Re-Tooling the Estimator's Approach to Escalation Forecasting

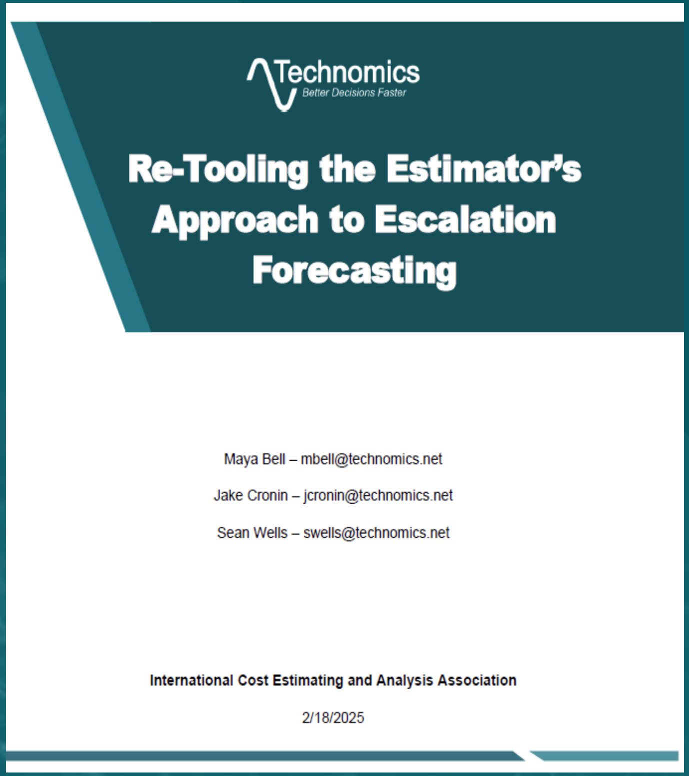
Technomics
Better Decisions Faster

International Cost Estimating and Analysis Association (ICEAA)

2025

Agenda

Check out our
Long-Form Research Paper!


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

- Introduction
- Data and ARIMA Background
- Forecast Models
- Analysis and Forecast Results
- Literature Review and Uncertainty Process
- Recommendations & Next Steps

The Team

Sean Wells

Lead Analyst

Sean has been with Technomics since July 2023 and currently provides cost support to Navy Integrated Warfare Systems 6.0 and the Canadian National Shipbuilding Service. Sean has an MS in Contracts Acquisition and Management from the Florida Institute of Technology and has 5+ years of professional experience working Ford Class Carrier and Virginia Class Submarine estimates. He is an ICEAA Certified Cost Estimator/Analyst (CCEA).

Maya Bell

Senior Associate

Maya Bell is a Senior Associate who joined Technomics in November 2023 and provides cost analysis/estimation, data analysis, data wrangling/visualization, and statistical analysis to the Conventional Prompt Strike Program. She received her BS in Economics from the Catholic University of America in 2019 and her MS in Applied Economics from the George Washington University in 2022.

Jake Cronin

Associate

Jake has been with Technomics since September 2023 and currently provides cost and acquisition support to Navy Integrated Warfare Systems 4.0 and 11.0 Jake has a B.S. in Applied Mathematics from Penn State University and has experience in historical data analysis, cost modeling, sole source proposal analysis and modeling, and earned value management.

Introduction

Escalation Forecasting and Cost Estimating

- Supply Chain Shocks
- Trade Wars
- Increasing Sustainment Costs
- Tariffs
- Economic Uncertainty

2022 Secretary of Defense: "Focus on Economic Price Adjustments, Long-Range budget estimates, and high-quality cost models to support them"

How will the Cost Community Respond?

Introduction

Thesis – *"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."*

Our Approach

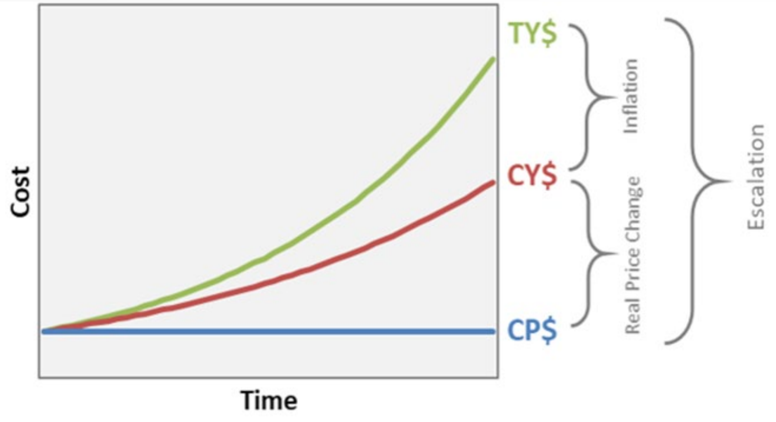
- BLS Moving Average Prediction / Global Insight Prediction (2024 Paper)
- CAGR Comparison
- ARIMA Check for Model Fit
- Risk Analysis

Background – Definitions/Key Terms

Inflation vs. Escalation

- **Inflation:** Economy-wide increase in prices (affects all prices similarly)
- **Escalation:** Price changes incorporating combination of factors
- **Real Price Change:** Specific price change of a good/service different from economy-wide trends

Per OSD CAPE Manual



Compound Annual Growth Rate (CAGR)

- Measure of average escalation rate over time
- Normalizes data for different index base years

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

where:

EV = Ending value

BV = Beginning value

n = Number of years

Root Mean Squared Error (RMSE)

- Average difference between a model's predicted values and actual values; smaller RMSE is better

Cost Example

Navigational Equipment at $1M Unit price in 2011 – how do you escalate it?

Inaccurate escalation forecasts **add up** across contracts, **compound** over time, and **erode** cost savings

Background – Data Sources

1. Chemicals
2. Construction
3. Defense and Aerospace
4. Electronics
5. Energy
6. Macro Price
7. Maintenance, Repair, and Operations
8. Metals
9. Non-Electric Machinery
10. Steel
11. Transportation
12. Wages

Bureau of Labor Statistics (BLS)

- Actual measured escalation values
- "Source of Ground Truth"

Global Insight (GI)

- Industry Forecast Data
- CAPE cites as a common data source for DoD

Table 1 - Sample Index Cross-Reference

Category	GI ID	GI Title	BLS ID	BLS Title
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

Power Query

BLS Moving Average Model

ID #

- ~200 Unique Indices

Start Year

- Earliest Year = 2008
- Latest Year = 2014

Data Years

Prediction Years

- 1, 2, or 3 years of Moving Average Data
- 1-10 Prediction Years

Table 2 - Sample Forecast Data

BLS ID	Start Year	Data Years	Predict Years	BLS CAGR	Model CAGR	GI CAGR
PCU322211322110	2010	1	1	4.04%	5.30%	6.36%
PCU322211322110	2010	1	2	2.10%	5.30%	3.82%
PCU322211322110	2010	1	5	3.02%	5.30%	1.81%
PCU322211322110	2010	1	10	2.58%	5.30%	1.78%
PCU322211322110	2010	2	1	4.04%	3.31%	6.36%
PCU322211322110	2010	2	2	2.10%	3.31%	3.82%
PCU322211322110	2010	2	5	3.02%	3.31%	1.81%
PCU322211322110	2010	2	10	2.58%	3.31%	1.78%
PCU322211322110	2010	3	1	4.04%	4.31%	6.36%
PCU322211322110	2010	3	2	2.10%	4.31%	3.82%
PCU322211322110	2010	3	5	3.02%	4.31%	1.81%
PCU322211322110	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%

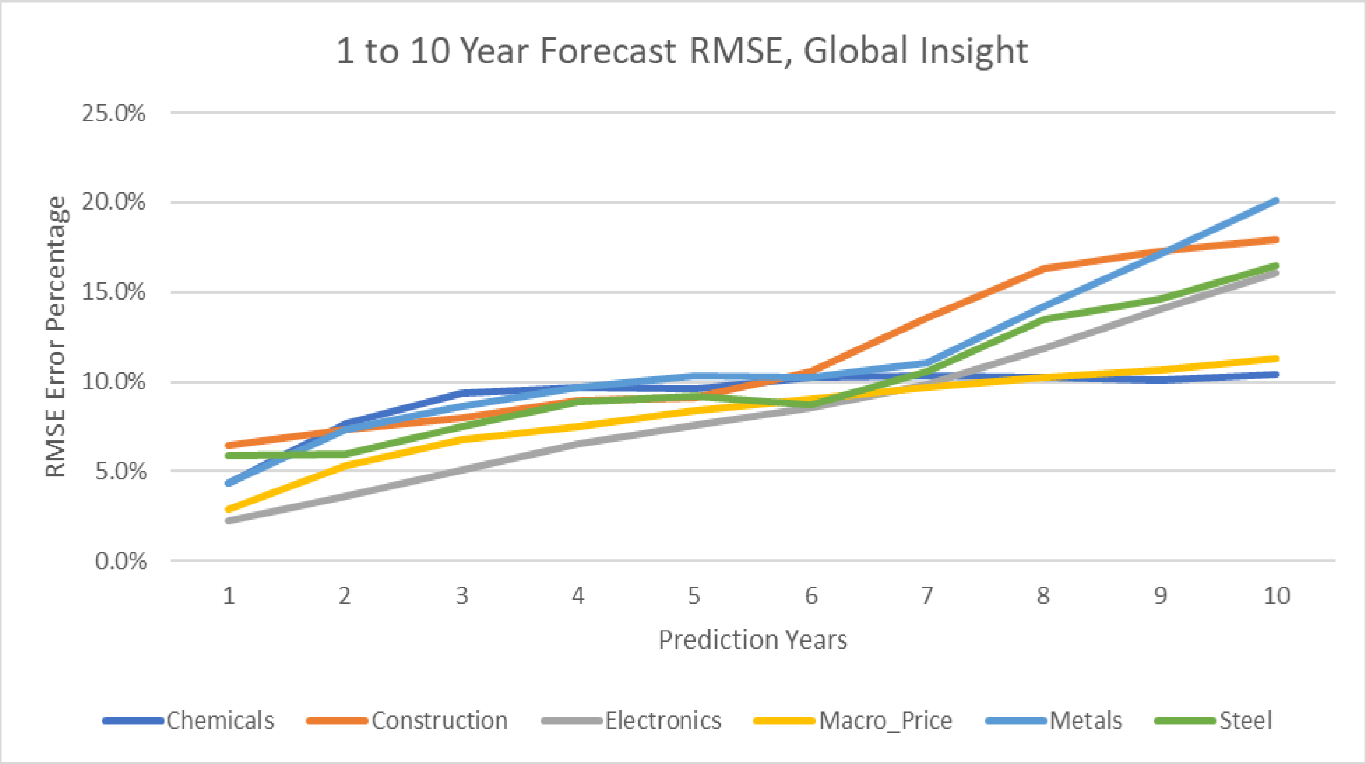
For sake of space, showing 1-2-5-10 years vice all years

Sector Performance (1)

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%

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% ★

Sector Performance (2)



- Global Insight outperformed the Moving Average Model by 2-3x for these 6 index categories
- Energy was the exception
  - Global Insights at **30-35%** RMSE at the 5-to-10-year mark
  - Moving Average at **40-70%** RMSE for 5–10-year range

Historical Variability in Forecasts is Measurable!

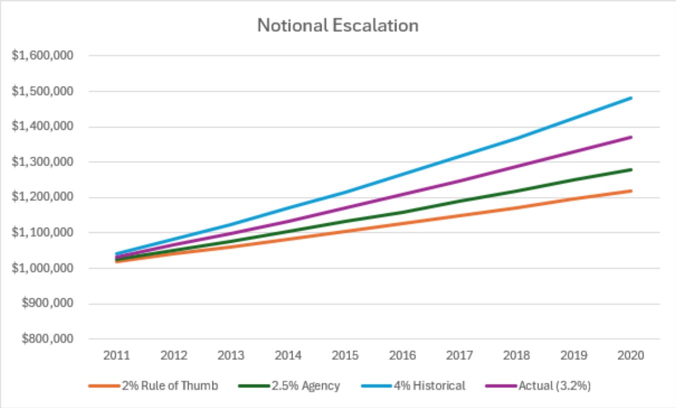
Background – ARIMA Overview

- The ARIMA (Auto-Regressive Integrated Moving Average) 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

The goal of the ARIMA model is to select the **best fitting model** to use for forecasting **time series** data.

The best fitting model will have the **lowest AIC**, **stationary data**, and data that will pass a **confidence interval** check at 95%.

The benefit of using an ARIMA model for escalation data is primarily focused on the fact that escalation data is **time series** data, where data trends can be cyclical, seasonal, or random



Background – ARIMA Overview (Continued)

- **Homoscedasticy** - variables that have constant variances
- **Stationarity** - do not contain a trend and has constant variance
- **Lags** - the time difference between two observations in a time series or the delay
  - They also help inform whether the ARIMA model will be comprised of an autoregressive (AR) component or a moving average (MA) component
- ARIMA model output expression (p,d,q):

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)
The parameter d	• Indicates the degree of differencing (referring to the differences between consecutive observations) required to achieve stationarity
The parameter q	• Is the size of the moving average window

Background – ARIMA Overview (Continued)

- ARIMA model unique test plots and transformations:

The Augmented Dickey-Fuller (ADF) plot *helps determine the d parameter by the type of trend decay and the confidence intervals*

The Autocorrelation Function (ACF) plot *helps identify the q parameter by showing the correlation between observations at different lags*

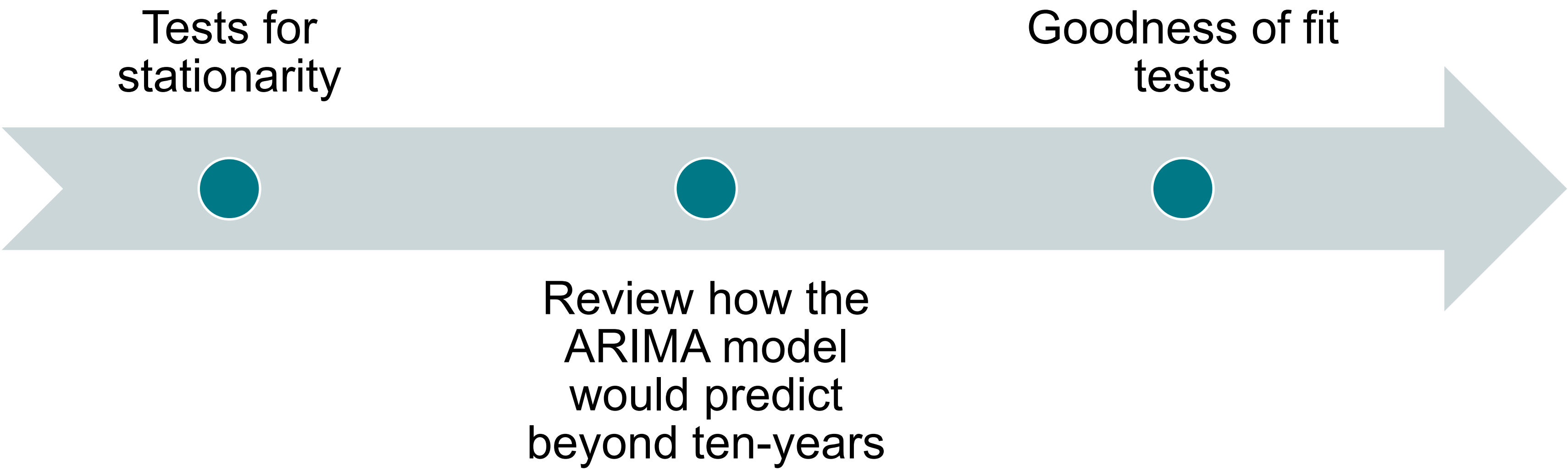
The Partial Autocorrelation Function (PACF) plot *determines the p parameter by showing the correlation of an observation with its lagged values after removing the effects of intermediate lags*

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*

The Box Cox transformation *is a power transformation that is used to adapt non-stationary data into stationary data*

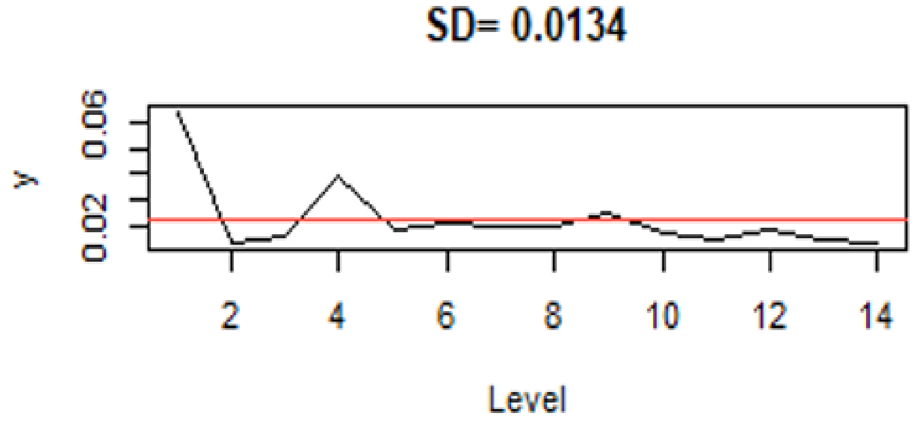
ARIMA Deep Dive: Introduction

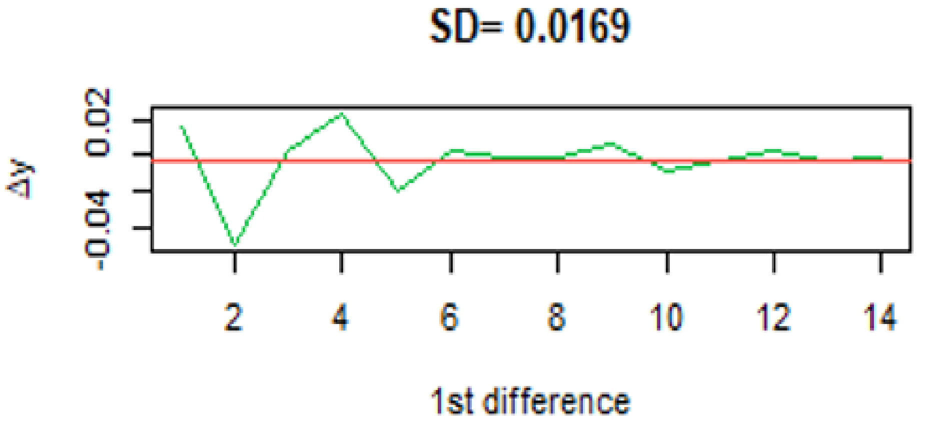
- **Index analyzed:** Aeronautical, Nautical, Navigational Instruments (PPI3345111)
- **Data Sets used:** Global Insights and Moving Average Model
- **ARIMA modeling steps:**



ARIMA Deep Dive: Testing for stationarity

- **The GI forecast is found to be stationary in three of the four initial ARIMA variance tests:**
  1. The variance is **visually constant** (no increasing trend or true random variation), **suggesting stationarity**
  2. The variance **does not halve over time**, when comparing the initial variance level to the 1st difference (log) variance, standard deviation (SD) equals 0.0134 and increases to 0.0169, **suggesting non-stationarity**



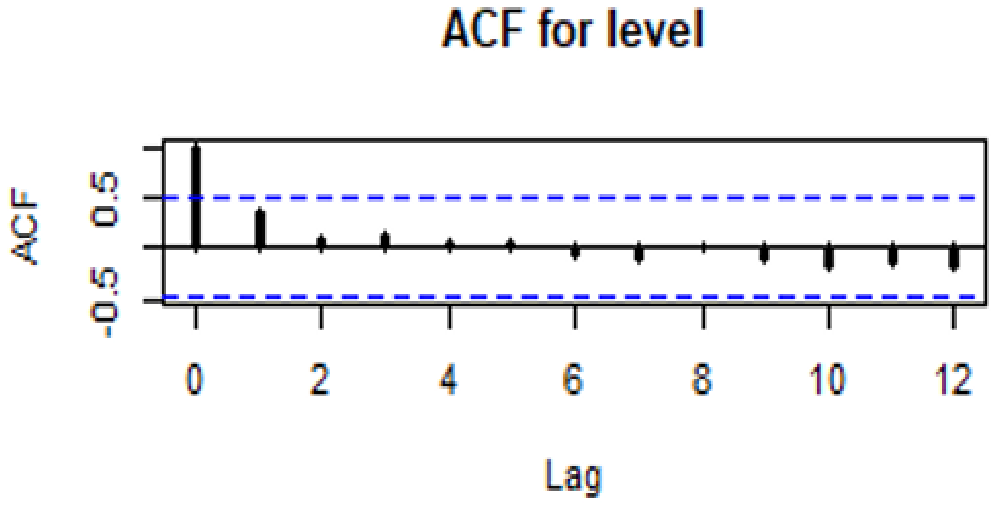


Technomics

ARIMA Deep Dive: Testing for stationarity (Continued)

- **The GI forecast is found to be stationary in three of the four initial ARIMA variance tests:**
  3. The **ACF decays to zero quickly**, **suggesting stationarity.**
  4. The ADF test results in a **p-value of 0.01**, **suggesting stationarity.**

Having a stationary model enables consistent goodness of fit testing and avoids seasonality



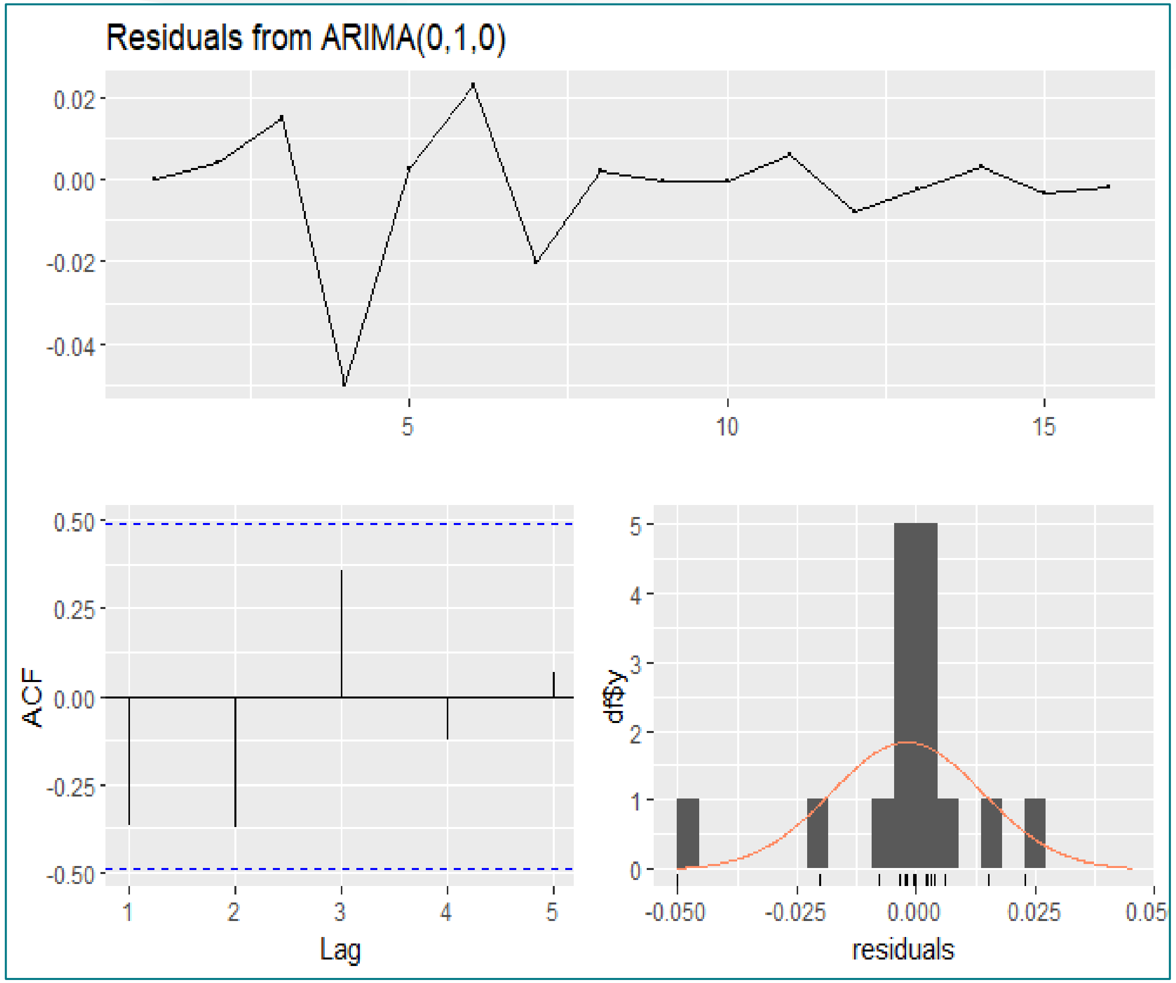

```
Augmented Dickey-Fuller Test

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

ARIMA Deep Dive: Goodness of Fit test

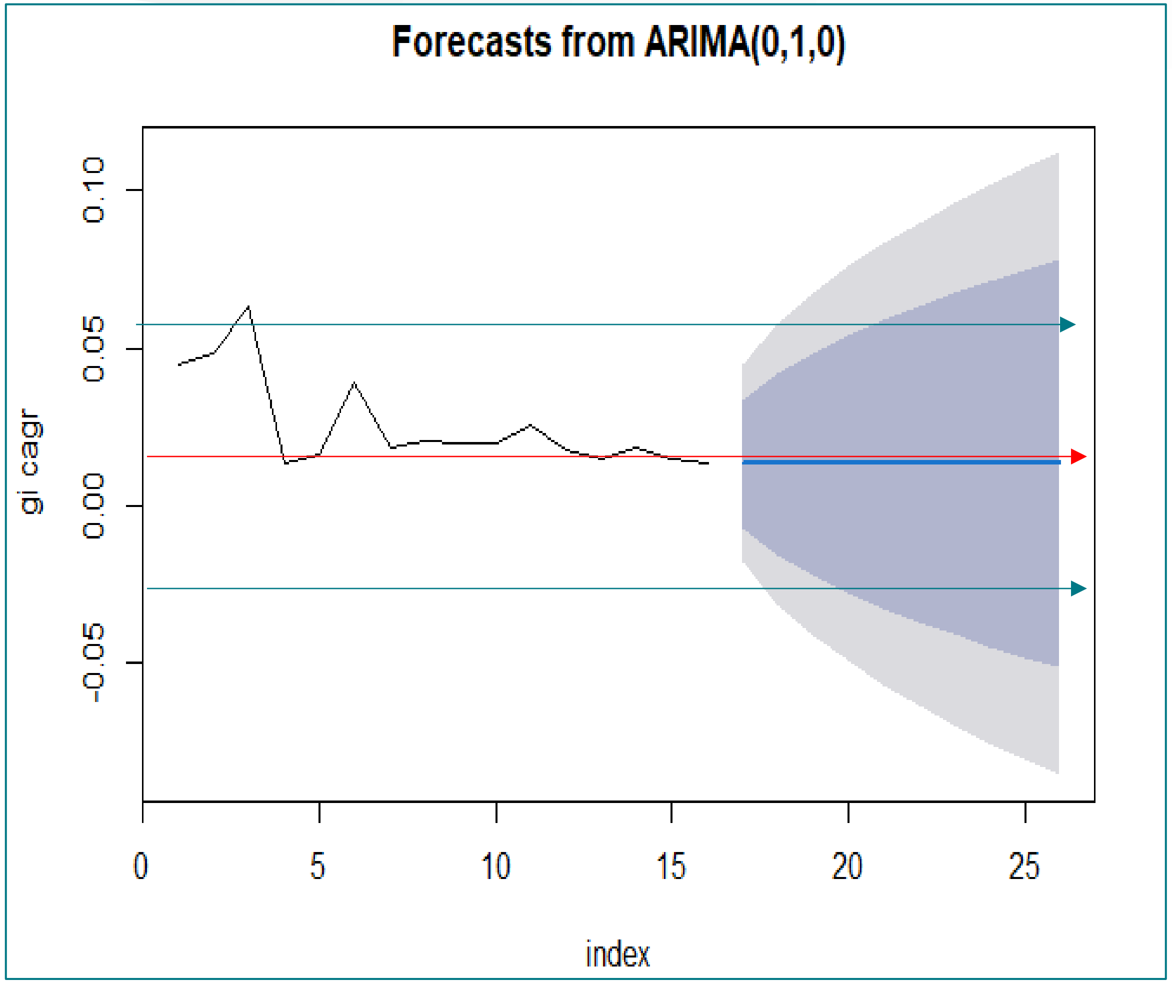
- **Resulting ARIMA Model:**
  - model of (0,1,0)
  - zero lags within the ACF test
  - autoregressive (AR) 1 model
  - zero lags of the PACF test
  - AIC of -79.58 which means the residuals are generally grouped together

Tight Spread of Residuals = Good Model Fit = Strong Predictions



ARIMA Deep Dive: Forecasting

- Forecasting ARIMA results:
  - ARIMA model forecast are aligned with the GI data base on the blue cone seen to the right
  - Cone of uncertainty helps with establishing both visual and calculated risk ranges
- Also tested Moving Average model:
  - Nonstationary on 3 tests
  - (0,0,0) model with 0.0028 standard error
  - For this case, Moving Average was the best fit



Modeling Limitations

- Methodology Limitations:
  - Limited years of data (available Global Insights files, some BLS indices not starting before the 2000s)
  - Used a simple moving average to demonstrate goodness of fit / risk range points; more complex models (like exponential weighted) might improve accuracy
- ARIMA Limitations:
  - Dataset needs to be non-correlated, sometimes an issue with escalation
  - ARIMA works best with large datasets (long forecast spans)

Literature Review - Overview

DeCarlo, Jabaley, & Druker 2012 (BAH)

- Monte-Carlo based method for prediction error assessment
- Historical standard deviation from actual inflation to calculate deciles around future estimates

CAPE Inflation Handbook, 2011

- Highlights risk of underestimating inflation
- Calculate deviation between actual/budget rate, then placed on probability distribution at P80

Joukar and Nahmens, 2016 (LSU)

- Leverage Value at Risk (VaR) technique to manage construction escalation
- Leverage ARIMA modeling to calculate VaR at 90-95% probability levels

Touran and Lopez, 2006 / 2012

- Highlights implicit link between schedule delays and increased impact of escalation
- Calculate moving average with additional weighting to last known index value (Exponentially Weighted Moving Average)

Literature Review – Key Trends

1. **Consistent** acknowledgement that estimators should specifically model escalation risk when conducting Monte Carlo trials
2. **Consistent** use of historical index data to derive distributions around published forecasts
3. **Consistent** sensitivity assessment of interaction between escalation, schedule, and material availability

Uncertainty Process

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 trial
4. Compare Trials With and Without Escalation Variance

1: Calculation of Historical Uncertainty

2A: Internally developed escalation factors

- Loop Historical data (5-10 years) through forecasting model
- Compare results to actual BLS values
- Output percentage deltas

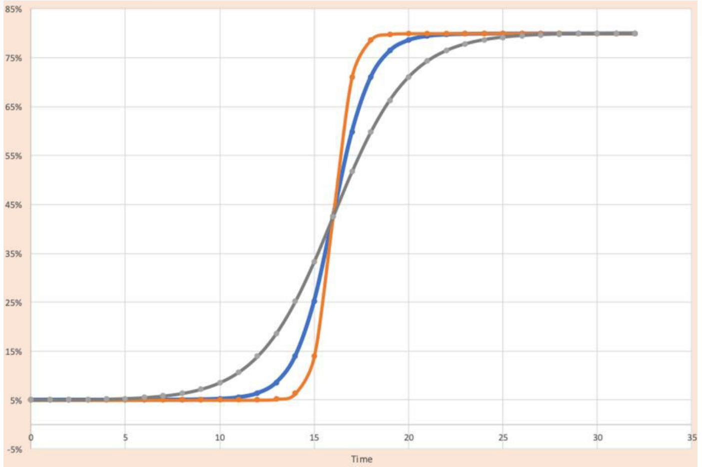
2B: Outsourced forecasts

- Retrieve past forecasts
- Align with actual BLS values
- Compare deviations
- Optionally incorporate moving average as sanity check

3: Consolidate findings to assess forecast accuracy and risk factors

Uncertainty Process

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 trial**
4. **Compare Trials With and Without Escalation Variance**



- Then-Year to Constant Year Conversion: Use a risk % or leverage decile sensitivity runs for escalation
- Assess Real Price Change while keeping baseline inflation constant
- Sensitivity Analysis between Schedule Risk Analysis / Discrete Risk Register and the Escalation Parameter

Enables traditional Monte Carlo analysis and/or escalation-specific sensitivity analysis in cost models

Conclusions and Recommendations

- Escalation forecasting and realistic modeling is as salient as ever!
- Defensible, Repeatable Risk Analysis Framework
  - Calculate the Historical Uncertainty in the Escalation Forecast
  - Develop an Uncertainty Distribution for that Forecast
  - Run Sensitivity and Risk Analysis using Monte Carlo trial
  - Compare Trials With and Without Escalation Variance
- Using ARIMA models can improve fidelity of custom escalation models and provide goodness of fit
- Improved understanding of the drivers of escalation is critical to accurate estimation!

Technomics
Better Decisions Faster

Questions?

swells@technomics.net
mbell@technomics.net
jcronin@technomics.net