Technomics
Better Decisions Faster

Well, That Escalated Quickly – A Novel Approach to Forecasting Escalation

Sean Wells - Swells@technomics.net

Kaitlyn Hagy - Khagy@technomics.net

International Cost Estimating and Analysis Association

2/20/2024

Abstract

Escalation rates are an important part of estimates, and as such, the provenance and derivation of indices should be regularly scrutinized, yet are rarely contemplated. This paper will compare a commonly used black-box escalation resource, IHS Global Insight, to a traceable, simplified forecasting method to determine if a purely mathematical model delivers improved forecasting accuracy. Our model relies on a curated set of Bureau of Labor Statistics (BLS) indices to develop a moving average forecast. With access to over 15 years of IHS forecasts dating back to 2006, spanning 800+ indices, this study has the unique opportunity to quantify the accuracy of IHS and moving average forecasts against historical BLS indices. Our paper will establish and explore various measures of forecast accuracy for use in creating defensible estimates. The goal is to provide a quick, transparent, and flexible way to develop tailored escalation projections without sacrificing accuracy.

Keywords: *Escalation, Risk, Modeling*

Table of Contents

Table of Figures

Table of Tables

1. Introduction

1.1 Applying Escalation Without Scrutiny

Escalation rates represent a crucial component of nearly every cost estimate. Whether applied broadly to a diverse set of materials or narrowly to a single component, escalation has a significant impact on projected costs. Forecasts influenced by escalation rates can span from several years in the case of a contract to decades in the case of a life-cycle cost estimate or sustainment review. Figure 1 demonstrates inflation fluctuations over the past decade, a period in which innumerable defense projects, construction ventures, and research activities were contracted.

Figure 1 – 2013-2023 Inflation Rates

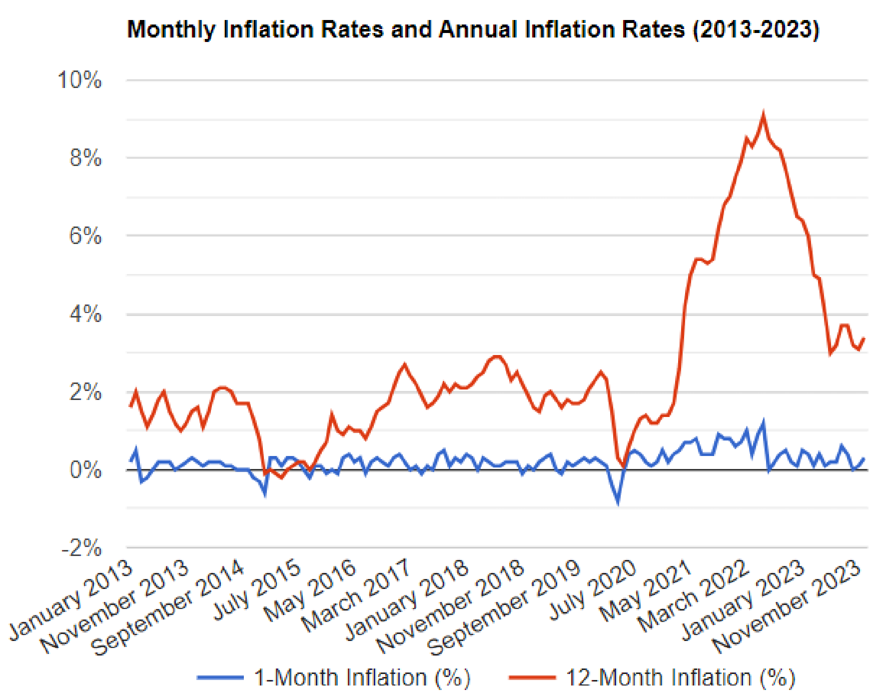


Figure 1 also poignantly alludes to how more recent economic phenomena, such as disruptions in supply chains—whether induced by the pandemic or resulting from poorly implemented logistics—have raised concerns within the estimating community about increased volatility of forecasts, implications on economic price adjustment (EPA) clauses, and the outdated concept of a flat percent year-to-year increase without any basis. In response, the Office for the Undersecretary of Defense (OSD) issued the following statement: “The current economic environment requires we understand the impacts of inflation to existing contracts and consider various approaches to manage risk of inflation to prospective Department of Defense (DoD) contracts” (OSD, Guidance on Inflation and Economic Price Adjustments, 2022).

Despite these growing concerns, cost estimators treat escalation as a check-box exercise. Escalation indices are applied with minimal-to-no consideration for escalation-specific uncertainty due to a lack of applied research measuring forecast accuracy for various indices and forecast horizons. Cost analysts can and should improve the credibility of their estimates via discrete treatment of escalation uncertainty based on historical forecast accuracy.

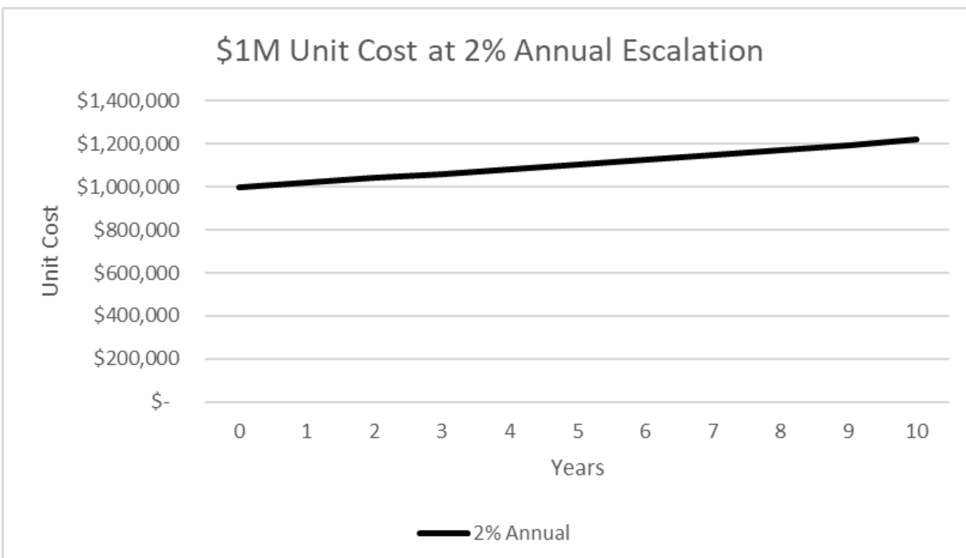
1.2 The Process for Measuring Accuracy

Our paper addresses this paradigm shift by applying three key steps: (1) collecting extensive historical data, (2) constructing a dynamic model to analyze this data, and (3) deriving meaningful initial results. This paper leverages over a decade of historical quarterly Global Insights (formerly IHS) escalation forecasts. These records are complemented with a simple moving average model based on Bureau of Labor Statistics (BLS) indices. Finally, we compare each forecasting model (simple moving average and Global Insights) against the actual BLS data to calculate error statistics over various prediction ranges and sample indices.

The simple moving average model is the foundational first step towards advancing to more complex models such as Exponentially Weighted Moving Average (EWMA) and Autoregressive Integrated Moving Average (ARIMA) models. The novelty of our approach lies in our ability to measure the forecast accuracy of a widely used source and compare it to “simple math” to determine the level of confidence an estimator should assign to their applied escalation forecasts. This methodology facilitates a data-driven approach to developing uncertainty factors by analyzing different commodity indices (ranging from fuels to semiconductors) and assessing accuracy ranges. Our simple moving average model and the subsequent escalation accuracy analysis reduces reliance on black box commercial models and leverages the authoritative source of inflation data to form the foundation for more complex future risk models and escalation projection techniques.

Before we dive headlong into models, data, and more, put yourself in the shoes of a 2011-era Contracting Officer (CO) procuring a navigation component with a unit price of one million dollars. At that time, a flat 2% year-to-year increase was the standard escalation approach, so if the CO needed to purchase a unit at five and ten years on a sustainment contract, that methodology would clear all leadership reviews. By leveraging a decade and a half of collected escalation data, this paper will determine just how risky that 2% assumption could be. Figure 2 visualizes this consistent escalation, with a 5-year unit cost of $1.104M and a 10-year unit cost of $1.219M.

Figure 2 – Example Inflation Application: 2% Annual Inflation Visual

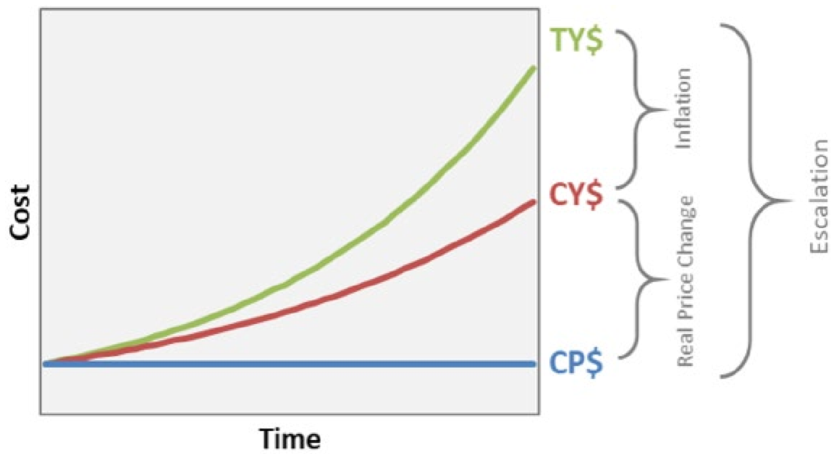


2. Background

2.1 Escalation vs. Inflation

Inflation represents an economy-wide increase in the average price level (affecting all prices in the same proportion). In contrast, **escalation** reflects price changes for specific goods and services (OSD, Inflation and Escalation Best Practices for Cost Analysis, 2021). Real Price Change refers to the concept that the price for a specific good or service might change differently than an economy-wide collection of goods and services. Escalation is the combined effect of inflation and real price change, which incorporates some value for inflation, a value for real price change, and a value for inflation on the real price change (see Figure 3).

Figure 3 – Inflation, Real Price Change, and Escalation



In the context of this paper, the cited indices are escalation indices for specific commodities, thus incorporating both economy-wide inflation and specific fundamental price changes for given commodities. Another key term is the Composite Annual Growth Rate, or "CAGR," a standard measure of the escalation rate over a given period (often ten or twenty years in financial forecasting).

$$CAGR = \left(\left(\frac{EV}{BV}\right)^{\frac{1}{n}} - 1\right) \times 100$$

where:

EV = Ending value

BV = Beginning value

n = Number of years

2.2 Considerations in Escalation Forecasting

Any evaluation of escalation forecast accuracy differences between models begins with understanding what information and factors feed these models. While the exact recipe for a Global Insights or Federal Reserve forecast is proprietary, some factors are generally accepted to contribute to the forecast. An informal search on "causes of inflation" yields results such as "supply shocks," "money supply," "expectations," and "the unemployment rate," among others (Frick, 2022). Real Price Change can be influenced by a variety of factors, including the supply of specific materials, economies of scale, workforce mix, and technological change (OSD, Inflation and Escalation Best Practices for Cost Analysis, 2021). Certain agricultural commodities may rely on weather forecasts, while a projected leadership change in an oil-rich nation could impact fuel price forecasts. Demand fluctuations, such as consumer fads for electronics, may cause subcomponent prices to spike.

Considering the breadth of factors influencing escalation rates, model inputs can be boiled down to two categories: (1) factors that have historically occurred and are projected forward (i.e., perceived rate of technological change), and (2) factors that may have some historical analogies but are reliant on scheduled future events (i.e., how predictions of upcoming elections may affect economic policy). A commercial model realistically leverages some combination of these factors, consistently revising the inputs to produce current forecasts. Contrast this with our simple moving average model, which is based solely on backward-looking historical changes in escalation indices. Ultimately, including recent history encompasses all the aforementioned historical factors (supply, rate of technological change, rate of market consolidation, overall volatility, etc.) in one composite number (the rate of change between index values).

The downside of such a “fair-weather” model is that future events cannot be discretely accounted for, and the model may be slow to respond when moving between inflationary paradigms. For example, if a three-year moving average period did not include the last major world conflict, then the impact of a future conflict has no predictive basis in the data. Fundamentally, any model with additional drivers beyond mere historical values should be more appropriate to use as a predictive model.

This paper seeks to test this assumption by comparing the accuracy of the Global Insights model (considering both forecasting categories previously referenced) with a simple moving average model (considering only backward-looking factors in forecasting). In reality, neither model can account for a third category of events: the “unknown unknowns,” such as a global pandemic, a cargo ship being stuck in a canal, or an assassination of a world leader. Commercial models might have some catch-all risk factors to account for this type of scenario, but our forecast comparisons will shed light on whether said factors really make commercial models less prone to forecast error.

3. Data Source Overview

3.1 Choice of Sample Indices

We selected a subset of ten indices out of the hundreds available to narrow the scope of this paper during model development and focus the statistical analysis. The indices shown below in Table 1 encompass a range of standard cost-estimating commodities. For the remainder of this paper, any reference to indices will be based on the "Short Title" naming convention.

Table 1 – Sample Escalation Indices for Trials

Category	Title	Short Title	BLS Reference Number	Global Insights Reference Number
Energy	Gas Fuels	Gas	WPU053	WPIP053S
AHE	Manufacturing	Manufacturing	CEU3000000008	CEU3000000008
Steel	Iron and Steel	Steel	WPU101	WPIWP101
Fabricated Metals	Aluminum Cans	Aluminum	WPU103103	WPIP103103
Chemicals	Industrial Chemicals	Chemicals	WPU061	WPIP061
Building Materials	Lumber and Wood Products	Wood	WPU08	WPIP08
Electrical Components	Semiconductor & Other Electronic Components	Semiconductors	PCU33441-33441-	PPI3344
Transportation Eqpt	Other Aircraft Parts & Auxiliary Equipment	Aircraft Parts	PCU336413336413	PPI336413
Paper	Corrugated and Solid Fiber Boxes	Paper Boxes	PCU3222113222110	PPI322211
Defense & Aerospace	Search, Detection & Navigation Instruments	Defense Instruments	PCU334511334511	PPI334511

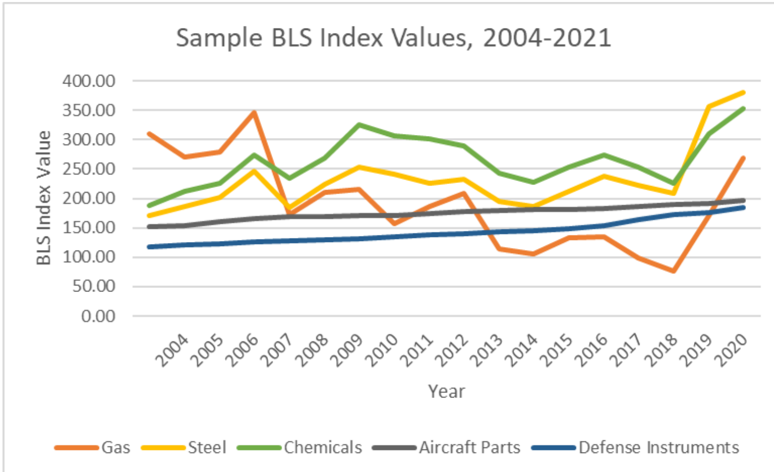
We curated a diverse sample of indices and strongly considered the relative volatility of each index. We selected a representative range of perceived and statistical volatility since moving average projections (and to an extent commercial models) are inherently influenced by historical volatility. Table 2 shows the Gas index having a 9.6% year-to-year average change with a 2.5% CAGR over the 14-yr data span. By contrast, Aluminum has a similar CAGR but only a 2.9% average year-to-year change, indicating that Aluminum is a less volatile commodity. Semiconductors show negative values due to the overall price of semiconductors falling between 2008 and 2022.

Table 2 – Sample Indices CAGR and Volatility Measures

BLS Index	Short Title	CAGR	Mean Monthly % Change	Mean Annual % Change
WPU053	Gas	2.5%	0.6%	9.6%
CEU3000000008	Manufacturing	2.8%	0.2%	2.5%
WPU101	Steel	3.1%	0.3%	6.9%
WPU103103	Aluminum	2.4%	0.2%	2.9%
WPU061	Chemicals	2.0%	0.2%	4.3%
WPU08	Wood	3.5%	0.3%	4.1%
PCU33441-33441-	Semiconductors	-1.0%	-0.1%	-1.4%
PCU336413336413	Aircraft Parts	1.3%	0.1%	1.4%
PCU3222113222110	Paper Boxes	4.6%	0.4%	4.5%
PCU334511334511	Defense Instruments	2.9%	0.2%	2.7%

Policymakers consider some indices more volatile than others, so the “core inflation” calculation often excludes items like fuel and food that are subject to significant price changes (Clark, 2021). Figure 4 shows several sample indices over time, some with peaks and valleys and others trending smoothly. We expect the moving average model to demonstrate improved accuracy on less volatile commodities (i.e., Aluminum, Defense Instruments, Aircraft Parts) and be less accurate for more volatile ones (i.e., Gas, Steel, and Chemicals). As alluded to, this expectation holds for commercial models as well, however, with a different (lower) magnitude of error (i.e., improved accuracy compared to the moving average model).

Figure 4 – Sample BLS Index Values, 2004-2021



3.2 BLS Indices

BLS data was collected from the BLS website for 2004 to 2022 for the previously identified ten indices. This date range was chosen to align with the available Global Insights date ranges (2006-2022), with a couple of additional years pulled for context. Each index contains monthly historical record values and we averaged each year's January through December figures to derive annualized values.

BLS indices (and escalation indices in general) are unitless and represent the change in overall commodity value since the index started (noted for each index on the BLS website with a starting year and value). Therefore, the comparison of index values across indices is a meaningless exercise. However, comparisons across periods within a specific index are tremendously insightful. BLS data is widely used by Original Equipment Manufacturers (OEMs) in proposal estimates, whether leveraging historical data to derive commodity-specific indices or merely applying BLS escalation forecasts to determine bid pricing.

3.3 Global Insights Indices

Global Insights data is published quarterly and provides details regarding the specific timeframe of the data and forecasts. Our team collected these historical files stored by NAVSEA Contracting Officers in the form of MS Excel files dating back to 2006. Each file contains quarterly and annual forecasts for a variety of indices. Global Insights forecasts approximately 2.5 years out for quarterly predictions and ten years out for annual predictions. Since the 2006-2007 data is incomplete relative to the subsequent years, our analysis focuses on the 2008-2022 timeframe. Global Insights forecasts are particularly suitable for comparison because they are widely used in DoD cost estimates and endorsed by government agencies and contractors. Even the OSD CAPE handbook cites and recommends their use (OSD, Inflation and Escalation Best Practices for Cost Analysis, 2021).

4. Method Overview

4.1 Moving Average Theory

A simple moving average (SMA) is defined by the equation predicting “Y” at time “t+1” using “m” periods of data up to time “t”, as demonstrated below (Nau, 2014):

$$\hat{Y}_{t+1} = \frac{Y_t + Y_{t-1} + \dots + Y_{t-m+1}}{m}$$

In simpler terms, a moving average is defined by how many prior periods of data are used to inform future predictions. To frame this, consider the two extremes explained below:

First, consider using every prior data point (greatest possible m), which could date back to 1985 in many index cases. This incorporates data that might be too far removed from current circumstances (for example, pre-internet, before the end of the Cold War, etc.) and incorporates such a high quantity of data points that period-to-period differences are smoothed out and functionally eliminated.

Second, consider using only the prior period’s data point (m = 1). This scenario assumes no change in the future from the current state. As comforting as this level of stability would be, overwhelming historical examples (and the second law of thermodynamics) demonstrate that change is inevitable. Thus, the “random walk” model is not an adequate predictor.

The moving average is the Goldilocks between these two extremes, taking a value of m that encompasses enough historical data to be predictive without accepting too much “noise.” Our paper analyzes the use of one-, two-, and three-year moving average ranges, considering the effectiveness of the sliding scale from more recency (1 year) to greater historical stability (3 years).

4.2 Model Overview

Our moving average model is a traceable Excel-based model used to calculate escalation forecasts and measure accuracy. Our model relies on a set of BLS indices to enable a measurement of accuracy between Global Insights and moving average forecasts. Our model mirrors the Global Insights forecast window parameters, calculating quarterly forecasts for two years in the future and annual forecasts for ten years in the future. Throughout this section, several abbreviations are used for conciseness (MA = moving average, Mo. over Mo. = Month over Month, Yr. over Yr. = Year over Year).

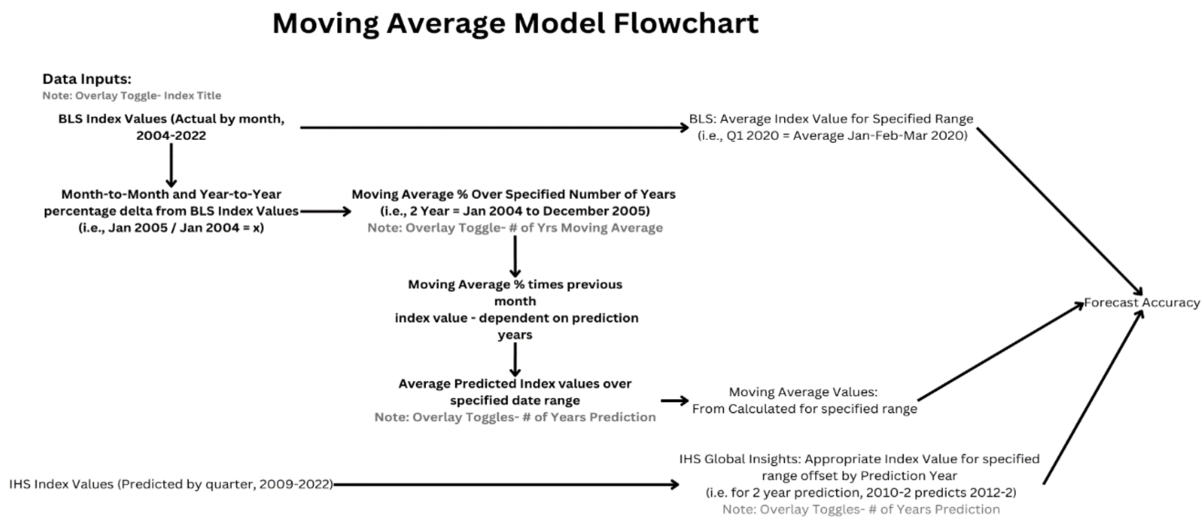
The model has three primary selection criteria which align directly with our three desired inputs:

1. **Index** – Select from the 10 sample indices curated (See Table 1 above).
2. **Moving Average Years** – Select the number of years of historical data to inform model predictions.
    - **Annual:** One, Two-, or Three-year moving average
    - **Quarterly:** One, Two-, or Three-year moving average
3. **Prediction Years** – Select the number of years the model should forecast.
    - **Annual:** One-, Two-, Five-, or Ten-year prediction
    - **Quarterly:** One- or Two-year prediction

Based on these parameters, the quarterly model will have 60 trials (10 indices, 2 prediction year options, 3 moving average year options) while the annual model will have 120 trials (10 indices, 4 prediction year options, 3 moving average year options).

Figure 5 demonstrates the visual flow of the model, starting with the data inputs, processing the moving average calculations, and providing key comparison metrics.

Figure 5 – Moving Average Model Flowchart



Further details about the model's mechanics are available in Appendix B, while Section 5.1 provides a visual walkthrough of model inputs and results for a sample combination of parameters.

4.3 Comparison Metric Overview

Forecast accuracy metrics are quantitative measures used to evaluate the precision and accuracy of predictions compared to the actual data. They are essential to evaluating the effectiveness of forecasting models, enabling organizations to make informed decisions and improve overall planning processes. For simplicity, we selected six statistical measures of the percent error between forecasted values and the actual BLS values:

1. Mean
2. Median
3. Standard Deviation
4. Range
5. Root-Mean-Squared-Error (RMSE)
6. Mean Absolute Percentage Error (MAPE)

Table 3 outlines each statistic, its calculation, and provides a brief explanation for their purposes.

Table 3 – Forecast Error Statistics Overview

Statistic	Calculation	Explanation
Mean	$\bar{x} = \frac{\Sigma xi}{n}$	Measures arithmetic mean of data (outliers contribute).
Median	$x = \left\{\frac{(n+1)}{2}\right\}^{th}$	Measures the 50th percentile of date (outliers do not contribute).
Standard Deviation	$\sqrt{\frac{\sum(x-\bar{x})^2}{n-1}}$	Measures spread of data.
Range	Maximum - Minimum	Measures the accuracy window for all outcomes in sample.
RMSE	$\sqrt{\frac{1}{n}\sum_{t=1}^{n}(y_i-\hat{y}_i)^2}$	Measures dispersion of forecast errors, sensitive to large errors (i.e., outliers); optimal values should be close to zero, large RMSE indicates high model error.
MAPE	$\frac{1}{n}\sum_{i=1}^{n}\frac{\lvert y_i-\hat{y}_i\rvert}{y_i}$	Measures dispersion of forecast errors, less sensitive to outliers, sensitive to near-zero input values; optimal values should be close to zero, large MAPE indicates high model error.

MAPE is sensitive to smaller starting values, an essential consideration for the manufacturing and semiconductor indices which have smaller starting index values than most other indices in our dataset. The summary analysis looks at which model performed best for each statistic over the total permutations of indices, moving average years, and prediction years. MAPE, and to a lesser extent RMSE, will be the most critical measures in the dataset. Average, median, standard deviation, and range will help from a more qualitative standpoint to show variation across the different models.

5. Data Analysis

5.1 Single Example – Defense Instruments (PCU334511334511)

To help demonstrate the calculations performed from index values to error statistics, the following tables and figure walk through the process for the Defense Instruments index (Search, Detection & Navigation Instruments). This trial is for a 2-year moving average and a 2-year prediction.

Table 4 – Moving Average Percentage Values (Defense Instruments, 2 Year MA, 2 Year Forecast)

Start	End	MoM	YoY
Jan-05	Jan-07	0.17%	2.01%
Jan-06	Jan-08	0.17%	2.12%
Jan-07	Jan-09	0.17%	2.06%
Jan-08	Jan-10	0.17%	2.16%
Jan-09	Jan-11	0.13%	1.70%
Jan-10	Jan-12	0.13%	1.36%
Jan-11	Jan-13	0.17%	2.01%
Jan-12	Jan-14	0.19%	2.46%
Jan-13	Jan-15	0.15%	1.82%
Jan-14	Jan-16	0.15%	1.58%
Jan-15	Jan-17	0.16%	1.78%
Jan-16	Jan-18	0.14%	1.85%
Jan-17	Jan-19	0.25%	3.00%
Jan-18	Jan-20	0.44%	5.30%
Jan-19	Jan-21	0.45%	5.73%
Jan-20	Jan-22	0.37%	3.39%
Jan-21	Jan-23	0.32%	3.55%

Table 4 represents the 2-year moving average percentages for this index. The calculation range for the first row is from January of the first year (2005) to one month before the end of the year listed (i.e., December 2006). This index has relatively small month-to-month increases and relatively standard (~2%) annual increases, with a spike in the 2020-2022 timeframe considering the COVID surge in general pricing.

Table 5 – BLS and Model Index Values (Defense Instruments, 2 Year MA, 2 Year Forecast)

Year	Quarter	BLS	Global Insights	MoM	YoY
2011	Q1	131.8	127.5	132.6	133.7
2012	Q1	135.1	134.0	135.9	137.1
2013	Q1	138.2	135.8	139.2	141.8
2014	Q1	140.1	139.1	141.6	143.2
2015	Q1	142.9	142.3	143.4	144.6
2016	Q1	144.9	142.8	146.9	148.0
2017	Q1	148.4	144.1	149.0	150.3
2018	Q1	152.9	147.7	153.9	157.4
2019	Q1	163.6	153.5	164.8	169.6
2020	Q1	172.4	155.8	175.6	182.8
2021	Q1	172.4	166.3	181.3	184.3

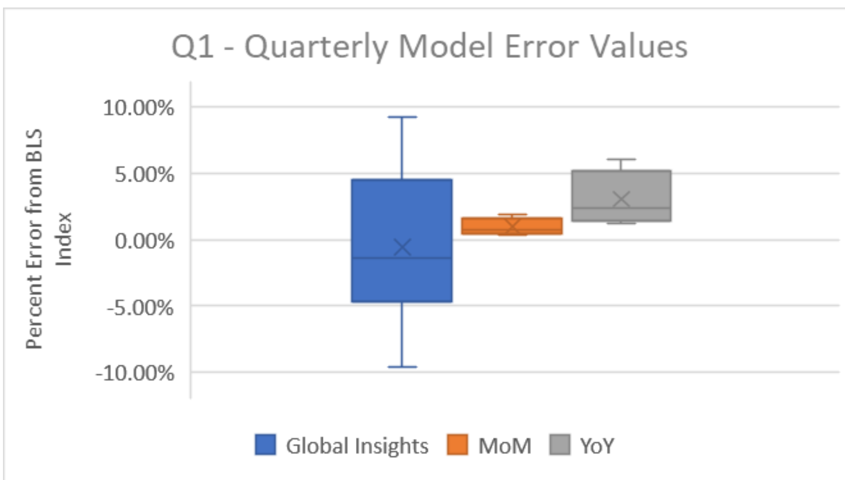
In Table 5, the BLS column reflects the historical actual values for each year in the first quarter, while each model column represents that model's forecasted values for each year. A cursory glance indicates all three models performed relatively well, but

Global Insights consistently underestimated escalation and our models consistently overestimated. Table 6 presents the percent deltas from the actual BLS values in each period and Figure 6 visually depicts these deltas.

Table 6 – Model Value Percent Error from BLS Index (Defense Instruments, 2 Year MA, 2 Year Forecast)

Year	Quarter	Global Insights	MoM	YoY
2011	Q1	-3.24%	0.60%	1.47%
2012	Q1	-0.81%	0.59%	1.49%
2013	Q1	-1.71%	0.73%	2.64%
2014	Q1	-0.74%	1.09%	2.25%
2015	Q1	-0.42%	0.37%	1.19%
2016	Q1	-1.45%	1.36%	2.15%
2017	Q1	-2.89%	0.40%	1.30%
2018	Q1	-3.40%	0.60%	2.91%
2019	Q1	-6.16%	0.72%	3.66%
2020	Q1	-9.64%	1.88%	6.05%
2021	Q1	-3.55%	5.19%	6.90%

Figure 6 – Q1 Model Error Boxplot (Defense Instruments, 2 Yr. MA, 2 Yr. Forecast)



Finally, Table 7 shows the calculated error statistics, which indicate that the MoM model seems to have performed the best. This is based on the model having the lowest MAPE and RMSE, the smallest median and average absolute values, and the tightest standard deviation. This result is pretty remarkable; while it is only one trial, a simple analysis of BLS data outperformed a commercial model.

Table 7 – Error Statistics (Defense Instruments, 2 Year MA, 2 Year Forecast)

Measure	Global Insights	MoM	YoY
Min	-9.64%	0.37%	1.19%
Max	-0.42%	1.88%	6.05%
Range	9.22%	1.52%	4.86%
Std Dev	2.88%	0.48%	1.48%
Average	-3.05%	0.83%	2.51%
Median	-2.30%	0.66%	2.20%
RMSE	6.74	1.48	4.60
MAPE	3.27	1.23	2.91

Table 8 shows all Q1 trial MAPE values for this index – across the six trial permutations, the MA YoY had the lowest MAPE in four trials, with the MA MoM performing best in the other two. Notably, while the Global Insights two-year prediction had two or three times the MAPE of the best-performing model, all three models had relatively low values for MAPE, indicating reasonable forecast accuracy for this index.

Table 8 – MAPE for all Q1 Trials for Defense Instruments

MA Yrs - Predict Yrs	Model	MAPE	MA Yrs - Predict Yrs	Model	MAPE	MA Yrs - Predict Yrs	Model	MAPE
1 MA 1 Pred	Global Insights	1.61	2 MA 1 Pred	Global Insights	1.61	3 MA 1 Pred	Global Insights	1.61
1 MA 1 Pred	MoM	1.34	2 MA 1 Pred	MoM	1.15	3 MA 1 Pred	MoM	1.10
1 MA 1 Pred	YoY	1.12	2 MA 1 Pred	YoY	0.57	3 MA 1 Pred	YoY	0.51
1 MA 2 Pred	Global Insights	3.08	2 MA 2 Pred	Global Insights	3.08	3 MA 2 Pred	Global Insights	3.08
1 MA 2 Pred	MoM	1.57	2 MA 2 Pred	MoM	1.59	3 MA 2 Pred	MoM	1.80
1 MA 2 Pred	YoY	2.82	2 MA 2 Pred	YoY	3.21	3 MA 2 Pred	YoY	1.04

The results in Table 9 for 5- and 10-year predictions are more favorable for Global Insights than those for 1- and 2-year predictions, where the moving average forecasts performed better. For this index, the difference between MA and Global Insights at 5 years is relatively slight; but at 10 years, Global Insights has only 60% the MAPE of the MoM and YoY models.

Table 9 – MAPE for all Annual Trials for Defense Instruments

MA Yrs - Predict Yrs	Model	MAPE	MA Yrs - Predict Yrs	Model	MAPE	MA Yrs - Predict Yrs	Model	MAPE
1 MA 1 Pred	Global Insights	1.44	2 MA 1 Pred	Global Insights	1.43	3 MA 1 Pred	Global Insights	1.53
1 MA 1 Pred	MoM	0.97	2 MA 1 Pred	MoM	0.95	3 MA 1 Pred	MoM	0.95
1 MA 1 Pred	YoY	1.28	2 MA 1 Pred	YoY	0.58	3 MA 1 Pred	YoY	0.49
1 MA 2 Pred	Global Insights	3.05	2 MA 2 Pred	Global Insights	2.98	3 MA 2 Pred	Global Insights	2.98
1 MA 2 Pred	MoM	1.98	2 MA 2 Pred	MoM	1.80	3 MA 2 Pred	MoM	1.75
1 MA 2 Pred	YoY	2.44	2 MA 2 Pred	YoY	2.59	3 MA 2 Pred	YoY	0.49
1 MA 5 Pred	Global Insights	7.98	2 MA 5 Pred	Global Insights	7.09	3 MA 5 Pred	Global Insights	6.40
1 MA 5 Pred	MoM	10.88	2 MA 5 Pred	MoM	10.83	3 MA 5 Pred	MoM	10.87
1 MA 5 Pred	YoY	10.52	2 MA 5 Pred	YoY	11.17	3 MA 5 Pred	YoY	8.64
1 MA 10 Pred	Global Insights	18.87	2 MA 10 Pred	Global Insights	18.16	3 MA 10 Pred	Global Insights	18.38
1 MA 10 Pred	MoM	30.07	2 MA 10 Pred	MoM	32.15	3 MA 10 Pred	MoM	33.26
1 MA 10 Pred	YoY	27.71	2 MA 10 Pred	YoY	31.47	3 MA 10 Pred	YoY	30.00

5.2 Data Analysis Overview

Appendix B contains the specific results across all 60 Quarterly and 120 Annual Trials (see Section 4.2 for trial permutation breakdown). With 180 trials to consider, we focused on the following questions to achieve the overall goal of assessing the forecasts:

1. How statistically accurate (or inaccurate) are escalation forecasts?
2. Where did the commercial model outperform the moving average model and vice versa? Was the difference visible based on index volatility, prediction years, or commodity type?
3. Did accuracy increase as moving average years increased and did accuracy decrease as prediction years increased, as would be expected?

5.3 Quarterly Data Analysis

Table 10 demonstrates the count of which model performed best in each significant metric. The macro-level trend across the population of trials is that the MA models typically (~80% of the time) outperformed the Global Insights model. Importantly, performance was similar across all four measures, indicating minimal outliers. While the prediction ranges were relatively short (one or two years), the error bounds were tight. One-year predictions across all models had a 2.2% average absolute error (indicating that, on average, forecasted indices were 2.2% higher or lower than the actual BLS values), with two-year predictions having a 4.1% average absolute error. Furthermore, the Gas index had errors across all models in the 20% range, pulling up that cumulative average error (notably, MA models had half the error of Global Insights forecasts for this index).

Table 10 – Best Performing Model for Quarterly Trials

Model	Average	Median	RMSE	MAPE
Global Insights	10	11	11	12
MA Mo.	23	26	29	29
MA Yr.	27	23	20	19
Totals:	**60**	**60**	**60**	**60**

Table 11 represents a “heatmap” view of the quarterly trials, in which each cell represents a trial and the colors represent the three models. For example, the heatmap indicates that the MoM model perform best in all six MA / Prediction Year permutations for index WPU061 (Chemicals); the YoY model performed best for seven out of ten indices for the 3 MA 2 Prediction Year trials. The three-year moving averages outperform Global Insights more than the one- and two-year versions (helping to answer the third question from the start of this section). Specific indices were more robust for Global Insights than others (notably the Manufacturing, Fuel, and Aircraft Parts indices). However, overall, the moving average model did exceptionally well compared to the expectations outlined in Section 3.1. On average, the difference between the Global Insights and MA error statistics was minimal, indicating that while there were clear cases where one model was significantly better than the others, they all performed similarly across the set of trials.

Table 11 – Lowest-MAPE Model for Quarterly Trials Heatmap

	MAPE					
Index	1 MA 1 Pred	1 MA 2 Pred	2 MA 1 Pred	2 MA 2 Pred	3 MA 1 Pred	3 MA 2 Pred
CEU3000000008	YoY	IHS	YoY	IHS	YoY	YoY
PCU3222113222110	MoM	MoM	MoM	MoM	MoM	YoY
PCU33441-33441-	YoY	YoY	YoY	MoM	YoY	YoY
PCU334511334511	YoY	MoM	YoY	MoM	YoY	YoY
PCU336413336413	IHS	IHS	YoY	IHS	YoY	YoY
WPU053	IHS	IHS	IHS	MoM	IHS	IHS
WPU061	MoM	MoM	MoM	MoM	MoM	MoM
WPU08	MoM	IHS	MoM	MoM	MoM	YoY
WPU101	MoM	IHS	MoM	MoM	MoM	YoY
WPU103103	MoM	MoM	MoM	MoM	MoM	MoM

5.4 Annual Data Analysis

Table 12 indicates that Global Insights outperforms in the ten-year range by a factor of more than two. Note the massive MAPE for 1 MA 10 Pred trials caused by the highly volatile fuels index; removing that index from the dataset reduces the MoM MAPE by a factor of 10. Concerning the second question posed earlier in this section, the Global Insights MAPE at 10 years was similar with and without the fuels index, indicating that volatility of inputs did not significantly impact commercial model performance.

Table 12 – MAPE for all Annual Trials

MA Yrs - Predict Yrs	Model	MAPE	MA Yrs - Predict Yrs	Model	MAPE	MA Yrs - Predict Yrs	Model	MAPE
1 MA 1 Pred	Global Insights	7.65	2 MA 1 Pred	Global Insights	6.33	3 MA 1 Pred	Global Insights	5.37
1 MA 1 Pred	MoM	5.23	2 MA 1 Pred	MoM	3.25	3 MA 1 Pred	MoM	3.75
1 MA 1 Pred	YoY	11.03	2 MA 1 Pred	YoY	4.84	3 MA 1 Pred	YoY	4.77
1 MA 2 Pred	Global Insights	10.49	2 MA 2 Pred	Global Insights	8.37	3 MA 2 Pred	Global Insights	7.40
1 MA 2 Pred	MoM	10.39	2 MA 2 Pred	MoM	4.57	3 MA 2 Pred	MoM	5.64
1 MA 2 Pred	YoY	18.77	2 MA 2 Pred	YoY	6.00	3 MA 2 Pred	YoY	4.19
1 MA 5 Pred	Global Insights	14.79	2 MA 5 Pred	Global Insights	12.68	3 MA 5 Pred	Global Insights	11.09
1 MA 5 Pred	MoM	53.98	2 MA 5 Pred	MoM	18.52	3 MA 5 Pred	MoM	17.59
1 MA 5 Pred	YoY	53.54	2 MA 5 Pred	YoY	21.72	3 MA 5 Pred	YoY	11.30
1 MA 10 Pred	Global Insights	16.38	2 MA 10 Pred	Global Insights	15.79	3 MA 10 Pred	Global Insights	16.25
1 MA 10 Pred	MoM	423.02	2 MA 10 Pred	MoM	36.64	3 MA 10 Pred	MoM	32.54
1 MA 10 Pred	YoY	116.39	2 MA 10 Pred	YoY	61.90	3 MA 10 Pred	YoY	32.74

That said, the performance for a 5-year prediction with two or three years of MA is often relatively strong (in the case of the three-year moving average, YoY was nearly identical in MAPE to Global Insights). The MA forecast fares far better in the 1- and 2-year prediction range, often outperforming Global Insights. As noted with the quarterly results, the deltas between Global Insights and MA in the shorter prediction ranges are relatively low in magnitude.

Table 13 – Best Performing Model for Annual Trials

Model	Average	Median	RMSE	MAPE
Global Insights	62	55	56	62
MA Mo.	19	23	35	34
MA Yr.	39	42	29	24
Total:	**120**	**120**	**120**	**120**

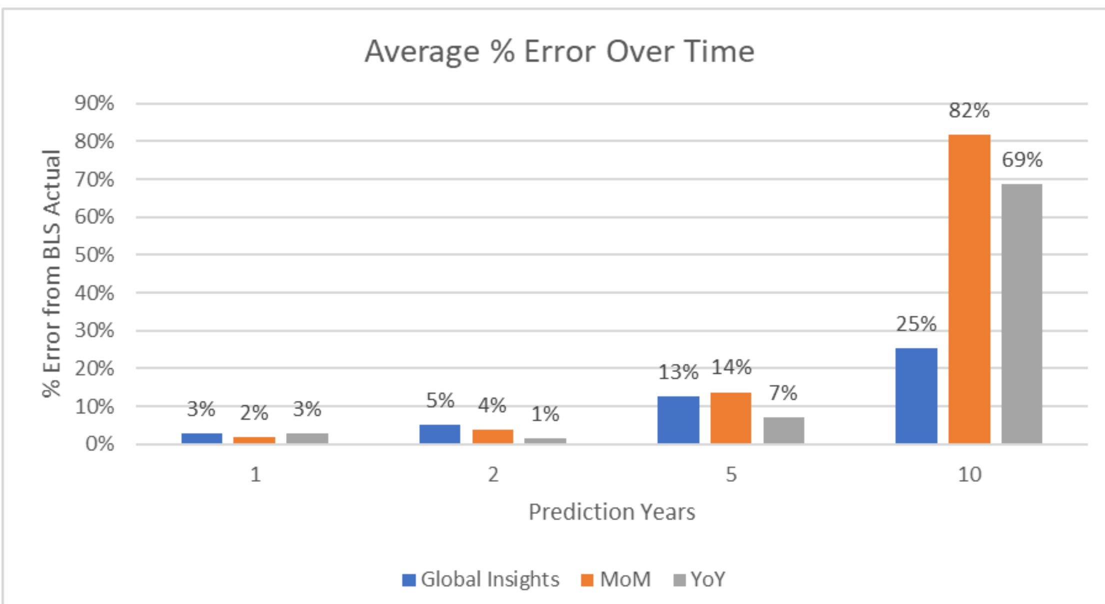
Table 13 demonstrates that Global Insights had the lowest average error (and lowest MAPE) in 51.6% of trials, but Table 14 sheds further light on the performance disparity between near-term and long-term estimates. Global Insights always had the lower MAPE at the 10-year prediction mark (and in 80% of 5-year prediction trials), while the MA models outperformed in 87% of 1- and 2-year trials. Notably, the MA models performed better in 100% of trials when there were 3 years of MA data (and improved consistently from one to three years). The commodity type (and volatility) had surprisingly minimal impacts on which trial performed best, solidifying that prediction year and the amount of historical data are the primary reasons for differences in model performance.

Table 14 – MAPE Heatmap for Annual Trials

	MAPE											
Index	**1 MA 1 Pred**	**1 MA 2 Pred**	**1 MA 5 Pred**	**1 MA 10 Pred**	**2 MA 1 Pred**	**2 MA 2 Pred**	**2 MA 5 Pred**	**2 MA 10 Pred**	**3 MA 1 Pred**	**3 MA 2 Pred**	**3 MA 1 Pred**	**3 MA 2 Pred**
CEU3000000008	MoM	IHS	IHS	IHS	YoY	IHS	IHS	IHS	YoY	YoY	IHS	IHS
PCU3222113222110	MoM	IHS	IHS	IHS	MoM	YoY	IHS	IHS	YoY	YoY	IHS	IHS
PCU33441-33441-	MoM	YoY	YoY	IHS	YoY	YoY	MoM	IHS	YoY	YoY	YoY	IHS
PCU334511334511	MoM	MoM	IHS	IHS	YoY	MoM	IHS	IHS	YoY	YoY	IHS	IHS
PCU336413336413	IHS	IHS	IHS	IHS	MoM	IHS	IHS	IHS	YoY	YoY	IHS	IHS
WPU053	MoM	IHS	IHS	IHS	MoM	MoM	IHS	IHS	MoM	YoY	YoY	IHS
WPU061	MoM	MoM	IHS	IHS	MoM	MoM	IHS	IHS	MoM	MoM	IHS	IHS
WPU08	MoM	MoM	IHS	IHS	MoM	MoM	IHS	IHS	MoM	YoY	IHS	IHS
WPU101	MoM	IHS	IHS	IHS	MoM	MoM	IHS	IHS	MoM	YoY	IHS	IHS
WPU103103	MoM	MoM	YoY	IHS	MoM	MoM	IHS	IHS	MoM	MoM	YoY	IHS

In reference to the discussion in Section 2.2 about what factors are (or aren't) baked into the historical BLS indices, how does that help explain the stark performance difference between MA and Global Insights over time? Some of it could be the sheer simplicity of the MA model used in this high-level comparison, which did not use any form of smoothing over time. However, the superior MA model performance in one- and two-year predictions might indicate a commercial model overreliance on the type of scheduled future events or other qualitative factors described in Section 2.2. Figure 7 demonstrates the rise in average percentage error of the forecast values over time, with five-year forecasts averaging between 7%-14% error and even short-term forecasts averaging between 2%-5% error. Additionally, note the magnitude of the delta in average percent error between 5- and 10-year forecasts: for Global Insights the increase is two-fold (12.5% to 25.2%) and for MoM the increase is six-fold (14% to 82%). For YoY, the increase is nearly twelve-fold (7% to 69%).

Figure 7 – Average % Error Over Time



6. Results

6.1 Summary of Major Trends

Our research objective was to explore various measures of forecast accuracy to improve the treatment of escalation uncertainty in estimates, thus enhancing their credibility. We collected Global Insights and BLS data and leveraged the latter to develop a traceable, simple moving average escalation forecast model. We then used a diverse set of forecast accuracy measures to compare the performance of forecasts from Global Insights against those of our moving average models and glean insights on how those results informed escalation uncertainty. We framed three key questions for analyzing the results in Section 5.2, which we can now use the forecast measure data to address:

1. How statistically accurate (or inaccurate) are escalation forecasts?

Figure 7 demonstrates a 2-5% average error in one- and two-year forecasts, increasing to 7-14% in five-year forecasts and upwards of 25% at ten years. This represents a significant level of error given the prevalence of escalation forecasts in cost models, as demonstrated in Section 6.2.

2. Where did the commercial model outperform the moving average model and vice versa? Was the difference visible based on index volatility, prediction years, or commodity type?

The commercial model outperformed the moving average model in 90% of five and ten-year Annual trials, while the moving average model outperformed in 80% of Quarterly trials and 87% of one- and two-year Annual trials.

The differences in performance were primarily due to the number of prediction years and years of moving average data in each trial, with index volatility/commodity only becoming a factor for one of the ten indices tested (Fuels).

3. Did accuracy increase as moving average years increased and did accuracy decrease as prediction years increased, as expected?

Accuracy did increase as moving average years increased: the 3-year MA models outperformed the commercial model in 80% of Quarterly trials compared to only 65% of trials for the 1-year MA models. Additionally, while the commercial models consistently outperformed in five-year trials, the difference in MAPE at 3 MA years was almost imperceptible (with only 0.8% difference in measured variance for the YoY model (11.30) compared to the Global Insights Forecast (11.09)).

Accuracy also decreased as prediction years increased, with the average forecast error consistently doubling between one- and two-year Quarterly trials.

The primary takeaway is not that simple moving average models are the future, but that escalation should not be treated as a check-box exercise, but one that warrants further research, modeling, and estimate consideration. Estimators should understand that there is data out there to help inform these considerations, and not become overly reliant on externally derived projections with no published measures of accuracy. These initial results are not broad enough to define policy, but shed light on the level of escalation forecast uncertainty in standard indices and inform recommendations for estimators (see Section 7.1).

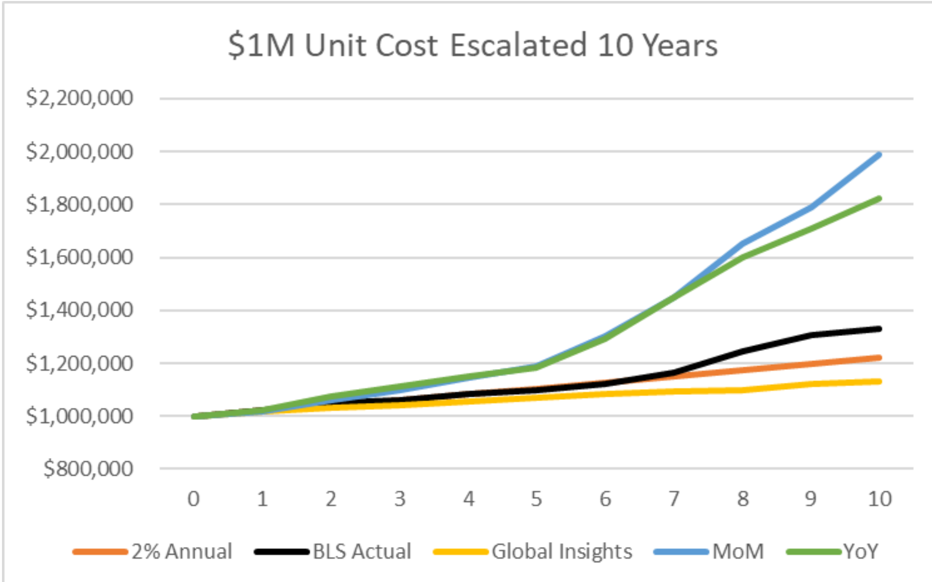
6.2 Cost Example

Revisiting the hypothetical contracting officer mentioned in Section 1.2, how would the navigation component procurement turn out compared to actual experienced escalation and our modeling approaches? For the Defense Instruments Index, the actual inflated cost from 2011 to 2021 for a $1M piece of equipment was $1,331,567. Figure 8 demonstrates that at 10 years, a 2% annual escalation forecast would underestimate the cost by $112,573, but Global Insights would have underestimated by $198,562 (and underestimated cost every year). That said, both moving average models (set at 2 years of MA data) would have overestimated the cost in the MA range of 30% (see Table 15 for details), indicative of the loss of explanatory power of shorter moving averages in the longer term.

Table 15 – Unit Cost Escalation Example for $1M Piece of Equipment

Years	2% Annual	BLS Actual	Global Insights	MoM	YoY
Base	$1,000,000	$1,000,000	$1,000,000	$1,000,000	$1,000,000
1	$1,020,000	$1,024,224	$1,016,503	$1,015,897	$1,020,439
2	$1,040,400	$1,049,962	$1,030,129	$1,062,074	$1,074,943
5	$1,104,081	$1,099,924	$1,071,083	$1,186,223	$1,183,195
10	$1,218,994	$1,331,567	$1,133,005	$1,990,916	$1,820,590

Figure 8 – Unit Cost Escalation Example for $1M Piece of Equipment



While the moving average indices show higher magnitude errors, we think the hypothetical contracting officer would need to be more comfortable knowing their commercial inflation model was the most accurate but underestimated inflation, possibly leaving them under-resourced. At 5 years, the same trends hold albeit at less significant margins. While this scenario is a single non-volatile index for a single set of model permutations across the breadth of defense contracting, these forecast variances can have massive consequences for key procurements.

7. Conclusion

7.1 Recommendations for Use

If our repeated finding is a need for increased risk awareness on escalation parameters, how should estimators look to incorporate risk into their models? While this paper's scope only covers a handful of escalation indices, our general recommendations are:

First, calculate the volatility (% change year to year) for each escalation index in a given model – this helps develop narrative parameters such as "this index often changes more than +/- 3% year to year" or "this index consistently changes ~2% a year, except for a particular historical instance".

Second, if possible, use the error metrics and techniques in this paper to calculate the historical accuracy of forecast models at an index or composite level.

Third, leverage the volatility and historical accuracy to derive a risk range. For example, if the historical forecasts for an index are low by 5% on average, this informs a right-skewed distribution; if the index tends to vary up or down by 4% over a given range, this informs the range of a triangular distribution.

Finally, incorporate the risk range around the escalation CAGR being applied (for example, a 5-year CAGR of 3% for a given escalation index used in a 5-year production contract). This risk range must be formatted as a Markov chain, as applying an inaccurate escalation value in the first year influences the application in the second year and so on. If the escalation values are applied individually by year (rather than a CAGR), then the risk range can be a more standard triangular format.

7.2 Next Steps

This paper only scratches the surface of this topic. As a result, the aforementioned recommendations are limited until additional research is performed. We intend to take the following next steps:

We would expand the sample dataset from ten indices to the several hundred available between Global Insights and BLS and from just 2008-2022 to a wider date range with more historical values. Expansion of the Global Insights forecast dataset would provide not only a larger, richer set of indices, but also more data points for the critical five- and ten-year trials. We would also leverage other commercial models or government forecasts (for fuels, particularly) to compare trends across the industry.

This paper used the simplest version of a moving average model. This raises the question, “How would adding EWMA, an ARIMA component, more moving average years, or optimizing the moving average range to reduce error impact the forecast quality?” More pragmatically, a future paper could go from merely justifying the need for escalation risk ranges to creating them by establishing clear guidelines for assessing forecast accuracy and assigning risk factors to different indices and programs.

Bibliography

Clark, T. (2021). Comparing Measures of Core Inflation. *Economic Review*.

Fernando, J. (2024, February). *Compound Annual Growth Rate (CAGR) Formula and Calculation*. Retrieved from Investopedia: https://www.investopedia.com/terms/c/cagr.asp

Frick, W. (2022, December). *"What Causes Inflation"*. Retrieved from Harvard Business Review: https://hbr.org/2022/12/what-causes-inflation

Nau, R. (2014, August). *Forecasting with Moving Averages*. Retrieved from Duke University, Fuqua School of Business: https://people.duke.edu/~rnau/Notes_on_forecasting_with_moving_averages--Robert_Nau.pdf

OSD. (2021, December). *Inflation and Escalation Best Practices for Cost Analysis*. Retrieved from https://www.cape.osd.mil/files/Reports/OSDCAPEEscalationHandbook2021.pdf

OSD. (2022, May). *Guidance on Inflation and Economic Price Adjustments*. Retrieved from https://www.acq.osd.mil/dpap/policy/policyvault/USA000999-22-DPC.pdf

Appendix A – Moving Average Model Details

Once the user has selected the criteria specified in Section 4.2, the model will dynamically calculate the MoM and YoY percent change values. The range end dates are fixed to January 2006 to January 2023, while the start date is adjusted based on the moving average years toggle (N-1, N-2, N-3). BLS values are pulled in based on the toggled index and given date. The MoM data calculates a percentage change of the BLS data from one month to the next (Month N minus Month (N-1)). Similarly, the YoY data calculates a percentage change of BLS data from one year to the next (Month N minus Month (N-12)). These percent change values are averaged over the toggled span (1, 2, or 3 years) to derive the MA value for each year.

For all years in the calculation range (2006-2022), the MoM section takes the BLS actual index value one month before the start date (i.e., December) and multiplies it by the MA value to derive the January value. In subsequent months, the MoM section multiplies the value one column to the left by the MA to achieve the one- and two-year forecasts. YoY operates similarly, calculating one year prior versus one month prior.

Next, the actual BLS values for each quarter and year (via averaging the monthly values for each quarter) are compared to the Global Insights prediction values for each quarter/year and the derived MoM and YoY MA values for each quarter/year. We then calculate the percentage difference between the actual BLS values and the Global Insights, MoM, and YoY for each assessed year. This facilitates calculation of the various consolidated error metrics: Minimum, Maximum, Range, Standard Deviation, Average, Median, Root-Mean-Square Deviation (RMSE), and Mean-Absolute-Percentage-Error (MAPE).

We used the month-to-month and year-to-year percentage change for our moving average. However, quarterly data was also analyzed using quarterly results (i.e., quarter one being the average of our model's values for January through March) both to reduce seasonal variance (better as a quarterly value than a monthly value) and align with Global Insights predictions, which are calculated for future quarters vice months. Theoretically, the moving average could have been taken for the quarterly

CAGR (i.e., January to March, April to June) with slightly different results. However, this might mask some month-to-month (or year-to-year) volatility that would prove beneficial to a moving average forecast. Hence, we decided to use this method instead of a quarterly CAGR.

Lastly, we made one simplification to avoid using a different moving average for each quarter's prediction. To further explain, a one-year moving average range for a one-year prediction beginning in January 2008 takes the average of the percent month-to-month differences between January 2007 and December 2007 and multiplies it by the BLS value for December 2007 to yield a prediction for January 2008. That prediction is then multiplied by the moving average to yield February 2008, and so on. This is based on the data available as of 1 January 2008. However, the second quarter 2008 prediction (1 April 2008) would have some data for January 2008 to March 2008 available if we shifted the moving average range to calculate from 1 April 2007 to 31 March 2008. While this would potentially add predictive value (particularly by the fourth quarter), we felt that (1) it introduced significant complexity into the calculations and (2) given the lack of insight into what data Global Insights had at the time of predictions, it was better to take the conservative approach and use only the prior year(s) data.

Although optimization of the "m" term is undoubtedly possible given sufficient data, we chose not to explore the concept in our initial model to allow focus on comparisons to a simple mathematical model. Additionally, techniques such as the Exponentially Weighted Moving Average (EWMA), which reduces the weight of each data point the further from "t" it is, can add predictive value. We decided not to incorporate advanced moving average models and smoothing techniques in this paper, but may consider doing so in follow-on work.

Appendix B – Data Tables

The following tables provide the results of our analysis for all trials across all ten indices, beginning with quarterly trials and moving to annual trials. For quarterly trials, Table 16 contains the error statistics for each Model / MA Yrs. / Predict Yrs., averaged across all 10 indices; Table 17, Table 18, and Table 19 represent heatmaps for Average, Median, and RMSE respectively (see Table 11 for MAPE heatmap). For annual trials, Table 20, Table 21, and Table 22 contain the error statistics for one-, two-, and three-year trials, respectively, while Table 23, Table 24, and Table 25 represent heatmaps for Average, Median, and RMSE respectively (see Table 14 for MAPE heatmap).

Table 16 – Quarterly Trial Error Statistics

MA Yrs - Predict Yrs	Model	Average	Median	Std. Dev	RMSE	MAPE
1 MA 1 Pred	Global Insights	0.024	0.019	0.077	14.37	5.64
1 MA 1 Pred	MoM	0.019	0.004	0.078	11.97	5.56
1 MA 1 Pred	YoY	0.023	0.018	0.095	17.05	7.77
1 MA 2 Pred	Global Insights	0.054	0.039	0.108	20.64	7.88
1 MA 2 Pred	MoM	0.043	0.017	0.141	21.99	10.19
1 MA 2 Pred	YoY	0.052	0.050	0.160	31.74	13.79
2 MA 1 Pred	Global Insights	0.024	0.019	0.077	14.37	5.64
2 MA 1 Pred	MoM	0.014	0.007	0.046	7.64	3.72
2 MA 1 Pred	YoY	0.013	0.013	0.056	10.04	4.54
2 MA 2 Pred	Global Insights	0.054	0.039	0.108	20.64	7.88
2 MA 2 Pred	MoM	0.023	0.012	0.057	10.56	5.04
2 MA 2 Pred	YoY	0.024	0.010	0.082	17.29	7.17
3 MA 1 Pred	Global Insights	0.027	0.021	0.085	15.78	6.19
3 MA 1 Pred	MoM	0.024	0.004	0.066	10.72	5.32
3 MA 1 Pred	YoY	0.026	0.011	0.078	13.12	6.08
3 MA 2 Pred	Global Insights	0.054	0.039	0.108	20.64	7.88
3 MA 2 Pred	MoM	0.042	0.016	0.083	14.91	7.26
3 MA 2 Pred	YoY	0.027	0.013	0.076	12.92	5.83

Table 17 – Quarterly Average Heatmap

Average						
Index	1 MA 1 Pred	1 MA 2 Pred	2 MA 1 Pred	2 MA 2 Pred	3 MA 1 Pred	3 MA 2 Pred
CEU3000000008	YoY	IHS	YoY	IHS	YoY	IHS
PCU3222113222110	YoY	IHS	YoY	IHS	YoY	YoY
PCU33441-33441-	YoY	MoM	YoY	MoM	MoM	YoY
PCU334511334511	YoY	MoM	YoY	MoM	YoY	YoY
PCU336413336413	IHS	IHS	YoY	IHS	YoY	YoY
WPU053	YoY	YoY	YoY	YoY	YoY	YoY
WPU061	MoM	MoM	MoM	MoM	MoM	MoM
WPU08	YoY	MoM	YoY	IHS	MoM	IHS
WPU101	MoM	MoM	MoM	MoM	MoM	MoM
WPU103103	MoM	MoM	YoY	MoM	MoM	YoY

Table 18 – Quarterly Median Heatmap

Median						
Index	1 MA 1 Pred	1 MA 2 Pred	2 MA 1 Pred	2 MA 2 Pred	3 MA 1 Pred	3 MA 2 Pred
CEU3000000008	IHS	IHS	IHS	IHS	IHS	IHS
PCU3222113222110	MoM	MoM	MoM	YoY	YoY	YoY
PCU33441-33441-	YoY	YoY	MoM	MoM	MoM	YoY
PCU334511334511	YoY	MoM	YoY	MoM	YoY	YoY
PCU336413336413	MoM	IHS	YoY	IHS	YoY	YoY
WPU053	MoM	MoM	MoM	YoY	MoM	MoM
WPU061	MoM	MoM	MoM	MoM	YoY	YoY
WPU08	IHS	YoY	YoY	IHS	YoY	YoY
WPU101	MoM	IHS	MoM	MoM	YoY	MoM
WPU103103	MoM	MoM	MoM	YoY	MoM	YoY

Table 19 – Quarterly RMSE Heatmap

	RMSE					
Index	1 MA 1 Pred	1 MA 2 Pred	2 MA 1 Pred	2 MA 2 Pred	3 MA 1 Pred	3 MA 2 Pred
CEU3000000008	YoY	IHS	YoY	IHS	YoY	YoY
PCU3222113222110	MoM	IHS	MoM	MoM	MoM	YoY
PCU33441-33441-	YoY	YoY	YoY	MoM	YoY	YoY
PCU334511334511	YoY	MoM	YoY	MoM	YoY	YoY
PCU336413336413	IHS	IHS	YoY	IHS	YoY	YoY
WPU053	IHS	IHS	MoM	MoM	MoM	YoY
WPU061	MoM	MoM	MoM	MoM	MoM	MoM
WPU08	MoM	IHS	MoM	MoM	MoM	YoY
WPU101	IHS	IHS	MoM	MoM	MoM	YoY
WPU103103	MoM	MoM	MoM	MoM	MoM	MoM

Note – for Annual forecasts below, the Global Insights forecasts is different for the one, two, and three MA year trials due to the difference in available model years (two-year trials do not include 2022, three-year trials do not include 2021-2022).

Table 20 – Annual Trial Statistics for 1 MA Trials

MA Yrs - Predict Yrs	Model	Average	Median	RMSE	MAPE
1 MA 1 Pred	Global Insights	0.003	0.013	23.18	7.65
1 MA 1 Pred	MoM	0.013	0.003	18.11	5.23
1 MA 1 Pred	YoY	0.021	0.015	29.82	11.03
1 MA 2 Pred	Global Insights	0.015	0.022	29.58	10.49
1 MA 2 Pred	MoM	0.072	0.020	55.88	10.39
1 MA 2 Pred	YoY	0.046	0.031	58.19	18.77
1 MA 5 Pred	Global Insights	0.075	0.097	39.32	14.79
1 MA 5 Pred	MoM	0.956	0.080	708.43	53.98
1 MA 5 Pred	YoY	0.605	0.012	529.93	53.54
1 MA 10 Pred	Global Insights	0.128	0.186	48.03	16.38
1 MA 10 Pred	MoM	91.581	1.354	57,102.63	423.02
1 MA 10 Pred	YoY	46.412	0.602	28,772.47	116.39

Table 21 – Annual Trial Statistics for 2 MA Trials

MA Yrs - Predict Yrs	Model	Average	Median	RMSE	MAPE
2 MA 1 Pred	Global Insights	0.016	0.016	19.72	6.33
2 MA 1 Pred	MoM	0.007	0.005	8.76	3.25
2 MA 1 Pred	YoY	0.008	0.005	13.41	4.84
2 MA 2 Pred	Global Insights	0.035	0.026	23.26	8.37
2 MA 2 Pred	MoM	0.029	0.021	12.46	4.57
2 MA 2 Pred	YoY	0.013	0.010	13.67	6.00
2 MA 5 Pred	Global Insights	0.107	0.113	34.29	12.68
2 MA 5 Pred	MoM	0.224	0.041	113.53	18.52
2 MA 5 Pred	YoY	0.102	0.001	61.48	21.72
2 MA 10 Pred	Global Insights	0.189	0.212	45.70	15.79
2 MA 10 Pred	MoM	5.874	0.465	2207.09	36.64
2 MA 10 Pred	YoY	1.338	0.539	511.33	61.90

Table 22 – Annual Trial Statistics for 3 MA Trials

MA Yrs - Predict Yrs	Model	Average	Median	RMSE	MAPE
3 MA 1 Pred	Global Insights	0.028	0.022	15.28	5.37
3 MA 1 Pred	MoM	0.018	0.010	8.91	3.75
3 MA 1 Pred	YoY	0.028	0.006	14.35	4.77
3 MA 2 Pred	Global Insights	0.050	0.031	19.95	7.40
3 MA 2 Pred	MoM	0.038	0.016	13.98	5.64
3 MA 2 Pred	YoY	0.015	0.003	10.36	4.19
3 MA 5 Pred	Global Insights	0.125	0.115	26.70	11.09
3 MA 5 Pred	MoM	0.137	0.047	52.36	17.59
3 MA 5 Pred	YoY	0.070	0.025	32.94	11.30
3 MA 10 Pred	Global Insights	0.252	0.236	45.89	16.25
3 MA 10 Pred	MoM	0.818	0.384	185.30	32.54
3 MA 10 Pred	YoY	0.688	0.289	151.77	32.74

Table 23 – Annual Average Heatmap

	Average											
Index	1 MA 1 Pred	1 MA 2 Pred	1 MA 5 Pred	1 MA 10 Pred	2 MA 1 Pred	2 MA 2 Pred	2 MA 5 Pred	2 MA 10 Pred	3 MA 1 Pred	3 MA 2 Pred	3 MA 1 Pred	3 MA 2 Pred
CEU3000000008	YoY	IHS	IHS	IHS	YoY	IHS	IHS	IHS	YoY	YoY	IHS	IHS
PCU3222113222110	MoM	YoY	IHS	IHS	MoM	YoY	IHS	IHS	YoY	YoY	IHS	IHS
PCU33441-33441-	MoM	MoM	MoM	IHS	YoY	IHS	IHS	IHS	YoY	YoY	IHS	MoM
PCU334511334511	YoY	MoM	IHS	IHS	YoY	MoM	IHS	IHS	YoY	YoY	IHS	IHS
PCU336413336413	MoM	IHS	IHS	IHS	MoM	IHS	IHS	IHS	YoY	YoY	IHS	IHS
WPU053	MoM	YoY	IHS	IHS	YoY	YoY	YoY	IHS	YoY	YoY	YoY	IHS
WPU061	IHS	IHS	IHS	IHS	YoY	YoY	MoM	IHS	MoM	YoY	IHS	IHS
WPU08	YoY	MoM	IHS	IHS	YoY	MoM	IHS	IHS	MoM	YoY	IHS	IHS
WPU101	MoM	IHS	IHS	IHS	YoY	IHS	IHS	IHS	MoM	YoY	YoY	IHS
WPU103103	IHS	YoY	IHS	IHS	YoY	MoM	YoY	IHS	YoY	YoY	YoY	IHS

Table 24 – Annual Median Heatmap

	Median											
Index	1 MA 1 Pred	1 MA 2 Pred	1 MA 5 Pred	1 MA 10 Pred	2 MA 1 Pred	2 MA 2 Pred	2 MA 5 Pred	2 MA 10 Pred	3 MA 1 Pred	3 MA 2 Pred	3 MA 1 Pred	3 MA 2 Pred
CEU3000000008	YoY	IHS	IHS	IHS	YoY	IHS	IHS	IHS	YoY	YoY	IHS	IHS
PCU3222113222110	MoM	MoM	IHS	IHS	MoM	YoY	IHS	IHS	YoY	YoY	IHS	IHS
PCU33441-33441-	YoY	YoY	MoM	IHS	YoY	YoY	YoY	IHS	YoY	YoY	YoY	IHS
PCU334511334511	YoY	YoY	IHS	IHS	YoY	MoM	IHS	IHS	YoY	YoY	IHS	IHS
PCU336413336413	YoY	IHS	IHS	IHS	MoM	IHS	IHS	IHS	YoY	YoY	IHS	IHS
WPU053	MoM	MoM	YoY	YoY	YoY	YoY	MoM	MoM	YoY	YoY	YoY	YoY
WPU061	MoM	IHS	IHS	IHS	MoM	MoM	IHS	IHS	MoM	YoY	IHS	IHS
WPU08	MoM	YoY	IHS	IHS	MoM	IHS	IHS	IHS	YoY	YoY	IHS	IHS
WPU101	MoM	IHS	MoM	IHS	YoY	YoY	MoM	IHS	MoM	MoM	YoY	IHS
WPU103103	IHS	MoM	MoM	IHS	YoY	YoY	IHS	IHS	YoY	YoY	IHS	IHS

Table 25 – Annual RMSE Heatmap

	RMSE											
Index	1 MA 1 Pred	1 MA 2 Pred	1 MA 5 Pred	1 MA 10 Pred	2 MA 1 Pred	2 MA 2 Pred	2 MA 5 Pred	2 MA 10 Pred	3 MA 1 Pred	3 MA 2 Pred	3 MA 1 Pred	3 MA 2 Pred
CEU3000000008	MoM	IHS	IHS	IHS	YoY	IHS	IHS	IHS	YoY	YoY	IHS	IHS
PCU3222113222110	MoM	YoY	IHS	IHS	MoM	MoM	IHS	IHS	YoY	YoY	IHS	IHS
PCU33441-33441-	MoM	YoY	YoY	YoY	YoY	YoY	MoM	IHS	YoY	YoY	YoY	IHS
PCU334511334511	MoM	MoM	IHS	IHS	YoY	MoM	IHS	IHS	YoY	YoY	IHS	IHS
PCU336413336413	IHS	IHS	IHS	IHS	YoY	IHS	IHS	IHS	YoY	YoY	IHS	IHS
WPU053	MoM	IHS	IHS	IHS	MoM	YoY	IHS	IHS	MoM	YoY	YoY	IHS
WPU061	MoM	YoY	MoM	MoM	MoM	MoM	MoM	MoM	YoY	MoM	MoM	MoM
WPU08	MoM	MoM	IHS	IHS	MoM	YoY	IHS	IHS	MoM	YoY	IHS	IHS
WPU101	IHS	IHS	IHS	IHS	MoM	MoM	IHS	IHS	MoM	YoY	IHS	IHS
WPU103103	MoM	MoM	IHS	IHS	MoM	MoM	IHS	IHS	MoM	MoM	YoY	IHS