	***Narrative*** [All the Extrapolations] The overall p-value for the OLSR is: **0.002*.** This FPE rationalizes accepting H_a: that the OLSR outperformed the RW re: the WRAE. **PIII: 63.6%[7/11]**		**OLSR**	**Total**
$t_{p=15}$[n=14]	6			**08**	14
$\boldsymbol{t_{p=16}}$**[n=14]**	3			**11**	28
$\boldsymbol{t_{p=21}}$**[n=14]**	2			**12**	**42**

Note The [*] are significance markers that the p-values are < 0.1 which is our judgment of an interesting inferential degree of significance

Computations of the PIII-Scoring; CoD[*91.6%*]-Table 2A

[*i*] *The Overall Mean[RAE[%]] of the CoD-Table:* This will be computed as follows:

*The **overall*** RW Mean[RAE[%]] is *BHH-Qualified, thus, the computation will be:*

$$z_{cal} = [ABS[\text{The overall RW Mean[RAE[\%]]}] - 50\%] / \sigma_{50\%}]$$

where: $\sigma_{50\%}$ = ([50% × (1 − 50%)] / n_{Sample})^0.5

For the Computations, we have:

$$z_{Cal} = [ABS[31/42 - 50\%] / ((50\% \times (1 - 50\%) / 42)^{\wedge}0.5] = 3.086$$

where: The Non-Directional N[0,1]: 3.086 gives FPE[Null[p-value]] of: **[0.002]***:

This is so noted in the CoD[96.1%] ***Narrative***]. **This yields 1 [*]s,**

[*ii*] The Welch FPE[Null[p-value]] is: ***0.260***. As this is > our **PIII**-cutoff of 0.01 there will **no [*],**

[*iii*] *The 95%RAE:CIs* If The Lower Limit of the 95%RAE:CIs is > 50% then an **[*]** will be affixed to that CI.

This yields 2 [*]s.

[*iv*] *The Three Extrapolation FPE*[*Null*[*p-value*]]*s in Col*[*4*]. *For the* respective Mean[RAE[%]]s in Col **[5]**, there is the following ***Alert***: For the computation of these p-values for Col[4], we are only using the $\boldsymbol{z_{Cal}}$ and so reporting the N[0,1] p-value-calibration. We are not checking the BHH-screen! These computations are:

$t_{p=14}$[n=14] :N[0,1]-PDF is: ABS[57.1% −50%] / [(25% / 14)^.5] = **0.5313**; p-value: N[0,1]PDF= 0.59.

Note: **No [*]** was affixed as: 0.59 > 0.1

$t_{p=15}$[n=14] :N[0,1]-PDF is: ABS[78.6% −50%] / [0.13363] = **2.140**; p-value: N[0,1]PDF= 0.03.

Note: An **[*]** was affixed as: 0.03 < 0.1

$t_{=21}$[n=14] :N[0,1]-PDF is: ABS[85.7% −50%] / [0.13363] = 2.673; p-value: p-value: N[0,1]PDF= <0.008.

Note: An **[*]** was affixed as: <0.008 < 0.1

This yields 2 [*]s.

[*v*] *For the Three* *Mean*[*OLSR%*]*s in Col*[5] the p-values, using the Excel-Data Platform are:

$t_{p=14}$[n=14] :$t_{df=13}$-PDF is: ABS[57.1% −50%] / [0.13725] = 0.520; p-value: $t_{df=13}$= 0.61.

Note: **No [*]** was affixed as: 0.61 > 0.1

$t_{p=15}$[n=14] : $t_{df=13}$--PDF is: ABS[78.6% −50%] / [0.11380] = 2.511; p-value: $t_{df=13}$= 0.026.

Note: An **[*]** was affixed as: 0.026 < 0.1

$t_{=21}$[n=14] : $t_{df=13}$-PDF is: ABS[85.7% −50%] / [0.09705] = 3.680; p-value: $t_{df=13}$-]PDF= <0.003.

Note: An **[*]** was affixed as: <0.003 < 0.1

This yields 2 [*]s.

[*vi*] *Thus,* there will be, in total, eleven [**11**] p-value computations. For the CoD[***91.6%***] Table there were, in total of 7**[*]**. Thus, the last-entry of the CoD[96.1%] Narrative is **PIII: 63.6%[7/11].**

4.3 The Following CoD-Tables We have offered details on the computations of the various **PIII**s for Table CoD[***91.6%***]. Following, we will present the three CoD-Tables that complete the accrued Panel-sets: CoD[***76.4%***] & CoD[***37.1%***] & CoD[***5.1%***], Finally, we will offer our judgmental summary of the PIII-Profiles relative to the selection of the FM in the Candidate Set{ RW FM & The OLSR FM}.

Table 2B *Coefficient of Determination* {*CoD*[*Mean* [***76.4%***]] *& 95%CI* [*74.9% : 77.9%*]}***A-Priori Suggested Classification***: ***Likely*** *some OLSR-Association of the Panels with the Time-Index* ***as well as*** *instances where the RW has the relative advantage. Panel Transformation [****ln****].* We have Shaded the RW-Partition as it overall may have dominated the OLSR FM.

Triage Codex of the ***WRAE*** Category	*If the* **WRAE[OLSR/RW]** is **>1.0,** *then the preferred FM indicated is* ***The RW:*** [*t-Test*]		**Two-tailed z**[p-value test] Mean% ***v.*** RAE50%	*If the* **WRAE[RW/OLSR]** is **>1.0,** *then the preferred FM indicated is* ***The OLSR***: [*t-Test*]	
Profiles n=23	*Mean*[*RW%*]	*95%CIs*	*Means v.50%*	*Mean*[*OLSR%*]	*95%CIs*
$t_{p=15}$[n=23]	56.5[p=0.544]	34.6% : 78.4%,	0.532	43.5%	21.6%:65.4%
$t_{p=16}$[n=23]	47.8[p=0.840]	25.7% : 70.0%	0.835	52.2%	30.1%:74.3%
$t_{p=21}$[n=23]	60.9[p=0.306]	39.3% : 82.5%	0.297	39.1%	17.6%:60.7%
Welch p-values	0.679	N/A	Both Mean-Sets	0.679	N/A
Actual Counts	**RW**	***Narrative*** [All the Extrapolations] The overall p-value for the OLSR is: ***0.399***. This FPE rationalizes accepting H_o: that the RW did not outperformed the OLSR re: the WRAE. However, the RW V-Bound is in evidence. **PIII 0%[0/11]**		**OLSR**	**Total**
$t_{p=15}$ **[n=23]**	**13**			**10**	**23**
$t_{p=16}$ **[n=23]**	**11**			**12**	**46**
$t_{p=21}$ **[n=23]**	**14**			**9**	**69**

Discussion **The overall p-value** for this Dataset for the **RW-Arm** was:
z_{Cal} = [ABS[38/69 – 50%] / ((50% × (1 – 50%) / 69)^0.5] = 0.8427
where: The Non-Directional N[0,1]: 0.8427 calibration is: 0.399.
Overall Experiential Result The RW-Profile, while in the form of a ***RW V-Bound***, has no inferential support to rationalize this as evidence. This is also the Welch Overall inferential result. As this is at the cusp of the Investopedia™ cut-point of 70%:
"- - -. *A fund with a low R-squared, at 70% or less, indicates that the fund does not generally follow the movements of the index.".*
this result is not unexpected as the RW-Arm p-value is: 0.399, clearly in the Non-Rejection Region indicating that there no inferential Triage for the RW *vis-à-vis* the OLSR.

Table 2C *Coefficient of Determination* {*CoD*[*Mean* [***37.1%***]] *& 95%CI* [*34.4% : 39.8%*]} ***A-Priori Suggested Classification***: **Likely** *RW-Early in the Extrapolation-Space and later in the Extrapolation-Space the OLSR would likely have the relative advantage assuming that the Panels have a positive trajectory. Transformation [****ln****].* We have Shaded the RW-Partition as it overall dominated the OLSR FM.

Triage Codex of the ***WRAE*** Category	*If the* **WRAE[OLSR/RW]** is **>1.0,** *then the preferred FM indicated is* ***The RW:*** [*t-Test*]		**Two-tailed z**[p-value test] Mean% ***v.*** RAE50%	*If the* **WRAE[RW/OLSR]** is **>1.0,** *then the preferred FM indicated is* ***The OLSR***: [*t-Test*]	
Profiles n=69	*Mean*[*RW%*]	*95%CIs*	*Means v.50%*	*Mean*[*OLSR%*]	*95%CIs*
$t_{p=15}$[n=69]	**66.7%***[p=0.005]	55.3% : 78.1%*	0.006*	33.3%	21.9%:44.7%
$t_{p=16}$[n=69]	53.6%[p=0.548]	41.6% : 65.7%	0.550	46.4%	34.3%:58.4%
$t_{p=21}$[n=69]	**62.3%***[p=0.038]	50.6% : 74.0%*	0.041*	37.7%	30.0%:49.4%
Welch p-values	0.289	N/A	Both Mean-Sets	0.289	N/A
Actual Counts	**RW FM**	***Narrative*** [All the Extrapolations] The overall p-value for the OLSR is: ***0.0018****. This FPE rationalizes accepting H_o: that the RW did not outperformed the OLSR re: the WRAE. However, the RW V-Bound is in evidence. **PIII: 63.6%7/11**		**OLSR**	**Total**
$t_{p=15}$**[n=69]**	**46**			**23**	**69**
$t_{p=16}$**[n=69]**	**37**			**32**	**138**
$t_{p=21}$**[n=69]**	**43**			**26**	**207**

Discussion **The overall p-value** for this Dataset for the **RW** FM was:
z_{Cal} = [ABS[126/207 – 50%] / ((50% × (1 – 50%) / 207)^0.5] = 3.128
where: The Non-Directional N[0,1]: of 3.128 forms a FPE[Null]p-value]] calibration of: **0.0018*.**

Overall Experiential Result The RW offers suggestive evidence that the RW is dominant relative to the OLSR, and also that the RW is in the form of RW V-Bound. However, this RW-experiential profile does test to support a ***RW V-Bound*** profile as scripted above. In this case, of a relatively Random-Link of the Slope Driver of the OLSR with the Panel-Points, the RW ***emerges*** as the dominant model.

Table 2D *Coefficient of Determination* {*CoD*[*Mean* [***≈5.1%***]] & *95%CI* [*4.1% : 6.0%*]}***A-Priori Suggested Classification***: **Very Likely a** *RW and OLSR balance as the OLSR is more or less a random variate re: the Time-Index. Transformation* [***None***]. We have Shaded the RW-Partition as it overall ***experientially*** dominated the OLSR FM.

Triage Codex of the ***WRAE*** Category	*If the* **WRAE[OLSR/RW]** is **>1.0,** *then the preferred FM indicated is* ***The RW:*** [*t-Test*]		**Two-tailed z**[p-value test] Mean% ***v.*** RAE50%	*If the* **WRAE[RW/OLSR]** is **>1.0,** *then the preferred FM indicated is* ***The OLSR***: [*t-Test*]	
Profiles n=96	*Mean*[*RW%*]	*95%CIs*	*Means v.50%*	*Mean*[*OLSR*]	*95%CIs*
$t_{p=15}$[n=96]	**64.6***[p=0.004]	54.8% : 74.3%*	0.004*	35.4%	25.7%:45.2%
$t_{p=16}$[n=96]	41.7[p=0.103]	31.6% : 51.7%	0.103	58.3%	48.3%:68.4%
$t_{p=21}$[n=96]	**59.4***[p=0.066S]	49.4% : 60.4%	0.066*	40.6%	30.6%:50.6%
Welch p-values	**0.004***	N/A	Both Mean-Sets	0.004	N/A
Actual Counts	**RW FM**	***Narrative*** [All the Extrapolations] The overall p-value for the OLSR is: ***0.077*.*** This FPE rationalizes accepting H_o: that the RW did not outperformed the OLSR re: the WRAE. However, the RW V-Bound is in evidence. **PIII: [63.6%[7/11]**		**OLSR**	**Total**
$t_{p=15}$**[n=96]**	**62**			**34**	**96**
$t_{p=16}$**[n=96]**	**40**			**56**	**192**
$t_{p=21}$**[n=96]**	**57**			**39**	**288**

Discussion **The overall p-value** for this Dataset for the **RW** was:
z_{Cal} = [ABS[159/288 – 50%] / ((50% × (1 – 50%) / 288)^0.5] = 1.768
where: The Non-Directional N[0,1]: of 1.768 forms a FPE[Null]p-value]] calibrations of: ***0.077****

Overall Expectation The RW FM tests to be interestingly more effective in the forecasting context than was the OLSR FM for the RAE-Measure where the CoD is ≈5%. Again, there was ***RW V-Bound*** Profile-configuration; it passes the Welch-test. We will make the Tukey-Kramer HSD-assessment, *anon*.

5. TRIAGE ANALYSIS: THE SYNTHESIS & DE-CONSTRUCTION & SUMMARY

5.1 Overview We evaluated Two Forecasting Model Candidates: {The OLSR TS & The Random Walk}, the intention being to use the intel provided by *Their RAE-Protocol Profiles* [*controlling for three specific Forecasting Extrapolations*] to develop Efficient and Effective forecasting protocols. At this point, we ***recommend*** creating the following critical elements to ***probe*** the four CoD-Profiles above to make a reasoned selection of the FM to use in the coming forecasting cycle**.** In this regard, the Analyst should:

I. **Score** the **PIII**-Variables of the CoD-Profiles produced by The RAE FM[Triage] Protocol—such as those displayed in the above four CoD-Profiles. ***Rationale*** This will allow informed Analysts to use their Experiential-intel to create meaningful information in forming the plan for the deployment of the FM-selected from the FM-Candidate Set, then

II. **Populate** the following three Tables:
Table 3: *The RAE FM Triage* ***Initial Synthesis*** *Capsule, &*
Table 4: *Tukey-Kramer [HSD] Ordered Difference Report*: ***De-construction*** *Capsule &*
Table 5: *Initial Synthesis Summary and* ***Revision***: *The* ***Final*** *Triage Capsule*

Rationale These three Tables form an integrated-Lattice where there is intel that is used to form the next Table. This integrative Lattice-Probing is fundamental to arriving at a reasoned judgement of the FM that is best matched with the CoD-Panel. We find these *Tables* to be essential in coalescing the **PIII**-intel in the CoD-Profiles into useful guidelines to arrive at a reasoned selection of the FM best aligned with the: (i) The Nature of the Data, and (ii) The Planned Extrapolation-set.

Let us now discuss these Tables and their profiles.

5.2 The Vetting Screens & Recalibrations & Revisions Most all of the Tabular-intel that is presented in the four CoD-Profiles can be captured by focusing on the eleven **PIII** [*]-results—which was our Judgement-Screen. With this information, *The RAE FM Triage **Initial Synthesis** Capsule* will be populated by the Analysts as their ***initial*** Impressions of the **PIII**-scored RAE-Profiles. Following is Table 3 that we have populated:

Table 3 *The RAE FM Triage **Initial** Synthesis Capsule*

CoD Group	Sum Vetted	Extrapolation $\{t_{15} : \boldsymbol{t_{16}}: t_{21}\}$	%RW Preferred	%OLSR Preferred	Synthesis & Codex by Panel **PIII** x-Score[[Overall [**x**/11]]]
[91.6%]	Yes	**OLSR** >> RW [57%<79%<86%]*	No 26%	**Yes 74%**	**Possibly Transitive for: OLSR** Inferential Profile: **63.6% [7/11]**
[76.4%]	Yes	**RW ≈ OLSR** [57%>***47%***< 61%]	**Yes 55%**	No 45%	**A Standard Profile for: RW or OLSR** & A possible RW V-Bound Profile **[0% [0/11]**
[37.1%]	Yes	**RW** >> OLSR [67%>***54%***<62%]	**Yes 61%**	No 39%	**Likely Transition to: RW >> OLSR** & A possible RW V-Bound Profile **[63.6% [7/11]**
[5.1%]	Yes	**RW** >> OLSR [65%>***42%***<60%]	**Yes 55%**	No 45%	**Modified State of Nature: RW v. OLSR** & A possible RW V-Bound Profile **[63.6% [0/11]**
Total Scored PIII for the RAE					**The Scored PIIIs = 47.8% 21** [7+0+7+7] of 44

*We have Scripted the "likely" Nature of the Profile using the directional-conditionals [< & >]. At this point they have no inferential context.

Discussion We offer Table 3 as the ***Synthesis*** of the **PIII**-scored elements that we profiled as intel of the RAE-measure used by *The RAE FM*[*Triage*] *Protocol* re: The Triage of the Forecasting Models in the Candidate-Set: {The OLSR & The RW} FMs. Our ***Synthesis*** is:

> The experiential evidence for these four CoD-Profiles is that they are ***likely** in sync* with the Math/Stat drivers of the RW- & OLSR-generating process(es) as profiled in CoD-Tables of the M-Study. Specifically, we note:

Both the M-Study of 1,001 TS-Panels ***and*** our four CoD [S&P500]-Panels presented above, have, in the main, RAE-Profiles that are ***indicative*** of the following:

In the ordered-Extrapolation-Space for these Tables,

Initially, for Extrapolation[t_{n+1}], the Actual Values Y_{n+t} are more aligned with the RW FM than they are aligned with the OLSR FM, ***then***

for Extrapolation[t_{n+2}], there is a transition where the Actual Values Y_{n+t} are more aligned with the OLSR FM, than they are aligned with the RW FM, ***finally***

for Extrapolation[t_{n+7}], the Actual Values Y_{n+t} are ***again*** more aligned with the RW FM than they are aligned with the OLSR FM.

This has the ***same profile*** as was observed for the M-Series in Table 1 labeled as a ***RW V-Bound***.

To explore the very interesting set of profiles in the CoD-Tables, we offer the following extended inferential analyses.

5.2.1 The Integrated Re-Vision Lattice: Vetting Intel

Experientially, we have learned, over many decades, the relevance and utility of the **Carpenter's Rule: "Measure Twice—Cut Once**." For this reason, we value forming a vetting-opinion from the above, ***Stage I***, ***Synthesis*** of the CoD[**PIII**s] profiles as presented in

Table 3. The Carpenter's Rule is our ***Stage II*** vetting of Table 4 using: The Tukey-Kramer [HSD] Ordered Difference ***De-Construction & Re-assessment*** Profile. Our experience has been that there is a profound analytical benefit from the following Integrated Re-***Vision*** Lattice:

> ***Stage I***: The scored **PIII**s of the CoD-Tables are used to from a judgmental ***Synthesis*** of each CoD-ordered Collective: This is noted in Column [Synthesis & Codex by Panel **PIII** x-Score[Overall [**x**/11]]] of Table 3. This is the ***initial***-assessment, and
>
> ***Stage II***: This is the ***vetting*** of Stage I. Experientially, there seems to be a benefit to forming an opinion and then re-visiting that initial investigative opinion by re-assessing the Stage I Intel.

5.2.2 Illustrations of the ***Stage II Vetting*** for the CoD-Collectives

In this case, we will probe the Tukey-Kramer [HSD]-intel for all four of the CoD-Panels. This HSD-information is presented in Table 4 following. Then, we will create the information that we deem necessary to inferentially-vet the Stage I intel that was used to create the Synthesis Table 3. Finally, we will present a ***Re-visions***: Table 5 where we indicate any and all revisions in our Table 3 Synthesis-opinions.

Table 4 *Tukey-Kramer [HSD] Ordered Difference Report*: *De-construction Capsule**

CoD=Panels	ABS[RAE Difference]	T-K [p-value]	Implication: re: Mean[RAE]
CoD[***91.6%***]	Extrapolation[OLSR [57%: 79%: 86%]		***Transitivity Not in Evidence***
t_{21} v. t_{15}	[0.857] – [0.571]= 0.286	0.21	Accept H_a: t_{21}^{RAE} likely >t_{15}^{RAE}
t_{16} v. t_{15}	[0.786] – [0.571]= 0.215	0.41	Accept H_a: t_{16}^{RAE} may be >t_{15}^{RAE}
t_{21} v. t_{16}	[0.857] – [0.786]= 0.071	0.90	Accept H_o: t_{21}^{RAE} clearly = t_{16}^{RAE}
CoD[76.***4%***]	Extrapolation[RW [57%: *48%:* 61%]		***RW V-Bound NOT in Evidence***
t_{21} v. t_{15}	[0.609] – [0.565]= 0.044	0.95	Accept H_o: t_{21}^{RAE} clearly = t_{15}^{RAE}
t_{16} v. t_{15}	[0.478] – [0.565]=– 0.087	0.83	Accept H_o: t_{16}^{RAE} clearly = t_{15}^{RAE}
t_{21} v. t_{16}	[0.609] – [0.478]= 0.130	0.66	Accept H_o: t_{21}^{RAE} clearly = t_{16}^{RAE}
CoD[37.1***%***]	Extrapolation[RW [67%: *54%:* 62%]		***RW V-Bound NOT in Evidence***
t_{15} v. t_{16}	[0.667] – [0.536]= 0.130	0.26	Accept H_a: t_{15}^{RAE} likely >t_{16}^{RAE}
t_{21} v. t_{16}	[0.623] – [0.536]= 0.087	0.55	Accept H_o: t_{21}^{RAE} likely = t_{16}^{RAE}
t_{15} v. t_{21}	[0.667] – [0.623]= 0.044	0.86	Accept H_o: t_{15}^{RAE} clearly = t_{21}^{RAE}
CoD[5.1***%***]	Extrapolation[RW [65%:> ***42%***: <60%]		***RW V-Bound in Evidence***
t_{21} v. t_{15}	[0.594] – [0.646]= 0.052	0.742	Accept H_o: t_{21}^{RAE} likely = t_{15}^{RAE}
t_{15} v. t_{16}	[0.646] – [0.417]= 0.229	**0.004**	Accept H_a: t_{15}^{RAE} clearly > t_{16}^{RAE}
t_{21} v. t_{16}	[0.594] – [0.417]= 0.177	**0.034**	Accept H_a: t_{21}^{RAE} clearly> t_{16}^{RAE}

Note:** The T-K[HSD] inferential profiler does not check for Population Order Requirements; and, The T-K FPE[Null[p-value]] is Non-directional—the Directional-conversion is: ***The [T-K[p-value]]/2. We used the SAS™[JMP™[v.13] for this test. In our testing profile these results are conservative as our sample sizes are equal over the Extrapolations for particular CoD-Panel sets.

Discussion To continue the Probing/Vetting of the Synthesis Profiles [Table 3], we will examine the T-K[HSD]-profiles as a ***vetting*** of the Synthesis Profiles of Table 3 for the four CoD-**PIII** Tables. Following, as an illustrative computational illustration, we will note the inferential steps that seemed to be needed. Other Analysts may certainly have used others. Recall we have formed a set of conditions for the judging the inferential evidence for the **PIII**-Profiles.

CoD[**91.6%**] In considering the **PIII**-intel of Table 2A for the CoD[***91.6%***], we noted that this profile [57% < 79% < 86%] was: **Possibly Transitive** for the **OLSR FM.**

[HSD]:Vetting Given the directional p-value intel offered in the T-K[HSD] Table, and the related assessment in Col[*Implication: re: Mean[RAE*]], we have modified our initial **Stage I** conjecture as follows:

> *We proffer that the **Nature** of the Inferential **Transitivity** for the two ordered paired comparisons*:
> $\{t_{15}[RAE] < t_{16}[RAE]\}$ & $\{t_{16}[RAE] < t_{21}[RAE]\}$
> *will be **Strong iff** the directional T-K[p-values] are: <0.1, or*
> ***Suggestive iff** the directional T-K[p-values] are: <0.25.*

Analysis *The directional T-K[p-values] do not meet these vetting-tests as indicated in Col*[*3*]. Thus, we note: There is **No Revision** of our initial analysis. **Final Analysis**: ***Not suggestive of a Positive-Transitive-Trajectory over the Three Extrapolations.***

CoD[**76.4%**] In considering the **PIII**-intel of Table 2B for the CoD[*76.4%*], we noted that this profile [[57% > *48%* <61%] was: **A Standard Profile for RW ≈ OLSR. We made this assessment as there were no PIIIs that were inferentially interesting over the Extrapolations.**

After examining the T-K[HSD] Table 4, we did ***not*** modify our conjecture. ***Rationale***:

> The CoD[**76.4%**[profile was: [[57% > *48%* < 61%] This is, in profile, not unlike a ***RW V-Bound***. Thus, the question of interest is: *Is there inferential support for a **RW V-Bound***? **Results** The Welch[p-value] was > 0.25 and there is no T-K inferential support for a ***RW V-Bound*** as all of the directional T-K p-values are > 0.25. Thus, we note: There is **No Revision** of our initial analysis. **Final Analysis**: **A Standard Profile for RW ≈ OLSR**

For the CoD[**37.1%**]:[67% > *54%* < 62%], initially we noted: **Likely Transition to: RW >> OLSR. In addition,** we observed a possible indication of a ***RW V-Bound.*** **Results** For the RW V-Bound profile, the Welch[p-value] is > 0.25 as well as the respective directional T-K[p-values] for the [t_{21} v. t_{16}] are > 0.25. Thus, we note: There is ***No Revision*** of our initial analysis. Final Analysis: **Likely Transition to: RW >> OLSR.**

Finally, for the CoD[**5.1%**]:[65% > *42%* < 60%], we observed that there are indications of a ***RW V-Bound***. All of the conditions for a ***RW V-Bound*** are satisfied. Specifically, the Welch [p-value] is < 0.1; the directional T-K[p-values] for a RW V-Bound are both < 0.1. This offers a ***Strong*** indication of a ***RW V-Bound***. Thus, we note: There is a Revision of our initial analysis. ***Final Analysis***: **Strong Indication of a RW V-Bound.**

Given the above T-K[HSD]information and our re-assessment calculations we offer the following:

Table 5 *Initial Synthesis Summary and **Revision**: The **Final** Triage Capsule*

CoD	Synthesis **after** the CoD[**PIII**]-Analysis	Table 5 De-Construction & ***Revision***
CoD[**91.6%**]	**Possibly Transitive for the OLSR** Inferential Strength: [63.6% [**7**/11]]	OLSR dominated the RW; **Transitivity was not inferentially in evidence.**
CoD[**74.6%**]	**A Standard Profile for RW or OLSR** Inferential Strength: [0% [**0**/11]	We considered a ***RW V-Bound***; however, there was **no** revision. **A Standard Profile for RW or OLSR.**
CoD[**37.1%**]	**Likely Transition to: RW >> OLSR** Inferential Strength: [63.6% [**7**/11]	We considered a ***RW V-Bound***; however, there was **no** revision. **Likely Transition to: RW >> OLSR**
CoD[**5.1%**]	**Modified State of Nature: RW v. OLSR** Inferential Strength: [63.6% [**7**/11]	The CoD[***5.1%***] is judge as having a **Very Strong** indication of a **RW V-Bound**.

RAE[Triage] FM-Candidates: {RW & OLSR} Suggested Triage for Forecasting	
CoD[**91.6%**]	FM-RAE-Triage {The clear choice is The ***OLSR FM***. Likely No Transitivity}
CoD[**76.4%**]	FM-RAE-Triage {**Use Occam's Razor**: Select ***The RW FM***}
CoD[**37.1%**]	FM-RAE-Triage {Both FMs: {***RW & OLSR***} will likely be engaged in producing forecasts}
CoD[**5.1.%**]	FM-RAE-Triage {Both FMs: {***RW & OLSR***} Clear evidence of a ***RW V-Bound***}

6. THE *APRÈS-ANALYSIS*: OUR RECOMMENDATION TO THE PLANNING COMMITTEE:

Our *suggestions* for the probing phases of: The Integrated Re-Vision Lattice Intel-Lattice:

I. *Review* of the **PIII**-Scored [CoD-Tables],
II. Create Table 3: The ***Synthesis*** as profiled in: *The RAE FM Triage Initial Synthesis Capsule*, and
III. Create Table 4*:* The ***De-construction & Revision*** of the above selected intel using: Table 4 Tukey-Kramer [HSD] *Ordered Difference Report*: *De-construction Capsule*, and
IV. *Create* Table 5: *Initial Synthesis Summary and* ***Revision****: The Final Triage Capsule* ***for*** the FM-Candidates: {RW FM & OLSR FM}

We offer these steps as reasonable guidance for the Tirage of a FM from the: Forecasting Model Candidates: {RW ***&*** OLSR}.

7. OUTLOOK

7.1 The Triage Protocol A few Caveats and Reality-Checks The RAE FM[Triage] Protocol seems a simple and specific screen for ferreting-out the Forecasting Model Candidates that are best used for the collection of Panels to be used in creating forecasts for the Planning Committee. We elected to group the Panels by their CoD-profiles. This was to test if the CoD-groupings affected the Triage. It certainly seems to have had effects; however, while the Math/Stat expectations seemed logical to us, overall they were rarely specific and so marginally effective as Expectations. This may call into question the information utility of musing over the Math/Stat profiles as a pre-cursor activity. Also, it does not seem that the grouping of the Panels by their CoD-profiles is required to use The RAE FM[Triage] Protocol. Actually, rarely will a Firms have sufficient Y-Variate Panels that CoD-grouping would be practical. In the past, for consultations, we simply used the datasets that the firm had collected. This was not a problem regarding the utility of the Triage results. However, we are committed to using the Forecasting Models in the Candidate-Set: {The OLSR & The RW} FMs as the initial Triage-screening. Most often these simple models outperform most of the other FMs in the panoply of the Forecasting Division on most of the usual forecasting acuity measures. Also, if the RAE is to be used as the Triage-Measure, we strongly recommend using the binary-category coding of the results of the individual panels—**Alert** Parametric Profiles of the RAE-results are an introduction into: *The Multiverse of Madness and Confusion*.

7.2 The Inferential Intel The ***RW V-Bound*** profiles for the M-Study were not vetted or tested inferentially by the Makridakis-collective or by us. The profiles of the OLSR[RAE] & RW[RAE] in CoD-Table**s** that, were inferentially tested, are not, in some cases, dissimilar to the general profile of the ***RW V-Bounds*** of the M-Study. Both are intriguing and beg an investigation. A *valued colleague* in providing feedback to a Draft of this Research Report noted[*verbatim*]:

> *I agree a "110%" that the M-Study 1,001-Panels for the OLSR & The "RW" FMs very likely resulted in what you are calling a* ***RW V-Bound****. This is indeed an interesting result. However, we are on the same page—How the RW V-Bound is a Population feature of your CoD-Panels belies any logic that I am able to summon-up to explain how this could* ***also*** *be a feature of your Panels—in*

> *particular the CoD[5.1%] Panel???!!! So, the M-Study RW V-Bound seems not unreasonable, however, IS IT POSSIBLE that your RAE-Screening Protocol* ***in fact fallaciously*** *created RW V-Bounds from your CoD-Panels? It is hard to imagine that your RAE-Screens have a "Feature" or Latent Bias that would create RW V-Bounds from your Panels—but this is the only creative thought that I am left with!!!*

True and well expressed. We anticipate, in a *Poirot-esque* Manner, the probing and possible resolution of these questions. If our readers are interested in our findings, please let us know; we will put you in our "Working Paper Loop"!

REFERENCES

Armstrong, J.S., & Collopy, F. (1992). The selection of error measures for generalizing about forecasting methods: empirical comparisons. *International Journal of Forecasting*, *8*, 69–80.

Box, G. E. P., & Cox, D. (1964). An analysis of transformations. *J. Roy. Statist. Soc: B*, *26*, 211–252.

Box, G. E. P., Hunter, W., & Hunter, J. S. (1978). *Statistics for Experimenters*. J. Wiley & Sons. Inc. ISBN-9780471093152.

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. https://doi.org/10.1287/mnsc.38.10.1394

Makridakis, S., Andersen, A., Carbone, R., Fildes, R., Hibon, M., Lewandowski, R., Newton, J., & Parzen, E. (1982). The accuracy of extrapolation (time series) methods: Results of a forecasting competition. *Journal of Forecasting*, *1*, 111 – 153. https://doi.org/10.1002/for.3980010202

Tamhane, A., & Dunlop, D. (2000). *Statistics and data analysis*. Prentice-Hall. ISBN 0-13-744426-5.

Theil, H. (1958). *Economic Forecasts and Policy*. North Holland Press, Amsterdam.

APPENDIX

The inferential profile of the Standard Format of this RAE-configuration of Nine Panels

Panels	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
Test	FMc/ FMs	FMc/ FMs	FMc/ FMs	FMc/ FMs	FMc/ FMs	FMc/ FMs	FMc/ FMs	FMc/ FMs	FMc/ FMs
RAE	**4.000**	**5.000**	**7.000**	**2.000**	1.000	**0.25**	**0.2**	**0.142857**	**0.5**
Prefer	FMs	FMs	FMs	FMs	**Either**	FMc	FMc	FMc	FMc

Where: ***FMc*** is The More Complicated FM & ***FMs*** is the Simple FM