Technomics
Better Decisions Faster

Re-Tooling the Estimator's Approach to Escalation Forecasting

Maya Bell – mbell@technomics.net

Jake Cronin – jcronin@technomics.net

Sean Wells – swells@technomics.net

International Cost Estimating and Analysis Association

2/18/2025

Abstract

Today's DoD contracting space is flush with economic price adjustment clauses, supply chain shocks, and half-century O&S plans. This environment necessitates that estimators treat escalation forecasts with more scrutiny than ever before. However, the toolbox for reviewing the volatility in escalation forecasts is severely limited. This paper builds off prior research and takes a two-pronged approach to enhancing estimators' ability to forecast escalation volatility, producing more impactful insights for decision makers. First, our team has developed an ARIMA model utilizing available BLS data that forecasts escalation indices seasonally, enabling identification of commodity-level trends spanning numerous periods. Second, this model is compared to industry escalation forecasts to highlight where forecasts are historically tumultuous. This modeling advances the tool suite and offers repeatable steps for estimators to develop risk guidance and tailored approaches to escalation forecasts and better inform estimate products reliant on them.

Keywords: *Data-Driven, Methods, Modeling, Risk, Uncertainty, Escalation*

Table of Contents

Table of Figures

Table of Tables

1. Introduction

1.1 The Challenges of Escalation

Any decent cost model, budget forecast, or analyses of alternatives includes factors to forecast escalation. However, recent supply chain shocks in the post-pandemic world are underscoring the inherent volatility in long-term forecasting. This era of uncertainty aligns with a recognition that guidance on inflation rates has dropped below the values seen in recent years. There is also an increased focus on the cost of sustaining units, made evident in the initiation of Sustainment Reviews in 2017 (Bonich, Kidwell, & Liss, 2023). This confluence of events has also led to policy change, with the Office of the Secretary of Defense (OSD) publishing a memo in 2022 reinvigorating a focus on economic price adjustments, long-range budget estimates, and the high-quality cost models necessary to support them (Office of the Secretary of Defense, 2022). The cost community faces the question of whether current escalation analysis procedures, particularly those involving risk analysis and long-term economic forecasting, are up to the challenge.

1.2 Empowering the Estimator's Toolkit

Our hypothesis is that with a few updates to the "Escalation 101" bag of tricks, the cost community can overcome this challenge while allowing analysts to innovate in estimating escalation. The use of historical data and advanced modeling techniques results in improvements to the quantification and handling of common escalation risks and should be adopted as common practice by the cost community. Building on a smaller-scale model presented at the International Cost Estimating and Analysis Association (ICEAA) workshop in 2024 (Wells & Hagy, 2024), this paper leverages hundreds of escalation indices and an Auto-Regressive Integrated Moving Average (ARIMA) model to derive insights into forecast volatility, historical accuracy, and index stability. The resulting framework is a powerful, data-driven, and repeatable method of assessing the risk associated with escalation indices to drive improved accuracy and defensibility of cost models.

The paper begins with a brief background on escalation, the dataset, the fundamentals of ARIMA modeling, and the statistical measures of performance leveraged in later analysis. Both the ARIMA model (used to assess forecast accuracy) and the simple Moving Average Model (used to develop predictions as a comparison to the commonly used Global Insight tool) are described in detail. This helps clarify the findings regarding index accuracy, volatility, and stability across varying sectors. Finally, armed with these findings, the escalation risk toolkit is built out with a step-by-step framework and application recommendations.

2. Background

2.1 Data Overview

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 CAPE, 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 1, pulled from OSD Cape).

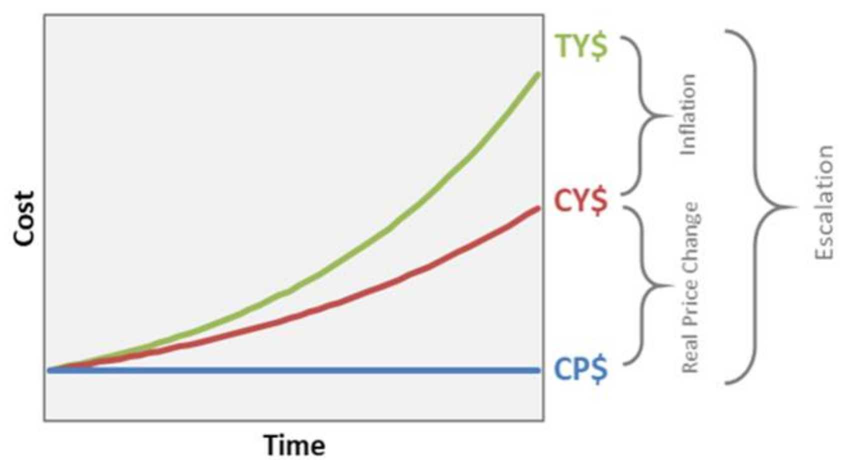

Figure 1 - Escalation Equation

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 those commodities. Another key term is the Compound Annual Growth Rate (CAGR), a standard measure of the escalation rate over a given period, which provides the average rate of compound growth excluding period-to-period variance (see Figure 2).

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

where:

EV = Ending value

BV = Beginning value

n = Number of years

Figure 2 - CAGR Formula

The data analyzed came from two sources: the Bureau of Labor Statistics (BLS) and Information Handling Services (IHS) Global Insight (GI) forecasts. The study team utilized the BLS for actual values and to derive moving averages, while Global Insight data was used for industry prediction values. The team sourced BLS data from 2005 to 2024 to support moving average calculations, ensuring alignment with Global Insight data over the same period for comparison. The Cost Assessment and Program Evaluation (CAPE) guidebook on inflation cites Global Insight as the most common private source used by the DoD for forecasts, as it includes forecasts of almost all indices produced by the BLS (OSD CAPE, 2017).

With thousands of indices to choose from in the Global Insight catalog, the list was narrowed based on several criteria: common usage, broad scope, accuracy/reliability (some indices had missing years), and the data range (some indices started after 2005, which would not have been sufficient for analysis). Even this consolidated list was long, so to consolidate the analysis into more manageable sectors, the indices were sorted into twelve categories based on the Global Insight classification: Chemicals, Construction, Defense and Aerospace, Electronics, Energy, Macro Price, Maintenance, Repair, and Operations (MRO), Metals, Non-Electric Machinery, Steel, Transportation, and Wages (see Appendix B – Index Catalog for further details). The majority of future analysis uses aggregations of these forecasts. For each index, the Global Insights index number was cross-referenced against several data sources to ensure the analogous BLS index number was identified, demonstrated in Table 1.

Table 1 - Sample Index Cross-Reference

Section	GI ID	GI Title	BLS ID	BLS Title
Average Hourly Earnings (AHE)	MEXLAB0016	Construction	WPUSI012011	Construction Materials
Building Materials	WPIP081204	PPI, Hardwood Flooring	WPU0812	Hardwood Lumber
Chemicals	WPIP0679	Misc. Chemical Products & Preparations	PCU3259--3259--	Other Chemical Product & Preparation Mfg
Electrical Components	PPI33441K6	Electronic Coils, Transformers & Other Inductors	PCU33441K33441K6	Electronic Coils, Transformers, And Other Inductors
Macro Price	V79309848	PPI, Machinery and Equipment	WPU11	Machinery And Equipment
Steel	WPIWP10	Metals & Metal Products	WPU10	Metals And Metal Products
Transportation Equip	PPI336412	Aircraft Engines & Engine Parts	PCU336412336412	Aircraft Engine And Engine Parts Mfg

Some difficulties within the escalation dataset included index IDs being tricky to align between the BLS and Global Insight and some BLS indices having earlier years of data omitted in the online search tool. This required data validation and continual revisiting of the index list, yielding the first mini-finding of the paper – carefully review the assumptions associated with any index used to ensure that the scope and available forecasts align with estimating needs.

2.2 Intro to ARIMA Modeling and Forecasting

The ARIMA model is used as a regression method for time series forecasting with the objective of describing autocorrelation, trends, and relationships between observational and residual errors (Statistics Easily, 2023). The goal of the ARIMA model is to select the best fitting model to use for forecasting timeseries data. The best fitting model will have the lowest AIC, stationary data, and data that will pass a confidence interval check at 95%. This paper will only scratch the surface of ARIMA models, focusing on factors affecting applicability such as key metrics, parameters, and assumptions necessary for cost analysis. The typical trends of the ARIMA model are stationary, non-stationary, and seasonality. Stationarity is a way to describe cyclical datasets that do not contain predictable patterns and have constant variance, while non-stationary datasets contain a trend. Seasonality is a repetitive trend that occurs in non-stationary data.

The ARIMA model used in this paper was developed entirely open-source, primarily using R Studio and the “auto-Arima” library (Hyndman & Khandakar, 2008). At their core, the benefits of using ARIMA modeling for escalation cost analysis are twofold: ARIMA models are specifically designed to assess the data-driven goodness-of-fit for time series data, and ARIMA models clearly demonstrate trends that can be replicated in long-term forecasts and cost models. The benefit of using an ARIMA model for escalation data is primarily focused on the fact that escalation data is timeseries data, where data trends can be cyclical, seasonal, or random. The ARIMA is a better choice as a model when compared to a linear regression, as the ARIMA is shows the data as it is observed and uses time components to forecast its data. The ARIMA model as has the potential of addressing independence between observations, normality of residuals, along with homoscedasticy (variables that have constant variances), unlike the linear regression model.

The team focused on Root Mean Squared Error (RMSE) and Root Mean Absolute Error (RMAE) as the primary measures of forecast accuracy. RMSE is a measure of the difference between the predicted values and actual values in a statistical model; a small RMSE indicates less of a difference and thus a lower error compared to a large RMSE (Kumar, 2024). RMSE is particularly beneficial for our escalation dataset because it works well for comparisons of model performance across different scales (as many indices have different bases). By comparison, RMAE has the advantage of being less sensitive to large outliers (as the errors are not squared), making it a robust metric in a dataset that includes real-world inflationary spikes. These metrics are general to any predictive model, while ARIMA-specific tests and statistics are covered in Section 3.2.

2.3 Flow of Analysis

Figure 3 outlines the four primary steps within our analysis.

Figure 3 - Flow of Analysis

First, the team queried the aforementioned BLS and Global Insight data to create prediction forecasts. For the BLS data, a simple moving average was leveraged to predict the future year index values. For the Global Insight data, the predicted values for each year were directly sourced from the provided Global Insight files. Second, the CAGRs were compared for the Moving Average Model forecast, the Global Insight forecast, and the actual BLS occurrences (i.e., the control representing actual historical data) to determine forecast error. Third, the ARIMA model was utilized to provide further insight into model goodness-of-fit and time series applicability. Finally, these insights were applied to the uncertainty process and draw conclusions about forecast accuracy.

3. Forecast Models

3.1 BLS Moving Average Forecast Model

The team leveraged Power Query to automate data transformation and calculate CAGR metrics for both the actual BLS data and the Global Insight predictions. The Global Insight data for a given year contains projections for each of the following ten years (for example, the 2009 file contains predictions for 2010 through 2019); this prediction range is mirrored for the moving average forecast.Figure 2 Growth rates were chosen for comparison rather than escalation rates because the base values of the BLS and Global Insight data were often inconsistent, which is mitigated by using a unitless growth rate.

The calculated expected annual growth rate was determined by developing a moving average from one, two, or three years of prior BLS data, which becomes the forecasting rate. This rate is then projected forward across a span of one through ten years. Table 2 below shows the forecast CAGR for 1, 2, 5, and 10 years for sake of conciseness. For more information on the data processing steps, see Appendix A – Moving Average Model Details.

Table 2 - Sample Forecast Data

BLS ID	Start Year	Data Years	Predict Years	BLS CAGR	Model CAGR	GI CAGR
PCU3222113222110	2010	1	1	4.04%	5.30%	6.36%
PCU3222113222110	2010	1	2	2.10%	5.30%	3.82%
PCU3222113222110	2010	1	5	3.02%	5.30%	1.81%
PCU3222113222110	2010	1	10	2.58%	5.30%	1.78%
PCU3222113222110	2010	2	1	4.04%	3.31%	6.36%
PCU3222113222110	2010	2	2	2.10%	3.31%	3.82%
PCU3222113222110	2010	2	5	3.02%	3.31%	1.81%
PCU3222113222110	2010	2	10	2.58%	3.31%	1.78%
PCU3222113222110	2010	3	1	4.04%	4.31%	6.36%
PCU3222113222110	2010	3	2	2.10%	4.31%	3.82%
PCU3222113222110	2010	3	5	3.02%	4.31%	1.81%
PCU3222113222110	2010	3	10	2.58%	4.31%	1.78%
CIU1020000520000I	2010	3	1	1.48%	2.21%	2.14%
CIU1020000520000I	2010	3	2	1.73%	2.21%	2.06%
CIU1020000520000I	2010	3	5	1.98%	2.21%	2.11%
CIU1020000520000I	2010	3	10	2.82%	2.21%	2.27%
CIU1020000520000I	2010	2	1	1.48%	1.87%	2.14%
CIU1020000520000I	2010	2	2	1.73%	1.87%	2.06%
CIU1020000520000I	2010	2	5	1.98%	1.87%	2.11%
CIU1020000520000I	2010	2	10	2.82%	1.87%	2.27%
CIU1020000520000I	2010	1	1	1.48%	1.73%	2.14%
CIU1020000520000I	2010	1	2	1.73%	1.73%	2.06%
CIU1020000520000I	2010	1	5	1.98%	1.73%	2.11%
CIU1020000520000I	2010	1	10	2.82%	1.73%	2.27%

3.2 ARIMA Model Parameters and Tests

Our ARIMA model builds off the simplistic Moving Average projections. Since our time series forecasts only span ten years (limited by the length of benchmark GI forecasts), we have a small number of residuals, resulting in a substantial amount of variance remaining in candidate predictive models.

The ARIMA model possesses several parameters including *p*, *d*, and *q*; these are determined through a series of tests listed below. They will inform the ARIMA model output expression (p,d,q), which will be shown in the processing steps.

1. The parameter *p* represents the number of lag observations included in the model (lag referring to the number of steps back in time a past observation is from the current time)
2. The parameter *d* indicates the degree of differencing (referring to the differences between consecutive observations) required to achieve stationarity
3. The parameter *q* is the size of the moving average window

An ARIMA model of (0,0,0) for (p,d,q) represents the best possible fit. ARIMA modeling also requires a deviation from more typical measures of success for OLS regression, leveraging unique test plots and transformations (Frost, 2025):

1. The Augmented Dickey-Fuller (ADF) plot helps determine the *d* parameter by the type of trend decay and the confidence intervals
2. The Autocorrelation Function (ACF) plot helps identify the *q* parameter by showing the correlation between observations at different lags
3. The Partial Autocorrelation Function (PACF) plot determines the *p* parameter by showing the correlation of an observation with its lagged values after removing the effects of intermediate lags
4. The Akaike Information Criteria (AIC) quantifies the goodness of fit and simplicity of a model in a single statistic, with a lower AIC considered better
5. The Box Cox transformation is a power transformation that is used to adapt non-stationary data into stationary data

The ARIMA model possesses lags. Lags are the time difference between two observations in a time series or the delay, they provide insight into trend identification in the ACF and PACF tests. They also help inform whether the ARIMA model will be comprised of an autoregressive (AR) component or a moving average (MA) component. For this paper we are only dealing with AR components, meaning that the ACF test visually quickly decays towards zero, the PACF test quickly drops to zero, and that the ACF shows a positive first lag.

The results of our ARIMA modeling will be detailed further in section 4.3 to provide additional insights into the behavior of a common DoD index.

4. Forecast Comparison

4.1 Sector Performance

As expected, as prediction years increased across all indices, forecast accuracy decreased. The RMSE associated with the GI CAGR error (delta between the actual BLS value and the GI predicted value) was normalized by multiplying by the number of prediction years to scale the time series (i.e., the five-year CAGR error was multiplied by five, ten-year CAGR error was multiplied by ten, etc.). The following five sectors had the lowest five-year CAGR RMSEs:

Table 3 - Top-Five Performing Sectors

Category	1-Yr CAGR RMSE	2-Yr CAGR RMSE	5-Yr CAGR RMSE	10-Yr CAGR RMSE
Defense and Aero	0.9%	1.2%	3.8%	11.3%
MRO	2.3%	3.8%	6.1%	12.8%
Non-Electric Machinery	1.1%	2.0%	4.7%	9.1%
Transportation	1.9%	3.0%	6.0%	7.7%
Wages	2.0%	1.8%	1.4%	1.1%

One outlier worth discussion in Table 3 is the drop in RMSE for the ten-year ‘Wages’ CAGR when analytic experience would exert more variance in a prediction twice as far out. Upon investigation this was determined to be an artifact of the date range (2008-2014 as starting years, thus forecasting out to 2018-2024); high wage increases in the 2021-to-2023-time frame meant that what was a slightly high forecast actually aligned better with reality. This should not be interpreted that a five-year wage estimate will consistently be more accurate than a two-year version, but instead as a reminder to both expect the unexpected and to model historical variance whenever possible.

In general, any of the sectors not referenced in the prior section were still in the 8-10% RMSE range at five years (encompassing of Chemicals, Construction, Electronics, Macro Price, Metals, and Steel). Those indices also all had only mild increases in error over the six-to-ten-year span, as demonstrated in Figure 4. These indices have an increased level of disparity between the Global Insight and Moving Average Model forecasts, with the MA model often having double or triple the forecast error of GI. Together these findings indicate that commercial forecasts may require a bit more risk attention and calibration but are still defensible for estimate purposes.

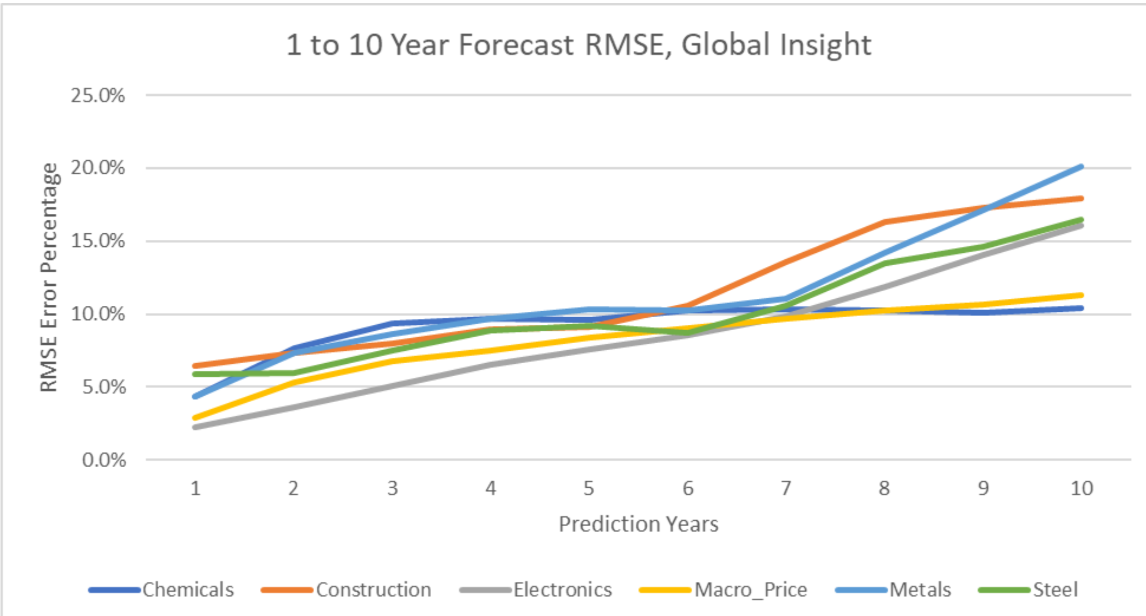

Figure 4 - GI Forecast RMSE

One notable exception is the Energy category (containing indices for Natural Gas, Petroleum, Gas Fuels, and Electric Power), which sits at anywhere from a 30-35% RMSE in the five-to-ten-year prediction span. The Moving Average Model performs about 25% worse at five years (43.2% RMSE for the model vice 30.8% for GI) and 100% worse at ten years (63.4% vice 32.1%). This finding falls in line with widespread economic guidance about energy volatility and agrees with its exclusion from common CPI measures. These error percentages equate to +/- 6% forecast error at the five-year mark, meaning that a program with $10M in annual fuel costs could be half a million over budget; for programs with 30+ year Operations and Sustainment budgets, those costs compound even further (half a million, plus the compound effect of escalation, in deficit per year over that span would be well over $20M).

The Federal Reserve Bank of Dallas's research on electricity price volatility provides a strong economic foundation for understanding why energy escalation indices often perform poorly (Saretto, Shcherbakova, & Lin, 2024). Their findings highlight that the transition from coal to natural gas in electricity generation has increased real-time price volatility, reinforcing the inherent instability of energy markets due to supply and demand fluctuations, structural changes, policy shifts, technological advancements, and geopolitical factors. Escalation indices like those from BLS and GI, which rely on historical trends, struggle to capture these sudden shifts, leading to significant forecasting errors. Rather than fueling frustration among program offices, these findings should empower estimators to perform sensitivity analysis for energy prices, as described in Section 5.2.

4.2 Common DOD Indices

4.2.1 DoD Index CAGR RMSE

Within the DoD, several indices are used commonly in cost estimates and analysis of alternatives to capture real price change in long-term projections. The following sample indices are shown in Table 4 using their GI title and demonstrate the five-year CAGR RMSE for both Global Insight and the Moving Average Model. With the previously noted exception of Energy, the Moving Average Model performed as well as or better than Global Insight at the five-year mark (though GI outperformed at ten years), indicating that moving average forecasts can be leveraged in some estimating situations with a reasonable degree of confidence.

Table 4 - Common DoD Indices

Index #	Index Title	5-Yr CAGR RMSE (GI)	5-Yr CAGR RMSE (Model)
PPI3345111	Aeronautical, Nautical & Navigational Instruments	10.4%	10.4%
WPIP053S	Gas Fuels	14.1%	28.1%
PPI333120	Construction Machinery	13.7%	13.1%
WPIWP101	Iron and Steel	11.6%	11.1%
PPI336611	Ship Building and Repair	10.9%	4.9%

4.2.2 ARIMA Model Deep Dive

The team conducted a deep dive into PPI3345111 (Aeronautical, Nautical, Navigational Instruments) as the index commonly used in the US Navy. By running the ARIMA model on the GI and Moving Average Model forecasts, further assessments can be made on which model is a better fit as well as providing a cross-check for predictions. The analysis first tests for stationarity, then for goodness of fit, and finally review how the ARIMA model would predict beyond ten-years. Similar tests were performed for the Moving Average Model, the results of which are described further down; the plots are shown in Appendix C.

The GI forecast is found to be stationary in three of the four initial ARIMA variance tests:

First, the variance is visually constant (no increasing trend or true random variation) as shown in Figure 5, *suggesting stationarity*.

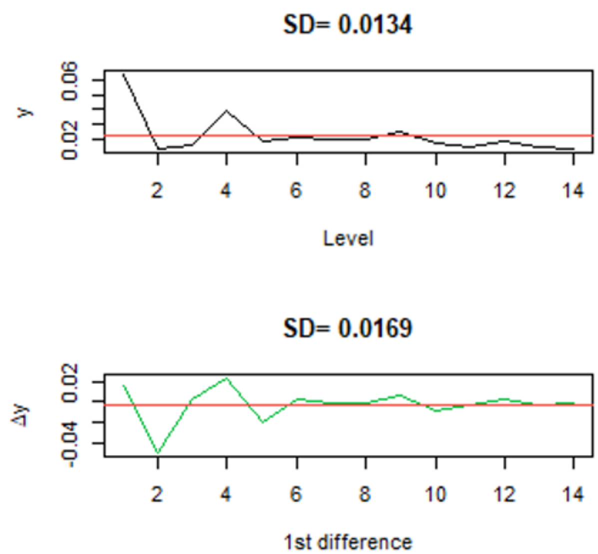

Figure 5 - GI ARIMA Variance Check

Second, the variance does not halve over time, when comparing the initial variance level to the 1st difference (log) variance, which is the necessary benchmark for determining stationarity. In the upper chart of Figure 5, standard deviation (SD) equals 0.0134; this actually increases to 0.0169 in the lower chart, *suggesting non-stationarity*.

Third, the ACF decays to zero quickly (demonstrated by the black tick marks flipping from positive to negative over the time span in Figure 6), *suggesting stationarity*.

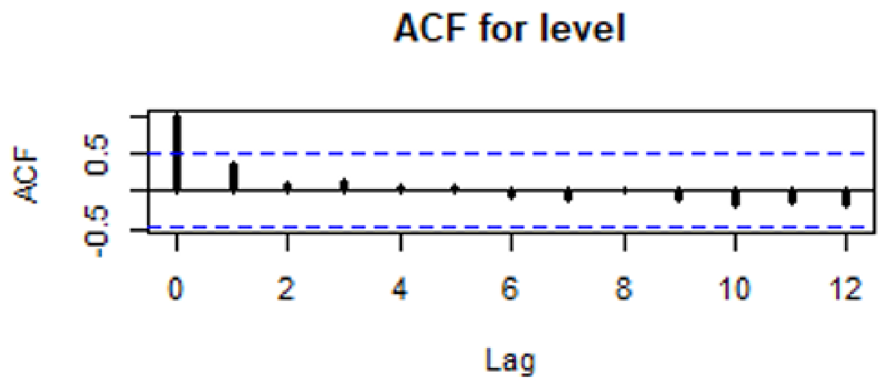

Figure 6 - GI ARIMA ACF Test

Fourth, the ADF test results in a p-value of 0.01 (this is demonstrated in the output shown in Figure 7), *suggesting stationarity.*

```
        Augmented Dickey-Fuller Test


Dickey-Fuller = -8.4787, Lag order = 2, p-value = 0.01
alternative hypothesis: stationary
```

Figure 7 - GI ARIMA ADF Test

Since the data was stationary in three of four tests, the ARIMA analysis was able to proceed. As shown in Figure 8, we have an ARIMA model of (0,1,0), an AIC of -79.58, and residuals that are generally grouped together. A (0,1,0) model means there were zero lags of the ACF test, an autoregressive (AR) 1 model was used, and there are zero lags of the PACF test.

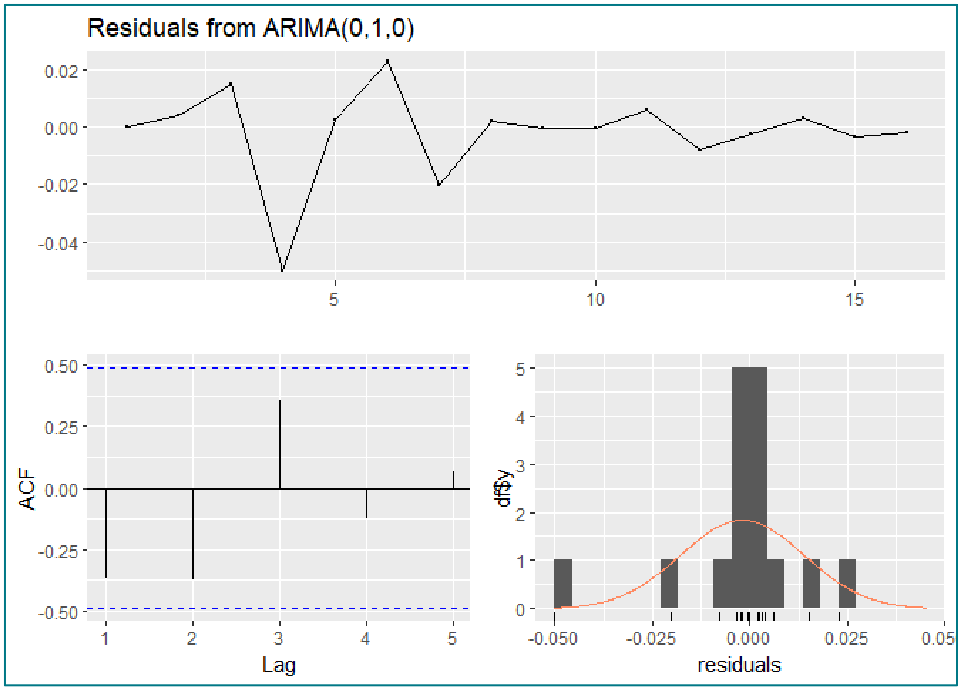

Figure 8 - GI ARIMA Results

These results all mean that the ARIMA forecast is aligned to the GI data, which green lights the ability to leverage this type of forecasting beyond the ten-year period from GI (as shown in Figure 9) Note that "index" represents the forecast periods, which in this model are years.

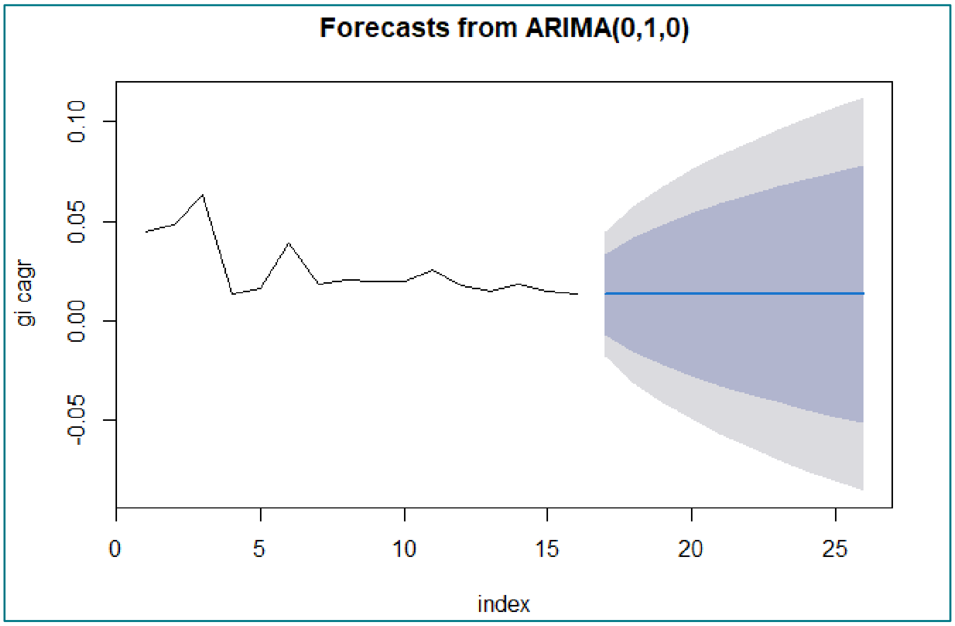

Figure 9 - GI ARIMA Model Forecast

The team repeated the tests for the Moving Average Model (in this case leveraging three years of data). It was found to be non-stationary in three of four tests and yielded a very well-fitting (0,0,0) ARIMA model with a 0.0028 standard error. Therefore, the original Moving Average Model is the best fit, so it should be used in lieu of an ARIMA model.

This brief survey into the ARIMA methodology highlights the ability for estimators to leverage common, proven, and reproducible statistical techniques to justify the use of specific forecast models. In this example, it helped to determine if the GI data is the best model to use for forecasting – for the specific defense index, a moving average was actually slightly superior, but in many cases the ARIMA model would test between numerous prediction sources.

4.3 Methodology Based Limitations

In reviewing the moving average methodology and dataset, three main limitations were identified:

1. The number of years of data available for analysis was limited. Since ten years of known BLS data were required to analyze the GI forecasts, only GI data files up to 2014 (which predict out to 2024) could be used. Additionally, the available GI dataset only extended back to 2006, creating a small range of available years for prediction.
2. Within this limited range, many of the eight-to-ten-year predictions coincide with COVID impacts. As referenced in Section 4.1, some of the high CAGRs ended up appearing more "accurate" in the outyears than in the four-to-six-year range, a result not frequently expected over other time spans.
3. To maintain a manageable number of forecasts, a simple moving average was used instead of a more complex model. While this approach did not necessarily hinder the ability to compare forecasts, a more refined forecasting model—such as an exponentially weighted version—would likely improve accuracy beyond the moving average approach.

Regarding the ARIMA, the team found two limitations that estimators should be aware of (in addition to the standard ARIMA conditions cited in Section 2.2):

1. The datasets need to be non-correlated, meaning that if an observation at one point in time in dataset 1 is increased, it will not cause the observation at same point in time in dataset 2 to also increase. This is not always the case for escalation indices, as they are typically based off the same historical observations.
2. The ARIMA model works best with large quantities of data (i.e., a significant number of prior years of data for forecast development or a forecast over a long span).

5. Risk Factor Derivation

5.1 Literature Review

This is certainly not the first foray by the estimating world into accounting for the risk of higher (or lower) escalation in long-term cost projections. The majority of prior research tends to be specific to the construction space and is focused on broad composite indices (as they pre-date the 2021-era CAPE guidance on development of custom indices tailored to the estimate scope). This review aims to glean common forecasting methods and best practices while also addressing areas that the “escalation toolkit” may need improvement.

A 2012 Booz Allen Hamilton presentation at the SCEA/ISPA workshop detailed a Monte-Carlo based methodology for assessing the prediction error within inflation forecasts (DeCarlo, Jabaley, & Druker, 2012). The impetus was a systematic under-estimation of the military construction (MILCON) escalation rates, leading the authors to calculate the historical standard deviation from the actual inflation rate (between 1-2%) and leverage that figure within Excel’s “Real Time Analytics” to calculate deciles around future estimates.

A 2011-era version of the CAPE Inflation Handbook highlighted the risk of underestimating inflation, noting that in four of seven fiscal years spanning 2004-2010, actual inflation exceeded budgeted inflation, resulting in programmatic shortfalls (OSD CAPE, 2011). The authors note both a “back of the napkin” approach where the deviation between the actual and budgeted inflation rate over a span is calculated, then placed on a probability distribution to find the 80th percentile (P80) of those deviations (in the timespan, this would have been 3.24%, although that is heavily biased by high 2007 rates). The more refined methodology is to fit a curve to the data and leverage a z-distribution, which in this case brings the P80 down to 3.05%.

The authors of a 2016-era Louisiana State University paper on escalation factors in construction leverages the Value at Risk (VaR) technique (common in the financial industry) to assess the rampant material escalation in the construction industry (Joukar & Nahmens, 2016). The methodology takes the common Engineering News Record (ENR) Construction Cost Index, creates an ARIMA model for the monthly escalation factors, then calculates the VaR at the 90 and 95 percent probability levels.

In response to a 2000s-era mandate from the Federal Transit Administration requiring risk analysis on the budget and schedule of transit projects, Touran and Lopez highlighted the implicit link between schedule delays and increased impact of any escalation factors (Touran & Lopez, 2012). The ensuing methodology involves calculating an escalation factor (using historical analysis of common sources like the ENR index to create a mean and standard deviation), then applying a Monte Carlo simulation to the total escalated project cost while varying the escalation factor distribution. In an earlier version of their research, the authors detail an additional method for deriving the escalation factor in which a moving average is calculated with additional weighting applied to the index value in the period prior to the last known index value, evoking an Exponentially Weighted Moving Average (EWMA) approach (Touran & Lopez, 2006).

Across the literature, several key trends stand out:

1. A consistent acknowledgement that estimators should specifically model escalation risk when conducting a Monte Carlo simulation
2. The use of historical index data to derive either a moving average forecast or a distribution around a published forecast

3. The use of Monte Carlo simulations to assess the out-year risk associated with both the fluctuation in escalation rates and the interaction between escalation rates and factors like schedule and material availability

Our “toolkit update”, described in the next section, aims to leverage these common themes in a repeatable and defensible process.

5.2 Uncertainty Process

The 2021 OSD CAPE Inflation and Escalation Handbook provides the foundation for the process of applying inflation and escalation to estimates, including identifying the input data type, normalizing the dataset, selecting the desired output type, and clearly labeling the results (OSD CAPE, 2021). Within the general instructions section, CAPE highlights the process for developing custom indices, which can include a weighted average of several indices (to align to a program’s scope), research into real price change, and the application of outlay profiles. The escalation uncertainty analysis process begins once these general steps are accomplished, as the OSD CAPE handbook is quiet on the subject.

5.2.1 Calculating the Historical Uncertainty in the Forecast(s)

The specific calculations for this step will vary depending on the forecast source. If the escalation factor was internally developed (i.e., an ARIMA model, EWMA, simple moving average, or even an agency-mandated escalation rate), leverage the same development process using data from five to ten years ago. The ARIMA steps outlined in Section 4.2.2 should be applied to ensure the best-fitting model is used for forecasts. The benefit here is that this factor can be compared to reality by lining up the resulting “forecasts” against the actual BLS values for those years. The raw values produced in this step should be a percentage delta for each year of historical forecasting; if the escalation factor is a composite of multiple indices, derive these values for each index and weight the percent deltas in accordance with the composite weighting.

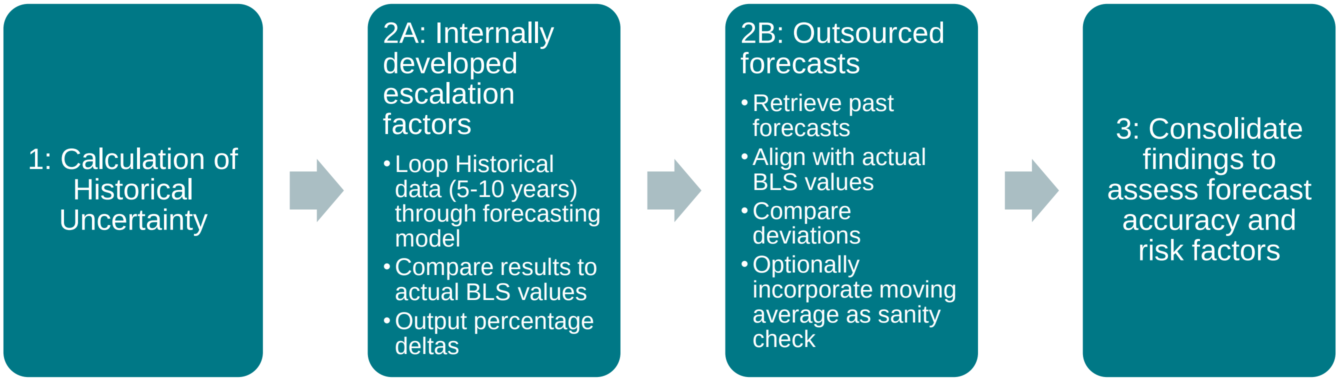

Figure 10 - Uncertainty Process

For organizations looking to move away from policy mandates and to data-driven, internal development of escalation factors, recent research on this topic provided strong evidence that more years of moving average data improves the forecast viability (Wells & Hagy, 2024). Optimization of the number of moving average years of data most appropriate for a given sector is eminently feasible given the quantity of historical BLS data and should be studied during the development of any custom escalation factors. The key is to use the same parameters (years of data, any manipulations to the forecast) in the test case (i.e., years where actual BLS data exists for accuracy comparisons) as in the true forecast. Our dataset demonstrated a median negative correlation of 95.9% between the number of moving average years and the CAGR RMSE (since a lower error is a good thing, the correlation against increased years of data is expected to be negative). Table 5 displays this data (the correlation was assessed between the 5-Year CAGR RMSE value and the number of years of moving average data).

Table 5 - Five-Year CAGR RMSE for Moving Average Model

Category	Data Yr = 1	Data Yr = 2	Data Yr = 3	Correlation
Chemicals	6.29%	4.52%	3.90%	-96.36%
Construction	7.84%	6.51%	6.24%	-93.41%
Defense & Aerospace	0.60%	0.63%	0.70%	97.79%
Electronics	1.94%	1.68%	1.54%	-98.55%
Energy	15.49%	10.76%	8.65%	-97.66%
Macro Price	3.78%	2.84%	2.49%	-96.75%
MRO	2.51%	2.21%	2.12%	-95.38%
Metals	4.74%	3.57%	3.39%	-92.04%
Non-Electric Machinery	1.28%	1.24%	1.32%	52.31%
Steel	6.47%	4.58%	3.53%	-98.68%
Transportation	2.30%	1.73%	1.61%	-93.60%
Wages	1.91%	1.59%	1.45%	-97.63%

If an outsourced forecast is used, review the prior year forecasts against the actual BLS values (i.e., locate the forecast file from five or ten years ago and line the predicted index values up against the BLS values). Developing a separate moving average forecast for the same period can be a useful sanity check when using outsourced forecasts, as it provides a barometer for the level of deviation between the outsourced forecast and what the historical trend represents. This “deviation” is by no means a bad thing, as forecasting agencies can provide significantly more calibration for future events, anticipated trends, and potential unknowns as compared to a simple moving average.

5.2.2 Develop the Uncertainty Distribution

Once the annual deviations from the actual BLS values are determined, several basic summary statistics help to shed light onto the overall uncertainty. First, taking the mean and standard deviation of the data provides the basis for the uncertainty distribution. Second, reviewing the minimum, maximum, and median, as well as performing sensitivity analysis by removing years of interest, helps to bound the uncertainty. As a best practice, estimators should perform the following steps with and without noted outliers (for example, the 2022 inflationary spike) to get a sense of both the forecast accuracy in stable periods and the level of deviation in unstable periods. If time and resources allow, running ARIMA analysis, based on the process laid out in Section 4.2.2, will provide additional measures of stationarity, seasonality, and goodness-of-fit.

5.2.3 Incorporating Escalation Uncertainty Distributions into Monte Carlo

This section focuses on the specific challenges and steps needed to vary escalation in a risk model, leaving the overall process and theory of risk to existing well-established resources such as the Joint Cost Schedule Risk and Uncertainty Handbook. In simplest form, the rate of escalation would just be one of the risk inputs that is varied across a triangular, normal, or other common distribution. However, three common scenarios merit additional consideration in the process:

One of the biggest challenges with including escalation as a risk parameter is the need to convert a then-year estimate into a constant-year value for comparison to program baselines. A simple solution to this challenge is to take the median escalation value across all trials and using it to de-escalate the estimate, though more complex adjustment methods exist. If users need to avoid risk on escalation in final estimates explicitly, one recommended way to assess the impact of escalation variance on the overall estimate is to run sensitivity trials for each decile of the escalation uncertainty distribution. This provides a discrete histogram of the escalation impact while continuously varying the other risk parameters in the model and thus capturing those interactions, which is sufficient data to inform most programmatic needs.

Additionally, many programs may want to assess real price change as a variable, but keep baseline inflation consistent with agency guidance (for example, the DoD Joint Inflation Calculator). In this case, it may be prudent to include a “real price change” value that is added to fixed standard inflation indices (for example, a 0.5% adder to 2.1% inflation becomes 2.6% escalation), then to apply the escalation risk distribution from prior steps only to that adder.

Another area to consider is the interaction between escalation and schedule, as project delays are common and buying labor and materials in the future is more expensive than doing so in the present. Whether the schedule risk is informed by a discrete risk register, a formal schedule risk analysis, or other common methods of expanding and contracting the period of performance, the parameter will exacerbate the impact of escalation risk. Therefore, it is recommended to run sensitivity analysis where all other parameters are toggled to remain constant whilst only varying the schedule and escalation to get a sense of their impact on the overall variance.

5.2.4 Compare Point and Varied Trials for Escalation

The final recommended step in assessing the impact of escalation variance is to create a counterfactual scenario in which there is no variance. This consists of an additional set of Monte Carlo trials in which the escalation risk parameter is toggled off while all other parameters are varied. Overlaying the S-curve for both sets of trials (see Figure 11) provides an additional indication of the impact of escalation risk on a project and enables a discrete calculation of the P80 value with and without said risk.

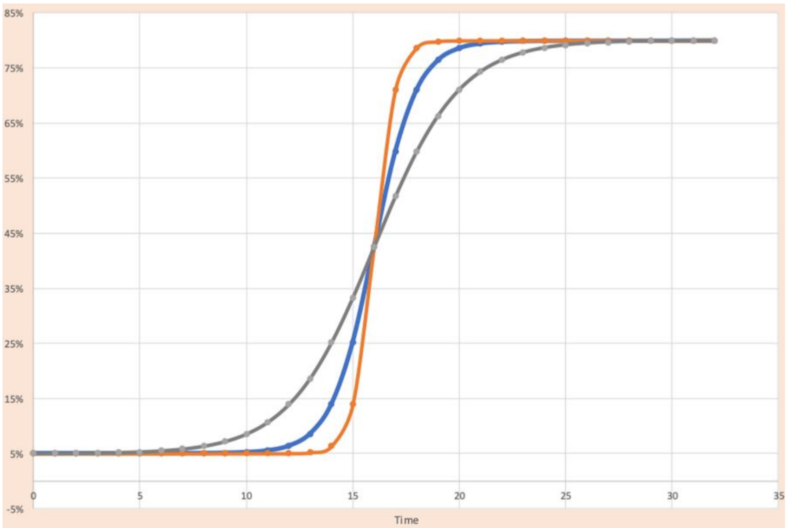

Figure 11 - Notional Overlayed S-Curves

The final presentable figures are dependent on which combination of the aforementioned risk techniques are used for a given estimate and what stakeholders are interested in seeing, but the bottom line is that providing insight into escalation risk is vital for programmatic success and using a suite of repeatable methods is the best way to provide that critical information.

6. Conclusion - Applying the Toolkit

A 2024 DODCAS presentation noted the challenge for analysts building their own weighted indices, as the process requires a number of non-standardized sources in addition to the application of outlay profiles (OSD CAPE, 2024). This briefing highlighted an emerging OSD-CAPE library of recommended indices to streamline this process. Building on these tailwinds, this paper helps answers the question of "once I have my customized index, how do I defend it?" by providing a defensible, repeatable risk analysis framework:

1. Calculate the Historical Uncertainty in the Escalation Forecast
2. Develop an Uncertainty Distribution for that Forecast
3. Run Sensitivity and Risk Analysis using Monte Carlo trials
4. Compare Trials With and Without Escalation Variance

This framework has a myriad of primary applications, spanning from risk and sensitivity in life cycle cost estimating, annual budgeting and executability planning, to negotiations on economic price adjustment (EPA) clauses in high-value contracts. Our findings on sector performance (with Energy standing out as the most volatile) helps scale the assumed risk as estimators evaluate overall uncertainty. The use of repeatable ARIMA modeling techniques can help quantify the goodness-of-fit for various time-series forecasts and improve defensibility of the coveted custom escalation indices that agencies often seek.

A secondary benefit of this framework is that analysts employing it will improve familiarity with the historical behavior of escalation indices and the drivers of escalation uncertainty, making for more grounded and defensible products. As the challenges of escalation, a globalized supply chain, and flattened budgets continue to emphasize the need for high quality, justifiable, and actionable cost analysis, the cost community should answer the call by leveraging available data and modernizing its approach. The toolkit now includes escalation risk analysis next to the other tried and true data-driven methods our community has leveraged in support of leadership for decades.

7. Bibliography

Bonich, Kidwell, & Liss. (2023). You've Invested, Now Sustain It: Insights Into the Sustainment Review Process. *International Cost Estimating and Analysis Association.* San Antonio.

DeCarlo, M., Jabaley, S., & Druker, E. (2012). Inflation Cost Risk Analysis to Reduce Risks in Budgeting. *SCEA/ISPA.*

Frost, J. (2025). *Autocorrelation and Partial Autocorrelation in Time Series Data*. Retrieved from Statistics by Jim: https://statisticsbyjim.com/time-series/autocorrelation-partial-autocorrelation/

Hyndman, R., & Khandakar, Y. (2008). Automatic time series forecasting: The forecast package for R. *Jounal of Statistical Software*, 1-22.

Joukar, A., & Nahmens, I. (2016). Estimation of the Escalation Factor in Construction Projects Using Value at Risk.

Kumar, A. (2024). *MSE vs RMSE vs MAE vs MAPE vs R-Squared: When to Use?* Retrieved from https://vitalflux.com/mse-vs-rmse-vs-mae-vs-mape-vs-r-squared-when-to-use/

Office of the Secretary of Defense. (2022). Guidance on Inflation and Economic Price Adjustments.

OSD CAPE. (2011). Department of Defense Inflation Handbook 2nd Edition.

OSD CAPE. (2017). Inflation and Escalation Best Practices for Cost Analysis Handbook.

OSD CAPE. (2021). Inflation and Escalation Best Practices for Cost Analysis Handbook.

OSD CAPE. (2024). Standardizing and Simplifying Escalation Indices. *DODCAS.* Alexandria.

Saretto, A., Shcherbakova, A., & Lin, J. (2024). *What Fuels the Volatility of Electricity Prices?* Federal Reserve Bank of Dallas.

Statistics Easily. (2023). *What is the ARIMA Model*. Retrieved from Statistics Easily: https://statisticseasily.com/glossario/what-is-arima-model-comprehensive-guide/

Touran, A., & Lopez, R. (2006). Modeling Cost Escalation in Large Infrastructure Projects.

Touran, A., & Lopez, R. (2012). Modeling Cost Escalation as a Risk Factor in Construction Projects.

Wells, S., & Hagy, K. (2024). Well, That Escalated Quickly - A Novel Approach to Escalation Forecasting. *International Cost Estimating and Analysis Association.* Minneapolis.

8. Appendix A – Moving Average Model Details

To clean the data and calculate inputs required for ARIMA analysis, the team performed several data cleaning and modeling steps, including:

1. Cross referenced each Global Insight index with a corresponding BLS index (based on the index titles and available cross-reference data sheets for PPI indices on the BLS website)
2. Gathered data from the Bureau of Labor Statistics website (the BLS Series Report allows entry of multiple indices at a time, designated time spans, and either excel or csv outputs)
3. Pivoted the index data to create a table in which each row represented a unique index title, year of measurement, and index value for every index across the time span of our data analysis
4. Added a column that specifies how many years of prior data that would be to calculate our growth rate from historical data (1, 2, or 3 years in our particular example, but a larger number of moving average years or an optimization is feasible)
5. Calculated the compound annual growth rate (CAGR) of each index from one, two, and three years ago to present; this served as the forecast rate for our Moving Average Model
6. Added a column that specifies how many years the model is forecasting (one through ten years in our case study, as that is as far as Global Insight predicts)
7. Calculated the expected index values for one through ten years into the future by utilizing the calculated compound annual growth rate and the starting year BLS index value
8. Mapped in the actual realized BLS index in the prediction year and the actual BLS index in the start year, then calculated the actual CAGR for comparison purposes
9. Pulled in the present and future Global Insight predicted escalation rates and calculated its predicted CAGR for comparison purposes

Several challenges were faced during the data cleaning process required to gather and organize the indices used in our analysis, as well as during the calculation of annual growth rates. The lessons learned not only provide insight into our analysis but also serve to develop general cautions during the risk process.

- The utilization of Power Query to clean and organize our data worked well up to a certain point; however, it lacked flexibility in handling outliers. In particular, several of the Global Insight files contained hard-coded “#N/A” or “#DIV/0” errors that caused significant query challenges. Power Pivot was found to be a more suitable environment for calculating growth rates and tracking performance metrics.
- During the data gathering and cleaning process, the team encountered multiple cases where it was not immediately clear which Bureau of Labor Statistics (BLS) index corresponded to which Global Insight index. This ambiguity introduced uncertainty into our dataset and required some estimator judgement in creating a cross-reference.
- Additionally, for the same index and/or commodity, the databases from BLS and Global Insight often tracked escalation rates over different time periods. Some indices did not have BLS data prior to the early 2010s, while others had years lacking data entirely. These discrepancies negatively impacted our calculations and limited the scope of data the team could effectively utilize. Carefully review the assumptions associated with different indices to ensure appropriate context is provided in the analysis.

9. Appendix B – Index Catalog

The following table includes the total list of indices used in our analysis, shown with their Global Insight Index Number, Title, and Category.

Table 6 - Index Catalog

Category	Global Insight ID	Global Insight Name
Chemicals	PROPCH	Propylene, Chemical Grade
Chemicals	WPIP061	Industrial Chemicals
Chemicals	HDPEB	Polyethylene Resin, High-Density, Blow Molding
Chemicals	WPIP0679	Misc. Chemical Products & Preparations
Chemicals	PPI325612	Polish & Other Sanitation Goods
Chemicals	PPI326211	Tire Manufacturing Except Retreading
Chemicals	PPI339991	Gaskets, Packaging, And Sealing Devices
Chemicals	WPIP072	Plastic Products
Chemicals	PPI325212	Synthetic Rubber
Chemicals	WPIP0622	Paint Materials
Chemicals	WPIP06	Chemicals & Allied Products
Chemicals	PPI322211	Corrugated And Solid Fiber Boxes
Construction	WPIP0811	Softwood Lumber
Construction	WPIWP1332	Concrete Pipe
Construction	PPI3241210	Asphalt Paving Mixtures & Blocks
Construction	PPI3241221	Roofing Asphalts, Pitches, Coatings & Cements
Construction	WPIP08	Lumber & Wood Products
Defense&Aerospace	PPI336612	Boat Building And Repair
Defense&Aerospace	PPI3345111	Aeronautical, Nautical & Navigational Instruments
Defense&Aerospace	PPI3345113	Search, Detection, Navigation, & Guidance Systems
Electronics	PPI335313	Switchgear & Switchboard Apparatus
Electronics	PPI335312	Motors & Generators
Electronics	PPI335932	Noncurrent-Carrying Wiring Devices
Electronics	PPI334210	Telephone Apparatus
Electronics	PPI334220	Radio & Television Broadcast & Wireless Comm. Equip.
Electronics	PPI335911	Storage Batteries
Electronics	WPIP117906	Engine Electrical Equipment
Electronics	PPI3353133	Panelboards
Electronics	PPI3342	Communication Equipment
Electronics	PPI334111	Electronic Computers
Electronics	PPI3344131	Integrated Circuits
Electronics	PPI334417	Electronic Connectors
Electronics	PPI3341	Computers And Peripheral Equipment
Electronics	PPI334510	Electromedical Apparatus Manufacturing
Electronics	PPI334511	Search, Detection, Navigation, Guidance, Aeronautical, And Nautical Sys And Instruments
Electronics	PPI334512	Automatic Environmental Controls
Electronics	PPI334515	Electrical Measuring & Testing Instruments
Electronics	PPI334516	Analytical Laboratory Instruments
Electronics	PPI333314	Optical Instruments And Lenses
Energy	WPIP0531	Natural Gas
Energy	WPIP0552	Commercial Natural Gas
Energy	PPI324110	Petroleum Refineries
Energy	WPIP0576	Lubricating Oils And Greases
Energy	WPIP0543	Industrial Electric Power
Energy	WPIP053S	Gas Fuels

Category	Global Insight ID	Global Insight Name
Energy	WPIP0542	Commercial Electric Power
Energy	PPI3241101	Gasoline
Macro Price	CUSAFNS	Cpi, Food & Beverage
Macro Price	CUSASNS	Cpi, Services
Macro Price	WPIS20S	Intermediate Materials, Supplies, And Components
Macro Price	WPIPIND	PPI, Industrial Commodities
Macro Price	WPIP03T15M05	PPI, Industrial Commodities Excluding Energy
Macro Price	WPIP05	Fuels & Related Products, & Power
Macro Price	WPIWP10	PPI, Metals And Metal Products
Macro Price	ECIPWTNS	Employment Cost Index, Wages And Salaries, All Private Industry Workers
Macro Price	ECIPBTNS	Employment Cost Index, Benefits, All Private Industry Workers
Macro Price	CWSA0NS	Cpi, All Items, All Urban Wage Earners And Clerical Workers
Macro Price	CUSAHNS	Cpi, Housing
Macro Price	CUSATNS	Cpi, Transportation
Maintenance, Repair, and Operations	PPI327910	Abrasive Products
Maintenance, Repair, and Operations	WPIWP133	Concrete Products
Maintenance, Repair, and Operations	WPIP134	Structural Clay Products
Maintenance, Repair, and Operations	WPIP1144	Industrial Material Handling Equipment
Maintenance, Repair, and Operations	PPI335931	Current-Carrying Wiring Devices
Maintenance, Repair, and Operations	WPIP0915	Other Converted Paper Products
Maintenance, Repair, and Operations	WPIP1542	Photographic Supplies
Maintenance, Repair, and Operations	PPI325510	Paints And Coatings
Maintenance, Repair, and Operations	PPI325520	Adhesives
Maintenance, Repair, and Operations	PPI325611	Soap & Other Detergents
Maintenance, Repair, and Operations	PPI326	Rubber And Plastics
Maintenance, Repair, and Operations	PPI3324393	Metal Barrels, Drums And Pails
Maintenance, Repair, and Operations	PPI333414	Heating Equipment Excl Warm Air Furnaces
Maintenance, Repair, and Operations	PPI333911	Pumps And Pumping Equipment
Maintenance, Repair, and Operations	PPI332991	Ball And Roller Bearings
Maintenance, Repair, and Operations	PPI3339121	Air And Gas Compressors
Maintenance, Repair, and Operations	PPI325910	Printing Ink
Metals	PPI332911	Industrial Valves
Metals	PPI332919P	Other Metal Valve And Pipe Fitting
Metals	PPI332313P	Fabricated Platework
Metals	WPIP107	Fabricated Structural Metal Products
Metals	PPI332321	Metal Windows & Doors
Metals	PPI332722	Bolts, Nuts, Screws, Rivets & Washers
Metals	PPI332913	Plumbing Fixture Fitting And Trim
Metals	PPI332996	Fabricated Pipe And Fittings
Metals	PPI335228	Other Major Household Appliances
Metals	PPI332	Fabricated Metals
Metals	WPIP106	Heating Equipment
Metals	PPI3324311	Steel Cans
Metals	PPI332510	Hardware
Metals	PPI335122	Commercial Lighting Fixtures
Metals	PPI3314913	Titanium And Titanium-Base Alloy Mill Shapes, Excluding Wire
Metals	WPIP1026	Nonferrous Cable And Wire
Metals	PPI335929A	Electronic Wire & Cable
Metals	WPIP1022	Primary Nonferrous Shapes
Metals	WPIP102501	Aluminum Mill Shapes
Metals	JTI6A4V	Titanium Ingot
Non-Electric Machinery	PPI3331321	Rotary Oil Field & Gas field Drilling Machinery & Equipment

Category	Global Insight ID	Global Insight Name
Non-Electric Machinery	WPIP1166	Other Special Industry Machinery
Non-Electric Machinery	PPI333120	Construction Machinery
Non-Electric Machinery	PPI333993	Packaging Machinery
Non-Electric Machinery	WPIP1143	Fluid Power Equipment
Steel	PREBAR	Concrete Reinforcing Bar, Carbon Steel
Steel	WPIWP101	Iron And Steel
Steel	PPI331222	Steel Wire And Drawing
Steel	PPI331210	Iron & Steel Pipe And Tube From Purchased Steel
Steel	PPI3312213	Cold Finished Steel Bars And Bar Shapes
Steel	PPI33151	Ferrous Metal Foundries
Steel	PPI331513	Steel Foundries
Steel	PGALVSHEET	Hot Dipped Galvanized Sheet
Steel	WPIP1015	Ferrous Foundry And Forge Shop Products
Steel	PPI331511	Iron Foundries
Transportation	PPI482111	Railroads - Line Haul Operations
Transportation	PPI483111	Deep Sea Foreign Transport Of Freight
Transportation	PPI4831132	Coastal And Great Lakes Freight Transportation
Transportation	PPI483211	Inland Water Freight Transportation
Transportation	PPI3361101	Passenger Cars And Chassis
Transportation	PPI3361102	Trucks, Truck Tractors, & Truck Chassis, 14,000 Lbs. And Less,
Transportation	PPI336510	Railroad Rolling Stock
Transportation	PPI336412	Aircraft Engines & Engine Parts
Transportation	PPI336413	Other Aircraft Parts & Auxiliary Equipment
Transportation	PPI3366	Ship And Boat Building
Transportation	PPI336611	Ship Building And Repair
Transportation	PPI336110	Automobile & Light Duty Motor Vehicles
Transportation	WPIP141106	Trucks, Over 14,000 Lbs., Gross Vehicle Weight
Transportation	WPIWP14	Transportation Equipment
Wages	CEU0600000008	Goods Producing Industries
Wages	CEU1021100008	Oil & Gas Extraction
Wages	CEU3100000008	Durable Goods Manufacturing
Wages	CEU3132100008	Wood Product Manufacturing
Wages	CEU3132700008	Nonmetallic Mineral Product Manufacturing
Wages	CEU3132730008	Cement & Concrete Product Manufacturing
Wages	CEU3133100008	Primary Metal Manufacturing
Wages	CEU3133200008	Fabricated Metal Product Manufacturing
Wages	CEU3133320008	Industrial Machinery Manufacturing
Wages	CEU3133350008	Metalworking Machinery Manufacturing
Wages	CEU3133360008	Engine, Turbine, & Power Transmission Equipment Manufacturing
Wages	CEU3133400008	Computer & Electronic Product Manufacturing
Wages	CEU3133420008	Communications Equipment Manufacturing
Wages	CEU3133440008	Semiconductor & Other Electronic Component Manufacturing
Wages	CEU3133441908	Other Electronic Component Manufacturing
Wages	CEU3133450008	Navigational, Measuring, Electromedical, & Control Instruments Manufacturing
Wages	CEU3133500008	Electrical Equipment, Appliance, & Component Manufacturing
Wages	CEU3133530008	Electrical Equipment Manufacturing
Wages	CEU3133600008	Transportation Equipment Manufacturing
Wages	CEU3133610008	Motor Vehicle Manufacturing
Wages	CEU3133630008	Motor Vehicle Parts Manufacturing
Wages	CEU3133640008	Aerospace Product & Parts Manufacturing
Wages	CEU3133641108	Aircraft Manufacturing
Wages	CEU3133700008	Furniture & Related Product Manufacturing

Category	Global Insight ID	Global Insight Name
Wages	CEU3133910008	Medical Equipment & Supplies Manufacturing
Wages	CEU3231100008	Food Manufacturing
Wages	CEU3231300008	Textile Mills
Wages	CEU3231500008	Apparel Manufacturing
Wages	CEU3232200008	Paper Manufacturing
Wages	CEU3232300008	Printing & Related Support Activities
Wages	CEU3232500008	Chemical Manufacturing
Wages	CEU3232510008	Basic Chemical Manufacturing
Wages	CEU3232540008	Pharmaceutical & Medicine Manufacturing
Wages	CEU3232560008	Soap, Cleaning Compound, & Toilet Preparation Manufacturing
Wages	CEU3232600008	Plastics & Rubber Products Manufacturing
Wages	CEU3232610008	Plastics Products Manufacturing
Wages	CEU4348400008	Truck Transportation
Wages	CEU4422000008	Utilities
Wages	CEU4422110008	Electric Power Generation, Transmission & Distribution
Wages	CEU5051700008	Telecommunications
Wages	CEU6000000008	Professional & Business Services
Wages	CEU6054000008	Professional, Scientific, & Technical Services
Wages	CEU6054130008	Architectural, Engineering, & Related Services
Wages	CEU6054150008	Computer Systems Design & Related Services
Wages	CEU6054180008	Advertising & Related Services
Wages	CEU6056132008	Temporary Help Services
Wages	CEU6056160008	Investigation & Security Services
Wages	CEU6562000008	Health Care & Social Assistance
Wages	CEU7072200008	Food & Drinking Places
Wages	CEU1000000008	Natural Resources & Mining
Wages	CEU2023800008	Specialty Trade Contractors
Wages	CEU6055000008	Management Of Companies & Enterprises
Wages	ECICCTNS	Compensation, Civilian Workers
Wages	ECIPCPARNS	Compensation, Professional And Related
Wages	ECIPCNRCNS	Compensation, Nursing And Residential Care
Wages	ECIPCTNS	Compensation, Total Private Industry
Wages	ECICCHCSNS	Compensation, Civilian, Health Services
Wages	ECICWTNS	Wages And Salaries, Civilian Workers
Wages	ECICBTNS	Benefits, Civilian Workers
Wages	ECIPWSONS	Wages And Salaries, Service Occupations
Wages	ECIPCOASNS	Compensation, Office And Administrative Support
Wages	ECIPWOASNS	Wages And Salaries, Office And Administrative Support
Wages	ECIPCMBFNS	Compensation, Management Business And Financial
Wages	ECIPWMBFNS	Wages And Salaries, Management Business And Financial
Wages	ECIPBMBFNS	Benefits, Management Business And Financial
Wages	ECIPWPARNS	Wages And Salaries, Professional And Related
Wages	ECIPCSARNS	Compensation, Sales And Related
Wages	ECIPBSARNS	Benefits, Sales And Related
Wages	ECIPCCEONS	Compensation, Construction And Extraction
Wages	ECIPWCEONS	Wages And Salaries, Construction And Extraction
Wages	ECIPBCEONS	Benefits, Construction And Extraction
Wages	ECIPWPRODNS	Wages And Salaries, Production
Wages	ECIPCTMMNS	Compensation, Transportation And Material Moving
Wages	ECIPWTMMNS	Wages And Salaries, Transportation And Material Moving
Wages	ECIPWFINNS	Wages And Salaries, Private, Financial Services

Category	Global Insight ID	Global Insight Name
Wages	ECIPWUMNS	Wages And Salaries, Union Workers, Manufacturing
Wages	ECIPWAIRNS	Wages And Salaries, Aircraft Manufacturing
Wages	ECICWHCSNS	Wages And Salaries, Civilian, Health Services
Wages	ECICCHOSNS	Compensation, Civilian, Hospitals
Wages	ECICWHOSNS	Wages And Salaries, Civilian, Hospitals

10. Appendix C – Additional ARIMA Model Visuals

For the Moving Average Model, three of the four tests *suggest non-stationarity*, as demonstrated in **Error! Reference source not found.** and Figure 13. The variance was visually inconsistent (*suggesting non-stationarity*), the variance does not halve (*suggesting non-stationarity*), the ACF test does quickly decay to zero (*suggesting stationarity*), and the p-value is 0.5129 (*suggesting non-stationary*).

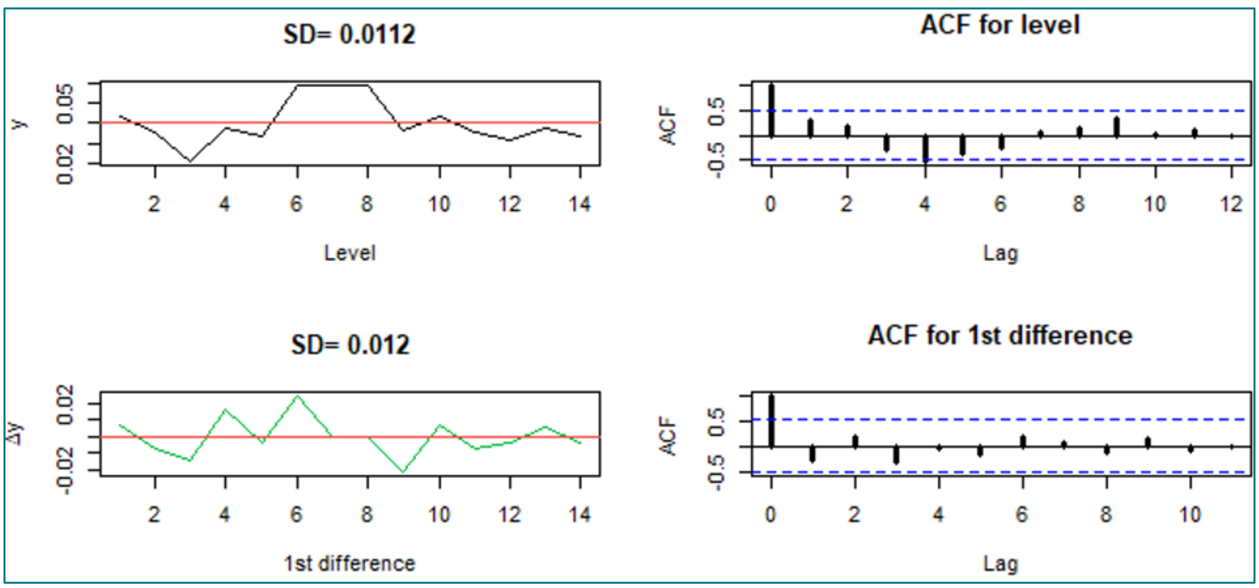

Figure 12 - Model ARIMA Variance

```
        Augmented Dickey-Fuller Test


Dickey-Fuller = -2.1562, Lag order = 2, p-value = 0.5129
alternative hypothesis: stationary
```

Figure 13 - Model ARIMA ADF Test

Since the Moving Average Model was found to be non-stationary, a transformation was attempted using the Box Cox test (see Figure 14), which reduces the variability in the dataset. In this case, the data only contained a single lag, nullifying the PCAF and ACF tests; for this reason, the Box Cox transformation was discarded, and the initial dataset was used (however this technique has potential uses on other indices).

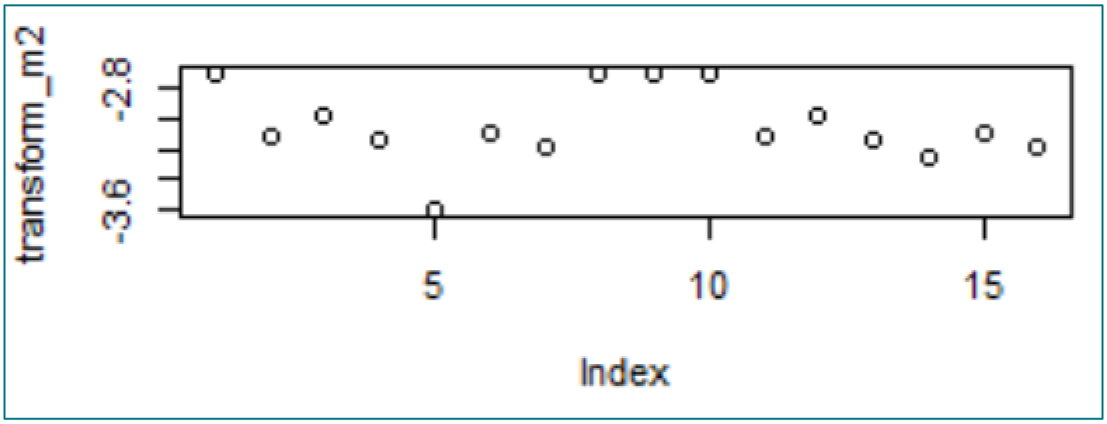

Figure 14 - Model ARIMA Box Cox

Finally, we used the auto-Arima package, which yields a (0,0,0) ARIMA (see Figure 15); this indicates that the model is a very good fit with standard error of 0.0028. Therefore, the original Moving Average Model is the best fit, so it should be used in lieu of an ARIMA model.

```
ARIMA(0,0,0) with non-zero mean

Coefficients:
        mean
      0.0412
s.e.  0.0028

sigma^2 = 0.0001309:  log likelihood = 49.34
AIC=-94.69   AICc=-93.76   BIC=-93.14
```

Figure 15 - Model ARIMA Output