



**RATE SETTING ANALYSIS: A STATISTICAL APPROACH TO OUTLIERS IN
THE RATE SETTING PROCESS WITHIN THE UNITED STATES
TRANSPORTATION COMMAND**

THESIS

Warren B. McGriff, 2nd Lieutenant, USAF

AFIT-ENS-MS-16-M-118

**DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY**

AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

DISTRIBUTION STATEMENT A.
APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

The views expressed in this thesis are those of the author and do not reflect the official policy or position of the United States Air Force, Department of Defense, or the United States Government. This material is declared a work of the U.S. Government and is not subject to copyright protection in the United States.

AFIT-ENS-MS-16-M-118

**RATE SETTING ANALYSIS: A STATISTICAL APPROACH TO OUTLIERS IN
THE RATE SETTING PROCESS WITHIN THE UNITED STATES
TRANSPORTATION COMMAND**

THESIS

Presented to the Faculty

Department of Operational Sciences

Graduate School of Engineering and Management

Air Force Institute of Technology

Air University

Air Education and Training Command

In Partial Fulfillment of the Requirements for the
Degree of Master of Science in Operations Research

Warren B. McGriff, BS

Second Lieutenant, USAF

March 2016

DISTRIBUTION STATEMENT A.
APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

**RATE SETTING ANALYSIS: A STATISTICAL APPROACH TO
OUTLIER ANALYSIS IN THE RATE SETTING PROCESS WITHIN THE
UNITED STATES TRANSPORTATION COMMAND**

Warren B. McGriff, BS

Second Lieutenant, USAF

Committee Membership:

LTC Brian J. Lunday, PhD
Chair

Dr. Jeffery D. Weir
Reader

Abstract

The rate setting process implemented by the USTRANSCOM J8 Program Analysis & Financial Management Directorate (USTC J8 or TCJ8) is of importance, as accurate rates yield proper and equitable recoupment of costs from customers and ensure USTC can achieve an accurate Net Operating Result (NOR) in each given year. This research sought to identify areas in which the current rate setting methodology can be improved. We initially examined the use of six months of historical cost data versus a full year of data to set rates, concluding that there is not a statistically significant difference with respect to their relative effect on the NOR; USTC should proceed with their current practice.

The research also identified outliers, first with regard to likelihood of historical rates not being set by the prescribed process and second with regard to whether the rates set by the prescribed process would be an outlier in terms of the marginal contribution to the net operating result. We found that approximately 8%, 10%, and 4% of the rates in FY14–FY16 were likely set using budget analyst experience in lieu of the prescribed method, for the most part imposing a reduction in the prescribed rates. Adapting classical Statistical Process Control (SPC) methods, we found that the prescribed rate setting method does work in aggregate but can induce recurrent outlier rates. However, a pattern in these outlier rates remains elusive – some are self-correcting – but the demonstrated methodology is shown to be useful for identifying outlier rates that do merit budget analyst experience-informed judgment for rate setting.

The final component of this research examined the combination of two factors used in the current methodology to adjust current average weighted costs to set future rates: the Accumulated Operating Result and Composite Rate Adjustment factors. Using historical data from FY08–FY15, we calculate the optimal combined factor values for each respective fiscal year to achieve an NOR equal to \$0. In doing so, we concluded that the combination of these two factors contributed to approximately 25% of the induced error in NOR. We suggest a more detailed examination of these rate computations for additional analysis.

Acknowledgments

I would like to express my gratitude to my faculty advisor, LTC Brian Lunday, for his leadership and dedication throughout the course of this thesis effort. The insight and knowledge shared over the course of this work is greatly appreciated.

Warren B. McGriff

Table of Contents

	Page
Abstract	iv
Table of Contents	vii
List of Figures	ix
List of Tables	x
I. Introduction	1
1.1 Background	1
1.2 Problem Statement	2
1.3 Research Objectives	3
1.4 Methodology Overview	4
1.5 Overview	5
II. Literature Review	6
2.1 Chapter Overview	6
2.2 The History of Transportation Cost	6
2.3 Demand Forecasting for Transportation	8
2.4 Rate Setting	12
2.5 Statistical Process Control	15
2.6 Summary	17
III. Methodology	19
3.1 Introduction	19
3.2 Current Practice	23
3.3 Evaluating Current Practice	25
3.4 Improving Upon Current Practice	27
3.5 Summary	36

IV. Analysis and Results	37
4.1 Chapter Overview	37
4.2 Evaluating Current Practice	37
4.3 Rates	38
4.4 Predicted Rate Contributions to the NOR	44
4.5 Summary	62
V. Conclusions and Recommendations	64
5.1 Overview	64
5.2 Conclusions	64
5.2 Research Significance	67
5.3 Recommendations for Action/Future Research	68
5.4 Summary	69
Appendix A: Script to compute the Weighted Average Costs via Equation (1)	70
Appendix B: A Script to Compute NORs via Equation (3)	73
Appendix C: A Script to Compute MNORs via Equation (7)	75
Appendix D: A Script to Compute the Composite AOR Factor via Equation (6)	77
Appendix E: Location Table of the Set I and J	79
Appendix F: QuadChart	81
Bibliography	82

List of Figures

Figure 1: Control Chart Example.....	32
Figure 2: Normal Probability Density Function	33
Figure 3: Individual Outlier Behavior Example	34
Figure 4: Behavior Over Time.....	46
Figure 5: Location Outlier Frequencies	51
Figure 6: Shipping Code Outlier Frequencies	54
Figure 7: 01 13 06, 01 13 11, 01 22 13 Outlier Behavior	57
Figure 8: 01 23 04, 01 43, 08, 01 46 08 Outlier Behavior	58
Figure 9: 02 15 06, 02 17, 16, 02 21 13 Outlier Behavior	59
Figure 10: 02 22 13, 03 23 12, 03 39 08 Outlier Behavior	59
Figure 11: 17 23 06, 22 23 08 Behavior Outlier	60

List of Tables

Table 1: Shipping Code Descriptions	20
Table 2: 6 Mo. vs. 1 Yr. Absolute Value Hypothesis Test	38
Table 3: Robust Examination of Rate Outliers	40
Table 4: Rate Outliers from 2014 Rates	41
Table 5: Rate Outliers from 2015 Rates	42
Table 6: Rate Outliers from 2016 Rates	43
Table 7: Mean of Means Results	45
Table 8: Data Falling 3 and 2 Standard Deviations Away from the Mean Hypothesis Test Results	48
Table 9: One-time outliers	50
Table 10: Partially Matched Outliers	50
Table 11: Location Code Frequency Breakdown	53
Table 12: Shipping Code Frequency Breakdown	55
Table 13: Recurrent Outliers over Fiscal Years (2008-2014).....	56
Table 14: Composite Method Hypothesis Test.....	62

RATE SETTING ANALYSIS: A STATISTICAL APPROACH TO OUTLIERS IN THE RATE SETTING PROCESS WITHIN THE UNITED STATES TRANSPORTATION COMMAND

I. Introduction

1.1 Background

The United States Transportation Command (USTC) J8 Staff Directorate is responsible for setting transportation rates for their customers within the Department of Defense (DoD) who ship who ship cargoes by ground and/or sea. The goal of the rate setting process is for USTC to have an annual Net Operating Result (NOR) of \$0. The NOR is defined by J8 as the sum of the differences between the rate charged and the actual cost of each individual shipments over all shipments in a fiscal year. Currently, rates for an upcoming fiscal year, which we will denote as “FY+1”, are generated for aggregated origin-destination pairs using the first six months of cost data from the current calendar year, which we denote as “FY0”. Within this approach, USTC J8 financial analysts have conjectured that the limited nature of this cost data yields inaccuracies in rate setting that, in aggregate, induce a NOR not equal to \$0.

A rate is generated for each aggregated origin-destination/destination-origin and commodity code combination in fiscal year FY+1 by first calculating the average per-measurement-ton cost in year FY0 using the first six months of accumulated cost data. The first six months of data of FY0 is used as it is all that is available at the time the rates must be turned in, as Congress mandates that all rates be finalized before a budget is set for year FY+1. The average cost is then multiplied by a Refresh Rate, as provided by

USTC's Finance Department, in order to adjust FY0's average cost, thereby accounting for inflation, to match the predicted fiscal climate. This average cost is then multiplied by an Accumulated Operating Result (AOR) factor, an adjustment intended to offset the NOR of the previous two years FY-1 and FY-2. To better understand the use of the AOR Factor, it is used to prevent a long-run NOR greater than or less than \$0 to occur. The implementation of a two-year running NOR serves to better align actual rates with customers' expectations and aid in their forecasting of transportation budgets (i.e., to prevent rates from having drastic swings). With regard to application, an AOR factor has a negative value when USTC's running two-year NOR is positive, and the AOR factor is positive when USTC's running two-year NOR is negative (i.e., AOR factor has an inverse relationship with the running NOR). Finally, the average cost is then multiplied by the Composite Rate Adjustment factor, which is used in order to balance the rates set back to the previous budget cycle.

The focus of this research will be on examining and improving upon inaccuracies induced by using sparse-cost data in a given calendar year (FY0) to set Transportation Working Capital Fund (TWCF) rates for the upcoming fiscal year (FY+1). Further the research will focus on the effect of using a full-year of data compared to that of six months, and it will identify outliers in the rates which require special attention in order to set accurate rate estimates.

1.2 Problem Statement

This research seeks to characterize and quantify the suspected sources of error in the existing TWCF rate setting process, and to develop, test, and recommend a method

(or methods) to improve their collective accuracy, as measured by the goal of obtaining a NOR of 'zero' in the following fiscal year.

1.3 Research Objectives

The objectives of this research are as follows:

- Objective #1. Use historical cost data sets and rate setting factors from FY08-FY-15 to analyze the impact of the current practice of using only the first six months of cost data, as compared to a proposed ideal of using all twelve months of cost data, from the current fiscal year (FY0) to develop TWCF rates for the upcoming fiscal year (FY+1) on the accuracy of achieving the goal of an NOR equal to zero.
- Objective #2. Examine the impact of other sources of error on TWCF rate setting accuracy. These include the Refresh Rate, the AOR factor, and the Composite Rate Adjustment.
- Objective #3. Examine the impact of outliers within the data, and determine what percentage of the data requires manipulation beyond the current rate setting methodology. These include origin-destination/destination-origin combinations that have large NORs as well as those rates emulated to do not align with rates set.
- Objective #4. Automate the methodology(ies) in a prototype student-level-of-fidelity tool used within the scope of this study that, with contractor support, can be further developed into a computer-based tool for customer use.

1.4 Methodology Overview

The objective of the research is to reduce the disparity between the rate set and actual cost of transportation in order to assist USTC in reaching their desired goal of an NOR of \$0. Three different methodologies were implemented in order to analyze the data provided by USTC J8 summarized below. Section 3 provides additional information regarding the methods provided here.

- Method 1: Evaluate rates set using six months vs. a full year of cost data by comparing resulting NORs. This method is utilized to meet Objective #1 of this research, and additional information can be found in Section 3.3.
- Method 2: Calculate the desired composite AOR factor and Composite Rate Adjustment for each year based off perfect information. This method is utilized to meet Objective #2, and additional information can be found in Section 3.4.
- Method 3: Evaluate outliers in the rates in order to determine origin-destination/destination-origin combinations in which require special attention due to large NORs or difference in magnitude between the rate set and the rate emulated. The method is utilized to meet Objective #3, and additional information can be found in Section 3.4.

All of the methods were implemented and carried out through the use of Microsoft's Visual Basic for Application tool in Excel® providing a prototype to meet Objective #4. Additional information on the prototypes can be found in Appendix (A).

1.5 Overview

The remainder of this thesis is organized as follows. Chapter 2 reviews the published literature relevant to this research among four categories: an overview, the history of transportation, demand forecasting in transportation, and rate setting. Chapter 3 details the methodologies utilized within this study, and Chapter 4 presents the results of computational experiments. Chapter 5 discusses major conclusions and recommendations to extend this work's contributions.

II. Literature Review

2.1 Chapter Overview

This review of existing literature seeks to gain insights into the history and techniques used in the rate setting process within the transportation industry to improve upon an existing rate setting methodology used by the United States Transportation Command (USTC). This review includes several works pertaining to rate setting in the transportation industry, to include the effects of technology and globalization, forecasting of demand, rate setting in the liner shipping industry and the implementation of Statistical Process Control.

2.2 The History of Transportation Cost

Since the dawn of the Industrial Revolution, competition has led to a technology driven rampage in the global trade realm. This has been caused by the rise in international trade which has, in part, been driven by the decline in international transportation costs (Hummels, 2007). Despite strong evidence linking the rise in globalized trade to reduced shipping costs, understanding the modern changes in transportation cost is a complex endeavor (Hummels, 2007). Variables such as the type of goods traded, the rate at which they are traded, and the method of transportation all have an important roles in determining the cost of such shipments.

As of 2007, roughly 23% of all world trade occurred between countries who share a land border (Hummels, 2007), sparking a change within the culture trucking industry seeking to provide more efficient lower cost methods of transportation. Research was conducted in an effort to improve fuel mileage. Increased number of trucks on public

highways resulted in damages and safety protocols causing cost increases of their own. Despite the majority of the world observing increased trucking flows, areas such as Africa, the Middle East, and Asia encountered between roughly one and five percent of their respective trade flow across their borders (Hummels, 2007). Addressing these numbers to their individual geographic regions sheds light into a few areas of potential interest for USTC. Asia's low percentage of bordering trade, trade flowing across one's borders via trucking (i.e., non-oceanic pathway trade), is expected as Asia is known for its cheap labor and large population living in coastal regions, indicating reduced ocean transport cost into Asia as well as increased transportation cost into remote locations within Asia. Africa and the Middle East's lack of trade across borders indicate a sign of lack of infrastructure and stability in the respective regions. With recent conflicts in both regions, this could warn of potential costs that need to be addressed by USTC, as conditions will force higher-cost methods of transportation to be implemented.

Asia's low percentage of trade across their borders coupled with the fact that it holds the largest exporting country in the world (i.e., China) introduces a growing realm of trade via bulk cargoes. Labor laws and unions in many countries have pushed the production of goods to other countries that offer a cheap source of labor due to having less restrictive rules and regulations. This has led to a large percentage of trade coming in the form of bulk cargoes, as manufactured goods are being shipped across ocean pathways.

USTC's mission is to provide air, land and sea transportation for the Department of Defense, both in times of peace and in war (Command, 2005). The DoD is not seeking to transport large quantities of raw materials; the DoD wishes to ship manufactured goods

quickly and efficiently. Despite air transportation offering timely shipping, the technology is not yet present to ship large quantities in a cost efficient manner. This is represented by air transportation comprising less than 1% of the transportation in the trade market (Hummels, 2007). These facts lead to an understanding of the importance of ocean pathway transportation in driving down shipping cost for USTC, and highlight the need for potential analysis on cost versus speed when choosing the method of transportation to implement.

This knowledge will lead to a better understanding of the aspects taken into consideration by USTC when setting rates. Further, being able to understand the types of transportation being discussed will lead to a better problem understanding and ultimately a better product for the customer.

2.3 Demand Forecasting for Transportation

I. Application

Demand forecasting is used in many different fields of business. Businesses that operate on product sales or services often use a demand forecast in order for the business to be prepared to meet a desired service level. Within the transportation industry, demand helps predicts cost. If a transportation company knows their respective demand to a degree of certainty, they also know a set of fixed costs to the same degree of certainty. For example, if a transportation repair shop knew they would have four trucks in the shop per week on average for repairs, management would not staff the shop to a level capable of repairing ten trucks per week. Instead they would staff the facility to repair at least four trucks, possibly five or six depending on the service level desired. This train of

thought applies to the transportation industry; if demand is known to a degree of certainty, then certain expenses that need to be covered are predictable.

II. Methodology

Various techniques in forecasting exist which are characterized as belonging to one of three different categories: qualitative techniques, casual models, or time series analysis (Chamber, Mullick, & Smith, 1971). Qualitative techniques focus more on situations in which data is not readily available and rely on subject matter experts and market research. Casual models are developed through the implementation of independent variables used as predictions of the dependent variables, with a foundation in regression analysis. Both of these methodologies are applied when data is not readily available and when the data is not thought to fluctuate with time. This section of the literature review will focus more on time series analysis which have been successfully applied to predict future demand within a specific certainty level for future demands.

The first time series methodology examined was the weighted average forecast. Similar to the exponentially weighted forecast, the weighted average forecast is used to predict future demands by taking past known demands and weighting them via probabilities that sum to one in order to generate a predicted demand (Mullick & Smith, 1971). The difference between the two models being that the weighted average forecast is not smoothed by the exponential function, making it more susceptible to variation in the data. This method is quite simple, and it has been used to successfully predict demand mainly for low volume items. This methodology has also been used extensively by the DoD in modeling attrition rates in our various services. Tomayo (2011) used this approach to successfully predict the attrition rates of the enlisted corps of the United

States Marine Corps. He was then able to validate his model by applying historical data, and calculated his probability of success using number of observations in his data set and the number of accurate predictions. Although simple, it is an approach that has worked in the past and could possibly be applied in predicting demands for transportation.

The second time series methodology examined is a commonly used technique called exponentially weighted forecasts. This methodology is based on weighting previous observations exponentially (Muth, 1960), and it has been used successfully in predicting demands within numerous different industries. The exponentially weighted forecast is known for its correction of persistent errors without responding to random variation in the data (Muth, 1960). The exponential function acts as a smoothing function, and the weights allow the user to base the prediction proportionally to previous years. For instance, if the user was to suspect time period i 's actual demand was affected by some unusual circumstance, then this period's demand could be weighted by a small proportion such that the random variation fails to have a significant impact on the upcoming period's demand. Such a methodology could prove useful to USTC, given past demands and relative information about that demand shipped is present. A shortcoming of this technique is that it must be applied for each specific variable being forecasted, and a situation with a large number of variables could become computationally exhaustive. The downfall of these two forecasting techniques is that they both fail to predict long-term behavior from the data and fail to identify turning points.

The third time series methodology commonly utilized was the Box-Jenkin methodology. The basic structure of the Box-Jenkin method is a univariate time series model in which the trend line, seasonal component and random irregular components are

slowly changed over time (Harvey & Todd, 1983). This concept is applied by utilizing a pattern recognition software that analyzes the data over time and identifies patterns in the data that can be modeled and accurately predicted (Hill & Woodworth, 1980). The benefits of the Box-Jenkins methodology is that it is very reliable for short-term (i.e., less than two years) predictions. A shortcoming of this methodology is that it requires pattern recognition software which can be financially expensive to the user, and it also would require supervision and upkeep by an employee educated on its use.

The fourth time series methodology commonly used was the Grey prediction model. This model is used in situations in which data is present, but not enough data is present for one to have an acceptable power corresponding to their prediction (Hsu & Chen, 2003). This technique combines residual modification with artificial neural networks. A shortcoming of this procedure is that it requires specific software packages that are capable of performing neural network analysis. Neural network analysis is a method of fitting the data using user-inputted randomness with various combinations of statistical and mathematical functions in order to accurately represent the data provided (Hsu & Chen, 2003). It is performed by reducing the randomness of the data such that demand can be predicted more easily (Hsu & Chen, 2003). Using 13 years of data to fit the model, Hsu and Chen were able to predict the two years of testing data with an average percent error of 3.88% compared to the actual demand. This demonstrated accuracy over such a long period of time shows the grey prediction model's ability to adapt to changes in demand over time. Such a model could in turn be applied to the transportation realm in order to accurately predict demand such that rates to better reflect actual cost.

The fifth and final methodology commonly used was bootstrapping. Unlike the previous methodologies, bootstrapping is not a forecasting technique. Instead, bootstrapping is a technique that utilizes the computational power of a computer in order to re-sample observed data. Persi Diaconis and Bradley Efron (1983) summarize the idea of bootstrapping by describing the technique as a tool to overcome the lack of data by constructing a sequence of fake data sets using only the data from the original sample (Diaconis, 1983). This is accomplished by essentially making an infinite amount of duplicates of your n observations and placing them all in a hat and then selecting n duplicates out of the hat creating a new sample with n observations (Diaconis, 1983). This provides a replicated data set in which desired statistical inferences can be drawn. The idea is that large amounts of the desired statistical inferences (the authors use 1,000) are drawn, providing confidence to inferences made. USTC currently sets rates using the average cost from the previous fiscal year as a baseline. Bootstrapping the data would allow them to have greater confidence in the costs used and have an estimate of the proportion of time in which cost could be larger than the cost applied in the methodology.

2.4 Rate Setting

I. Sea Rates

The study of liner, ocean/water pathway cargo, is one that is well documented in the academic field with basically the same construct applied to each methodology. In summary, liner freight rates can be systematically explained by costs and demand (Schneerson, 1976). According to Schneerson, the most important factors to shipping rates are the stowage factor (the ratio of volume to weight) and the unit values of each

commodity (Schneerson, 1976). Another key component found by Schneerson is that, although demand plays a major role in setting the rate of liner freight, the most important factor is accurately determining costs. This theory leads Schneerson to develop what he calls the Relevant Cost Concept for Pricing in which he breaks down costs between cost-in-port and cost-at-sea which are added together to develop the long run marginal cost (LRMC) and the long run average cost (LRAC), which are set equal to each other through the assumption of constant returns to scale. He then divides the cost-in-port into two components: direct and indirect handling costs, which allow him to use regression analysis to model the port cost and validate his assumptions on the important factors of cost-in-port. These values were then used in a microeconomic series of equations to calculate expected future port costs using predicted demands. The same technique was then applied using sea cost in order to calculate the expected future sea cost using predicted demands. Once the expected costs were calculated, rates were then set in a manner to cover these costs.

Another aspect of costs that may need to be evaluated is brought up by Forkenbrock (2001) who references cost incurred due to accidents. Many large companies self insure themselves with insurance companies' backing, meaning they cover all costs of accidents up to a certain threshold, saving them money in the long run by reducing insurance rates. Forkenbrock's work seeks to estimate the total cost incurred to a company via accidents, injuries, fatalities, and property loss as well as estimated emissions. External cost estimates of accidents are calculated by multiplying the number of fatal, personal injury, and property damage accidents by their corresponding per-event cost, and subtract the compensation. Dividing this cost by ton-miles allows for an estimate

of external cost per mile (Forkenbrock, 2001). “To generate comparable external cost estimates of accidents involving freight trains or trucks, then, we multiply the numbers of fatal, personal injury, and property damage accidents by the appropriate per-event cost and subtract the amount of compensation paid by the particular mode. Dividing the resulting external cost by the number of ton-miles allows us to estimate the per-ton-mile external cost for each mode” (Forkenbrock, 2001). “Our estimates of external costs for intercity general freight TL trucking and rail freight transportation imply that these costs are substantial. For general freight TL trucking, the external cost is 1.11 cent per ton-mile” (Forkenbrock, 2001). As shown in this analysis, this is a significant component to the rate set given such cost are taken on by the transporter. A 1.11 cent per ton-mile rate may not sound significant, but over the hundreds of thousands of miles transporters are responsible for each year, this cost grows substantially.

The final article reviewed depicted an interesting study on the potential for skewed cost data. A study performed by Brooks and Button (1996) examined the effects of shipping rates in the North Atlantic using directional data. Directional data references ocean travel either East to West, West to East, Southwest to Northeast, etc. The data provided, however, did not represent stowage factors or the approximate unit value of the particular goods in the data set found to be significant by Schneerson (1976). The authors, however, point out that these could be misleading variables, as variables with large stowage factors are often times loaded and shipped at incremental rates less than the actual cost to load the items in order to offset the ballast for higher priced goods which induce the cost. This could have potentially skewed Schneerson’s results had this practice been implemented at the time of his study (Brooks & Button, 1996). Further, this could

explain variations in cost data from year-to-year charged to USTC for liner shipments. Brooks and Button went on to conclude that the type of customer is a significant factor on rates indicating the value of quality business relationships. More applicable to our analysis, however, is that they found the direction of the shipping route to be a significant contributor to the rate charged. This could, in turn, be an indication of cheaper cost during shipments due to the flow of ocean currents, or could be impacting the model due to the quantity of goods being shipped is actually less on the routes with higher rates.

2.5 Statistical Process Control

Statistical Process Control (SPC) is a commonly implemented technique utilized to capture and control rare occurrences. It is a quality control method that can be manipulated within a large range of various processes in order to create control charts for managers to implement within their respective processes. In regards to USTC J8, this methodology can be implemented in order to define origin-destination/destination-origin and commodity code combinations which fail to fit the current rate setting methodology as outliers. Further, it can be implemented in order to define which origin-destination/destination-origin and commodity code combination rates had been set with undocumented art being applied (e.g., using experience-informed intuition by a budget analyst rather than the prescribed process).

The general approach to SPC is simple: define the desired statistic to be controlled, collect data, and then produce line charts in order to capture the variability in the data (StatSoft, 1984). If samples fall outside of the pre-specified limits, the process is declared to be out of control and action is taken to identify and correct the source of error

(StatSoft, 1984). In common practice, the most common implementations of SPC are referred to as the X-bar and R-charts.

The X-bar chart is a plot of means with the observations representing the horizontal axis and the values of the means representing the vertical axis. The center line is composed of the desired value of the statistic of observation, and the upper and lower control limits represent the threshold of acceptable variation of the statistic from the desired value. With regard to USTC J8, the X-bar chart could be utilized in order to analyze the behavior of the current methodology over the past several years. The desired observation would be a chart without large swings in NOR variation giving indication of a stable rate setting methodology. A chart of oscillating observances would indicate that the current methodology is not suited to adapt to the changes in the economic environment in which it acts.

The R-chart, similar to the X-bar chart, is a plot in which the observations represent the horizontal axis and the observed values represent the vertical axis. The difference is that the R-chart is a plot of ranges. A range of acceptable values of a statistic is determined, and the center line represents the middle of the range. The upper and lower control limits are represented by the upper and lower threshold values of the statistic. This methodology could be leveraged by USTC J8 in order to identify individual outliers within their rate setting process.

After correctly choosing the right statistic and chart to implement in order to accomplish the user's desired objective, the next most important decision is the choice of upper and lower control limits. In certain situations thresholds may be easily observed. With structural failures, a limit could easily be established and could already exist (e.g.,

the stress needed to break or warp a bolt). If not, it is common practice to utilize the Central Limit Theorem, due to large sample sizes, to make the assumption of normally distributed test statistics. Under this assumption, it is expected that 99.7% of the sample statistics will fall within three standard deviations of the mean (StatSoft, 1984), establishing the mean as the center line, and accepting the upper and lower control limits to be three standard deviations away from the mean.

SPC can be utilized and adapted to many fields. There are a wide range of various charts beyond what was discussed here. There are charts for variance control as well as charts for defective product control in production processes. Although simple in application, it is important to choose the correct statistic to observe as well as proper upper and lower control limits. This is a methodology that can be adapted to provide benefit to USTC J8 by identifying the outliers in their respective rate setting process. By identifying the outliers in their process, the NOR error in future years can potentially be reduced in magnitude by addressing recurrent outlier rates in order to set a rates having lesser marginal contributions to the NOR.

2.6 Summary

This review included a summary of several past works that depict the foundation of past rate setting in the transportation industry to include the effects of technology and globalization, forecasting of demand, and rate setting in the liner shipping industry. Although this thesis will focus on the forecasting of cost in order to set rates, it is important to gain insight into the field of business in which a solution is being generated for the problem. The research into the history of the transportation realm and rate setting

within the transportation industry provides background knowledge of the type of problem being examined. The research into forecasting of demand provides insight into the potential methods that have been applied in the past to solve similar problems. The combination of research on these three topics provides a more in-depth understanding of the problem under consideration.

III. Methodology

3.1 Introduction

The purpose of this chapter is to describe the methodology implemented in order to improve upon the existing rate setting methodology used by USTC J8 in order to set Transportation Working Capital Fund (TWCF) rates for their DoD customers. Improvement, in this scenario, is defined as a reduced absolute Net Operating Results (NOR) compared to the current baseline model. The remainder of this chapter is organized as follows. The introduction is followed by definitions, Section 3.2 discusses the current practice implemented by USTC, Section 3.3 discusses the method in which the current practice is evaluated, Section 3.4 discusses methods examined to improve upon the existing methodology, and the chapter is concluded with a summary.

Definitions:

Sets:

- $FY = \{\dots - 3, -2, -1, 0, 1, 2, 3, \dots\}$: fiscal year representation where $fy = 0$ corresponds to the current fiscal year, the next fiscal year is represented by $fy = 1$, the previous fiscal year is represented by $fy = -1$, and so forth.
- $I = \{01, 02, \dots, 59\}$: the origin or destination designation for a shipment, indexed by either i or j . This set I is alternatively represented as J . Shown in Appendix (E).
- $K = \{01, 02, 03, 04, 06, 07, 08, 09, 11, 12, 13\}$: shipping code designation, indexed by k .

Table 1: Shipping Code Descriptions

Commodity Code Designation	Nomenclature	Description
01	Reefer	Cargo on a breakbulk vessel kept under constant re Fridgeration.
02	Bulk	Non-packaged cargo (e.g., grain, coal, etc.)
03	POV	Individually owned vehicular cargo on a breakbulk vessel.
04	Hazardous	Hazardous cargo being moved by a breakbulk vessel.
06	General	Cargo being shipped that fails to fall into one of the other breakbulk categories.
07	Trailers	Trailers, van type semi-tailers, or stake and platform type semi-trailers moving on an roll on/roll off vessel.
08	Special	Any cargo item exceeding 10,000 lbs. or measuring more than 35 feet in any dimension.
09	Aircraft	Whole aircraft or aircraft fuselages being shipped by breakbulk or roll on/roll off vessels.
11	Reefer	Containerized perishable cargo that must be kept under constant refrigeration.
12	Vehicles	Containerized vehicular cargo.
13	General	General containerized cargo (e.g., not reefer or vehicular)

- M_{ijk}^{fy} , the set of observations/shipments from origin i to destination j of shipping code k in a given fy , $\forall i \in I, j \in J, k \in K$
- $\tilde{M}_{ijk}^{fy} \in M_{ijk}^{fy}$ represents the first *six months* of observations/shipments in M_{ijk}^{fy}

Parameters:

- C_{ijkm}^{fy} = cost (\$) per measurement ton, in $fy = 0$ for shipment m from origin i to destination j , of shipping code k , $\forall i \in I, j \in J, k \in K, m \in M_{ijk}^{fy}$
- $\tilde{C}_{ijkm}^{fy}$ = cost (\$) per measurement ton, in $fy = 0$ for shipment m from origin i to destination j , of shipping code k , $\forall i \in I, j \in J, k \in K, m \in \tilde{M}_{ijk}^{fy}$

- W_{ijk}^{fy} = weight (measurement tons) in $fy = 0$ of shipment m , from origin i to destination j of shipping code k , $\forall i \in I, j \in J, k \in K, m \in M_{ijk}^{fy}$
- $\tilde{W}_{ijk}^{fy}$ = weight (measurement tons) in $fy = 0$ of shipment m , from origin i to destination j of shipping code k , $\forall i \in I, j \in J, k \in K, m \in \tilde{M}_{ijk}^{fy}$
- R_{ijk}^{fy} = the actual rate set in fiscal year fy , in terms of (\$/measurement ton), for origin i to destination j of shipping code k , $\forall i \in I, j \in J, k \in K$
- a_{RR}^{fy} = Refresh Rate Adjustment implemented in fiscal year $fy \in FY$ (i.e. a rate of 5% corresponds to $a_{RR}^{fy} = 0.05$).
- a_{AOR}^{fy} = AOR Factor Adjustment implemented in fiscal year $fy \in FY$.
- a_{CRA}^{fy} = Composite Rate Adjustment implemented in fiscal year $fy \in FY$.
- F = total number of fiscal years examined

Calculated Variables:

- $\bar{C}_{ijk}^{fy}$ = average cost (\$) per measurement ton, to ship from origin i to destination j of shipping code k in fiscal year fy , $\forall i \in I, j \in J, k \in K$
- $\tilde{C}_{ijk}^{fy}$ = average cost (\$), for six months of cost data, per measurement ton, to ship from origin i to destination j of shipping code k in fiscal year fy , $\forall i \in I, j \in J, k \in K$
- $\overline{WC}_{ijk}^{fy}$ = weighted average cost (\$) per measurement ton, in $fy = 0$ for shipment origin i to destination j of shipping code k , $\forall i \in I, j \in J, k \in K$

- $\widetilde{WC}_{ijk}^{fy}$ = weighted average cost (\$), for six months of cost data, per measurement ton, in $fy = 0$ for shipment origin i to destination j of shipping code k , $\forall i \in I, j \in J, k \in K$
- r_{ijk}^{fy} = per unit rate dollars per measurement ton, to ship from origin i to destination j of shipping code k in fiscal year fy , $\forall i \in I, j \in J, k \in K$
- $\tilde{r}_{ijk}^{fy}$ = per unit rate dollars per measurement ton, for six months of cost data, to ship from origin i to destination j of shipping code k in fiscal year fy , $\forall i \in I, j \in J, k \in K$
- NOR_{ijk}^{fy} = dollar (\$) contribution to the net operating result from shipments from origin i to destination j of shipping code k , $\forall i \in I, j \in J, k \in K, fy \in FY$
- NOR_{Total}^{fy} = total net operating result in fiscal year fy , $\forall fy \in FY$
- $MNOR_{ijk}^{fy}$ = marginal contribution to the net operating result (\$/measurement ton) from origin i to destination j of shipping code k , $\forall i \in I, j \in J, k \in K, fy \in FY$
- $\overline{MNOR}^{fy}$ = average marginal contribution to the net operating result (\$/measurement ton) in fiscal year fy , $\forall fy \in FY$
- $\overline{\overline{MNOR}}$ = average of the average fiscal year marginal contributions (\$/measurement ton) to the net operating result.

- s_{MNOR}^{fy} = standard deviation of the marginal contribution to the net operating result for fiscal year fy , $\forall fy \in FY$
- O^{fy} = the total number of origin-destination/destination-origin and commodity code combinations in fiscal year fy , $\forall fy \in FY$
- $\bar{X}_{PT}^{fy}$ = point generated for the X-bar chart for each respective fiscal year fy in terms of the number of standard deviations away from the center line, $\forall fy \in FY$.
- PD_{ijk}^{fy} = the percent deviation of the calculated rate from origin i to destination j of shipping code k from the actual historical rate used by USTC in fiscal year fy , $\forall i \in I, j \in J, k \in K, fy \in FY$
- D^{fy} = the difference of the absolute value of the total NORs of six months of data and one year of data for each fiscal year fy , $\forall fy \in FY$
- $\bar{D}$ = the average difference of the absolute value of the total NORs of six months of data and one year of data across all fiscal years observed in the set FY .

3.2 Current Practice

USTC J8 sets rates for the projected fiscal year using approximately the first six months of cost data available during the current calendar year. The first six months of cost data from the current calendar year are cited for two reasons: it is believed that the current calendar year offers an accurate representation of negotiated shipping contracts for the upcoming fiscal year, and time restrictions due to shipping rates having to be set before the DoD budget is determined.

The first step in the current rate setting methodology is to calculate the weighted average cost for each origin-destination/destination-origin and commodity code

combination. For a given combination, the weighted average is calculated by summing all of the observed total cost and dividing by the total weight shipped via the given combination using Equation (1), as follows:

$$\overline{WC}_{ijk}^{FY} = \begin{cases} \frac{\sum_{m \in \widetilde{M}_{ijk}^{fy}} (w_{ijkm}^{fy} c_{ijkm}^{fy})}{\sum_{m \in \widetilde{M}_{ijk}^{fy}} w_{ijkm}^{fy}} & \forall i \in I, i = j, j \in J, k \in K, fy \in FY. \\ \frac{\sum_{m \in \widetilde{M}_{ijk}^{fy}} (w_{ijkm}^{fy} c_{ijkm}^{fy}) + \sum_{m \in \widetilde{M}_{jik}^{fy}} (w_{jikm}^{fy} c_{jikm}^{fy})}{\sum_{m \in \widetilde{M}_{ijk}^{fy}} w_{ijkm}^{fy} + \sum_{m \in \widetilde{M}_{jik}^{fy}} w_{jikm}^{fy}} & \forall i \in I, i \neq j, j \in J, k \in K, fy \in FY. \end{cases} \quad (1)$$

The weighted average is used so that each observation affects the average in proportion to its influence in the actual cost, thereby, preventing an outlier cost (e.g., a small volume shipment with an exorbitant cost) from having undue influence on the average cost used in the rate setting methodology.

Once the weighted average cost is computed, it is then adjusted using three factors: the Refresh Rate, the AOR Factor, and the Composite Rate Adjustment. The Refresh Rate, represented by a_{RR}^{fy} , acts in a similar manner to inflation in financial markets. It is intended to adjust the weighted average cost in the current calendar year to the projected cost environment for the upcoming fiscal year. The AOR factor, represented by a_{AOR}^{fy} , is used to balance the USTC's two-year running Net Operating Results (NOR). The goal of USTC is to achieve neither a surplus or deficit in each year. (Due to the inherently stochastic nature of forecasted rates, attaining a profit of zero is an unattainable task.) The AOR Factor is used to adjust the next year's rates to accommodate for the NOR of the previous fiscal year. For example, a large profit in a given year would result in an increased AOR factor over the next two years. This in turn

causes USTC to absorb some of the cost in the next fiscal year and, in essence, give back their profit to their customers. The Composite Rate Adjustment, represented by a_{CRA}^{fy} , is the final adjustment and is intended prevent rates from drastically changing year-to-year. It is believed that customers will become upset if they encounter large variations in the rate they pay each year, and so the Composite Rate Adjustment is intended to dampen the longitudinal variations in the rates on a year-to-year basis. Combining these adjustment factors with the weighted average cost produces the rate that is set for each origin-destination/destination-origin and commodity code combination via Equation (2), as follows:

$$\tilde{r}_{ijk}^{fy+1} = \widetilde{WC}_{ijk}^{fy}(1 + a_{RR}^{FY})(1 + a_{AOR}^{FY})(1 + a_{CRA}^{FY}); \quad \forall i \in I, i \leq j, j \in J, k \in K. \quad (2)$$

This rate calculation will be used as the baseline model throughout the remainder of this research.

3.3 Evaluating Current Practice

The second step taken in the research was to evaluate the current practice. This allowed for the current success of the methodology to be evaluated, as well as set a baseline to compare any future changes to the current methodology. This was done for cost data from six months of a calendar year as well as a full calendar year. A full calendar year was evaluated in order to determine if additional data resulted in more accurate cost predictions.

The evaluation phase was performed by first computing the rates for each of the fiscal years, 2009-2015. Once the rates were generated, they were then used to calculate the NOR contribution due to each commodity code and origin-destination/destination-

origin combination. Rates were only generated for combinations that had cost observations in both the six month and full year cost data. This was done to prevent evaluation from occurring between different commodity code sets. Further, NORs were only computed for data in the following fiscal year that had rates generated using cost data from the current calendar year. That is, predicting rates for commodity code and origin-destination/destination-origin combinations not having cost data in the current year was beyond the scope of this research.

With the rates calculated, the NOR was produced by multiplying the rate by the weight of the observation in the next fiscal year's data and then subtracting the product of the weight and unit cost of the next fiscal year's data using Equation (3), as follows:

$$NOR_{ijk}^{fy+1} = \sum_{m \in \tilde{M}_{ijk}^{fy}} (W_{ijkm}^{fy+1} r_{ijk}^{fy+1}) - (W_{ijkm}^{fy+1} C_{ijkm}^{fy+1}); \forall i \in I, i = j, j \in J, k \in K \dots$$

, $fy \in FY$.

$$NOR_{ijk}^{fy+1} = \sum_{m \in \tilde{M}_{ijk}^{fy}} (W_{ijkm}^{fy+1} r_{ijk}^{fy+1}) - (W_{ijkm}^{fy+1} C_{ijkm}^{fy+1}) + \sum_{m \in \tilde{M}_{jik}^{fy}} (W_{jikm}^{fy+1} r_{ijk}^{fy+1}) - (W_{jikm}^{fy+1} C_{jikm}^{fy+1}) \forall i \in I, i \neq j, j \in J, k \in K, fy \in FY. \quad (3)$$

The total NOR was then calculated by summing all the NOR components using Equation (4), as follows:

$$NOR_{Total}^{fy} = \sum_{i \in I} \sum_{j \leq i \in J} \sum_{k \in K} NOR_{ijk}^{fy} \quad (4)$$

This was accomplished for both six months and one year by changing the set of which m was drawn from.

To evaluate and determine if a significant difference occurred between six months of data and a full year, a paired *t*-test was employed. This test allowed for comparison amongst the differences between each year, since observations varied from year to year, and to make a statistically justified answer as to whether or not using a full year's worth of data added value to the process.

3.4 Improving Upon Current Practice

Objectives #2 and #3 of this research are to examine other sources of error in the rate setting methodology, as well to examine outliers in the data in order to determine where the art of rate projection should be applied and/or has been applied in the past. Section 3.4 depicts the methodology implemented in order to achieve those objectives. The following section is organized as follows: the introduction is followed by an explanation of the Composite Method Section A, then outlier analysis is introduced in Section B. Outlier Analysis consists of two Sections with B.1 introducing group behaviors and B.2 depicting individual outliers of marginal NORs followed by emulated rate outliers.

(A) The Composite Method

The Composite Rate method was first identified in discussion with J8; they are considering elimination of the Composite Rate Adjustment and allowing the effect to be accounted for within the AOR factor. As previously discussed, the weighted average cost for each individual commodity code is currently multiplied by a combined factor of the Refresh Rate, AOR Factor, and Composite Rate Adjustment, as shown in Equation (2), in order to produce rates. Provided that J8 sets both the AOR Factor and the Composite Rate

Adjustment and is considering eliminating the Composite Rate Adjustment, we combined the two factors into one using Equation (5), as follows:

$$(1 + \hat{a}_{AOR}^{fy}) = (1 + a_{AOR}^{fy})(1 + a_{CRA}^{fy}). \quad (5)$$

Replacing $(1 + a_{AOR}^{fy})(1 + a_{CRA}^{fy})$ in Equation (2) with Equation (5), and then utilizing this modified version of Equation (2) in Equation (3) in order to produce a new Equation (4), it yielded an equation with one unknown. The desired total NOR was then set equal to zero and $\hat{a}_{AOR}^{fy}$ was solved for using Equation (6), as follows:

$$\hat{a}_{AOR}^{fy} = \frac{\sum_{i \in I} \sum_{k \in K} \sum_{l=j \in J} \left(\sum_{m \in \bar{M}}^{fy} (w_{ijkm}^{fy+1} c_{ijkm}^{fy+1}) \right) \sum_{i \neq j \in J} \left(\sum_{m \in \bar{M}}^{fy} (w_{ijkm}^{fy+1} c_{ijkm}^{fy+1}) + \sum_{m \in \bar{M}}^{fy} (w_{jikm}^{fy+1} c_{jikm}^{fy+1}) \right)}{\sum_{i \in I} \sum_{k \in K} \sum_{i=j \in J} \left(\overline{WC}_{ijk}^{fy} (1 + a_{RR}^{fy}) \sum_{m \in \bar{M}}^{fy} w_{ijkm}^{fy+1} \right) \sum_{i \neq j \in J} \left(\overline{WC}_{ijk}^{fy} (1 + a_{RR}^{fy}) \sum_{m \in \bar{M}}^{fy} w_{ijkm}^{fy+1} + \sum_{m \in \bar{M}}^{fy} w_{jikm}^{fy+1} \right)} - 1 \quad (6)$$

Given an NOR equal to zero, this methodology allows the calculation of the exact AOR Factor for each fiscal year that would have resulted in the desired NOR. Ultimately, this method depends on perfect information that is not available at the time rates are set, but it may provide insight into more accurate AOR Factor estimations.

Despite the Composite Method's power of yielding NORs equal to zero, given perfect information, its flaws lie in its inability to correct for large individual NORs. Currently, this method seeks to set the total NOR equal to zero. In an ideal, a rate setting process would drive the NOR to zero for each shipping commodity and origin-destination/destination-origin combination. This would address the potential concerns for fair rate setting, specifically eliminating cases when customers pay more or less than their fair share of the total cost of shipping goods within the DoD.

(B) Outlier Analysis

Outlier analysis was introduced as a method of evaluating the current methodology on a year-to-year basis, as well as identifying origin-destination/destination-origin combination outliers in which the current methodology fails to set a rate that accurately recoups cost. The metric chosen to be examined was the NOR contribution per measurement ton or marginal NOR (MNOR) calculated using Equation (7), as follows:

$$MNOR_{ijk}^{fy} = \frac{NOR_{ijk}^{fy}}{\sum_{m \in \bar{M}_{ijk}^{fy}} \tilde{W}_{ijkm}^{fy}}; \forall i \in I, i = j, j \in J, k \in K, fy \in FY.$$
$$MNOR_{ijk}^{fy} = \frac{NOR_{ijk}^{fy}}{\sum_{m \in \bar{M}_{ijk}^{fy}} \tilde{W}_{ijkm}^{fy} + \sum_{m \in \bar{M}_{jik}^{fy}} \tilde{W}_{jikm}^{fy}}; \forall i \in I, i \neq j, j \in J, k \in K, fy \in FY. \quad (7)$$

The metric was chosen such that the MNOR of each origin-destination/destination-origin combination were examined preventing volume disparity from skewing the results of the analysis.

The following portion of Section 3.4 will depict the three types of outlier analysis performed in this research. The first method looks into the performance of the overall mean of the marginal NORs across fiscal years 2009-2015, the second method identifies outliers in each fiscal year and looks at their performance over the course of fiscal years 2009-2015, and the third method utilizes the same underlying concepts applied in the second method but is adapted to identify calculated rates which differ from historical rates used by USTC.

(B.1) Mean of Means Outlier Analysis

The first outlier analysis conducted examines the performance of USTC J8's rate setting performance in regards to marginal NORs over the course of fiscal years 2009-

2015. This was done by applying a variation of statistical process control called an X-bar stabilized chart.

The X-bar stabilized chart utilizes the average MNOR from each fiscal year calculated using Equation (8), as follows:

$$\overline{MNOR}^{fy} = \frac{\sum_{i \in I} \sum_{j \in J} \sum_{k \in K} MNOR_{ijk}^{fy}}{Ofy}, \forall fy \in FY. \quad (8)$$

The standard deviation of the MNORs from each fiscal year using Equation (9), as follows:

$$s_{MNOR}^{fy} = \sqrt{\frac{1}{Ofy-1} \sum_{i \in I} \sum_{j \in J} \sum_{k \in K} (MNOR_{ijk}^{fy} - \overline{MNOR}^{fy})^2}, \forall fy \in FY. \quad (9)$$

The center line of the control chart was generated by taking the average of the average MNORs from each fiscal year utilizing Equation (10), as follows:

$$\overline{\overline{MNOR}} = \frac{\sum_{fy \in FY} \overline{MNOR}^{fy}}{F}. \quad (10)$$

Since each individual year yielded different averages and standard deviations, uniform upper and lower control limits could not be generated. Thus, the difference between the average marginal cost of each fiscal year and the center line, the average of the average MNORs, was scaled by each individual fiscal year's respective standard deviation using Equation (11), as follows:

$$\bar{X}_{PT}^{fy} = \frac{(\overline{MNOR}^{fy} - \overline{\overline{MNOR}})}{s_{MNOR}^{fy}}, \forall fy \in FY. \quad (11)$$

The results from Equation (11) provided a scaled version of the results from Equation (8) in terms of the number of standard deviations the average MNOR contribution for a given fiscal year fell from the overall average MNOR

Figure 1, shown below, is an example of how the methodology is applied in order to evaluate the performance of the current rate setting methodology implemented by USTC J8. Equation (10) produces the center line. The individual points, calculated by Equation (11), are represented by the individual boxes. Any individual observation located above or below the respective y-values of 2 and -2 represent sporadic behavior in terms of marginal NORs of that given fiscal year compared to the other fiscal years as a whole. The values of 2 and -2 were chosen for upper (UCL) and lower (LCL) control limits as statistical process control utilizes the central limit theorem, due to large data samples, which assumes the data follows a normal distribution. Provided this assumption, approximately 95.45% of the data should fall within three standard deviations of the mean, or center line, making any point falling outside these bounds cause for concern. In this example, fiscal year 2011 would be cause for concern as it lies below the lower control limit. Trends in the data, either with a positive or negative slope, are also cause for concern as they represent a method that is either consistently producing increasing or decreasing marginal NORs. Ideally, the data would produce randomly dispersed points within the bounds of the upper and lower control limits, providing indication that the marginal NORs of each fiscal year are independent of one another and that the methodology is performing in a scientifically responsible manner.

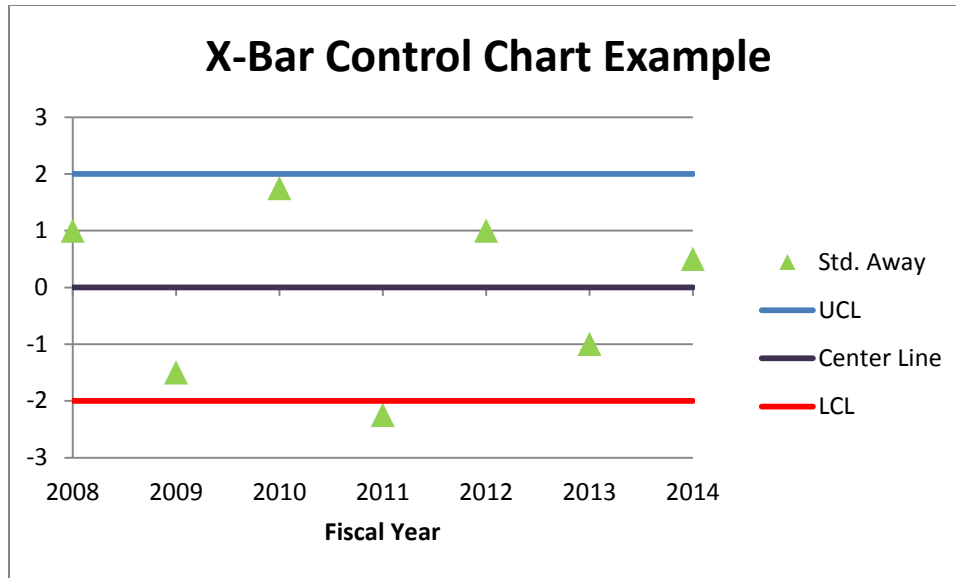


Figure 1: Control Chart Example

(B.2.1) Individual Fiscal Year Outliers

After analyzing the performance of the methodology's behavior within each individual year compared to the collective performance over the time period of this study, the individual outliers were examined using a similar approach to the one implemented in Section B.1.

To identify outliers, each origin-destination/destination-origin and commodity code combination's respective MNOR was calculated via Equation (7) for each individual fiscal year. These individual marginal NORs were then compared to two standard deviations away from their respective means. Any observation falling outside the bounds of two standard deviations away from the mean were identified as an outlier for that particular year.

Outliers in this particular method represent an origin-destination/destination-origin and commodity code combination in which the current methodology fails to set a

rate in which accurately recoups the actual cost of shipment per measurement ton in the given fiscal year. In particular, an observation falling more than two standard deviations away from the mean, as depicted in Figure 2, falls within a category in which we expect to occur less than 5% of the time. Further, a significant percentage of outliers falling more than two standard deviations away from the mean indicate that the methodology as a whole is failing to accurately address the cost needed to be recouped over the set of origin-destination/destination-origin and commodity code combinations for a given fiscal year.

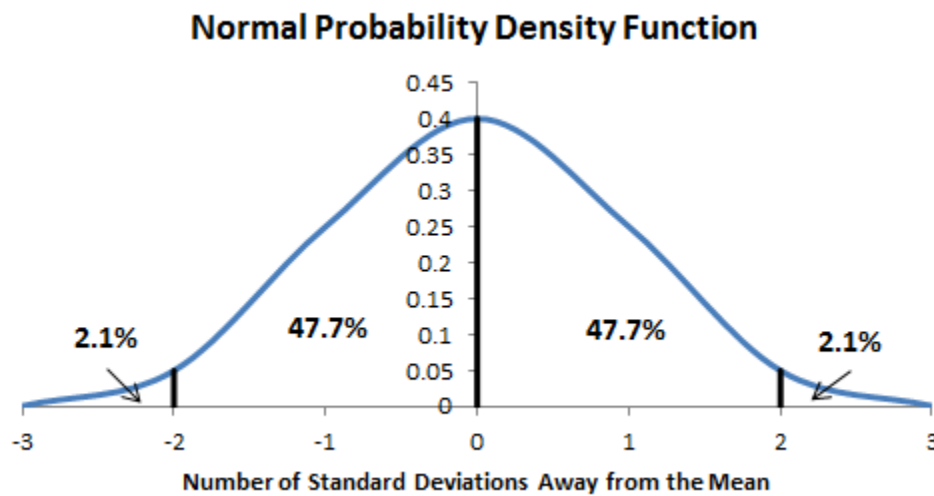


Figure 2: Normal Probability Density Function

After identifying the outliers in each given fiscal year, these outliers were then examined over the time period of the study using the same methodology reviewed in Section B.1 with the only change being now we are examining the individual outliers relative to the mean of the means instead of the average marginal NOR for the respective fiscal year. Examining just the outliers in this fashion, there are three distinct behaviors we anticipate observing. Those behaviors are (a) an outlier being corrected, (b) an outlier

being over corrected for, or (c) an outlier which remains an outlier. If the outlier is corrected for, as shown by the line with box-shaped points in Figure 3, the methodology is considered to be adequately representing the particular origin-destination/destination-origin and commodity code combination, and that the outlier was caused by a source of error other than the current methodology. If the outlier is overcorrected for, as shown by the line with circle-shaped point in Figure 3, the methodology is considered to be volatile to the economic conditions of the origin-destination/destination-origin and commodity code combination, and it requires special attention in order to dampen the swings. Finally, if the outlier is not corrected for, as shown by the line with star-shaped points in Figure 3, the methodology is not considered to adequately represent the given origin-destination/destination-origin combination and will required special attention in the rate setting process in order to adequately set the rate.

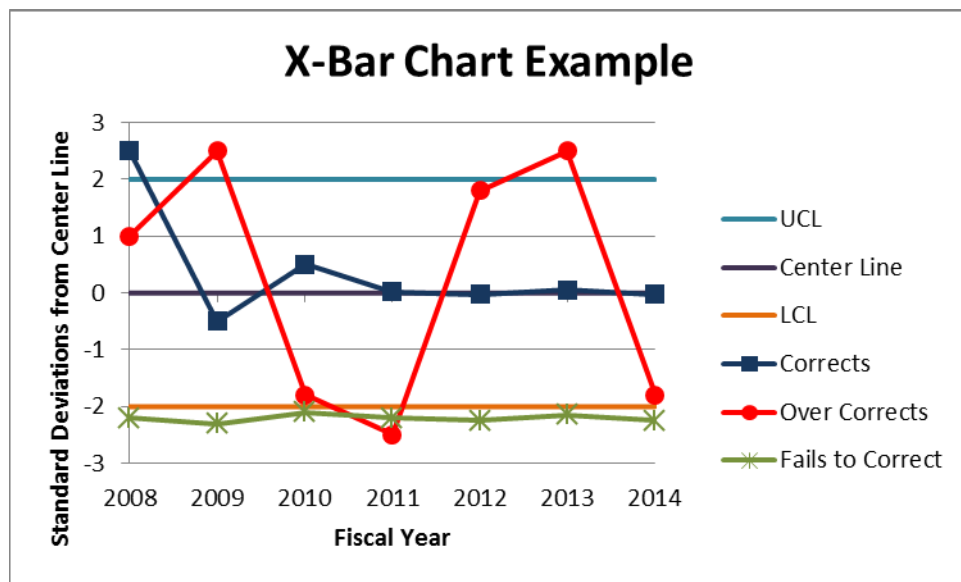


Figure 3: Individual Outlier Behavior Example

(B.2.2) Individual Rate Outliers

In discussions with USTC J8, it was noted that some portion of the current rates were set by deviating from the documented methodology. In order to determine where this attention was directed, statistical process control was applied. This approach was chosen as it cannot be expected for exact rates to be calculated during emulation, as the data sets used have a high probability of variation. Thus, ruling out looking at actual rates that have minimal deviation from the emulated rate as outliers.

Despite the possibility of variation within data sets used, it can be reasonably assumed that the emulated and actual rates will have a small deviation when the current methodology was applied and a larger deviation when the methodology was altered for a specific origin-destination/destination-origin and commodity code combination. Thus, the percent deviation from the actual rate set, Equation (12), was used instead of the marginal NOR in this outlier analysis. The average percent deviation, Equation (13), and the standard deviation of the percent deviations, Equation (14), were calculated for the control charts to be produced identically to the method prescribed to the marginal NORs.

$$PD_{ijk}^{fy} = \frac{(r_{ijk}^{fy} - R_{ijk}^{fy})}{R_{ijk}^{fy}} * 100 \quad \forall i \in I, j \in J, k \in K, fy \in FY \quad (12)$$

$$\overline{PD}^{fy} = \frac{\sum_{i \in I} \sum_{j \in J} \sum_{k \in K} PD_{ijk}^{fy}}{O^{fy}}, \quad \forall fy \in FY \quad (13)$$

$$S_{MNOR}^{fy} = \sqrt{\frac{1}{O^{fy}-1} \sum_{i \in I} \sum_{j \in J} \sum_{k \in K} (PD_{ijk}^{fy} - \overline{PD}^{fy})^2}, \quad \forall fy \in FY \quad (14)$$

The outliers for fiscal years 2014 and 2015 were identified using this methodology, as these specific fiscal year's rates were the only two years available. Despite the smaller sample, it is believed that identifying these outliers will provide

insight into how often and for which rules the prescribed methodology was not used. If more than two years of actual rates were available, the methodology provided in Section B.1 could be applied to look at the behavior such outliers over time. In this particular instance a direct comparison is appropriate as only two years were being compared.

3.5 Summary

Chapter 3 explained the varying methodologies implemented in this research. This chapter began with an introduction to the problem, and a description of all sets, parameters and calculated variables used in the research. This was followed by a description of the current practice implemented by USTC J8. Then the method of evaluating the current practice was explained by showing how rates were emulated and NORs were calculated for both six month and full year cost data sets. Following the evaluation of the current practice, methods of improving the current practice were introduced. This section introduced and explained the Composite Rate Method followed by an explanation of the various ways in which Outlier Analysis was applied to the data. In conclusion, this chapter introduced and explained the methodologies implemented within this research in order to accomplish the research objectives highlight in Section 1.4.

IV. Analysis and Results

4.1 Chapter Overview

This chapter details the results from the analysis conducted in this research on behalf of USTC J8. The chapter begins with the results gathered by evaluating the current practice before presenting the results from Outlier Analysis as well as the Composite Method.

4.2 Evaluating Current Practice

The evaluation of the current practice was accomplished by first calculating the rates for each fiscal year as outlined in Section 3.2. Upon completion of the calculation of the rates, the performance of the current methodology was evaluated by calculating the total NOR for each fiscal year and comparing the results based on the current utilization of six months of cost data to the results based on the utilization of one year of cost data. The theory behind this inquiry was that more data would yield better results in terms of a smaller absolute value of the total NOR, thus enabling the analysis of whether deviations from the NOR goal of zero were caused by using the smaller data samples.

To perform the evaluation, the absolute-value of each respective fiscal year's NOR using both six months of data and a full year of data was computed. Then a hypothesis test was conducted utilizing the differences between pairs of the absolute value of the total NORs. The absolute value of each respective total NOR was taken, since a negative NOR of a certain magnitude is equally as undesirable as a positive NOR of the same magnitude. The null hypothesis of the hypothesis test was that the difference was equal to zero, and the alternative hypothesis was that the difference was not equal to

zero. This hypothesis test allowed the research to determine whether or not there was a statistical advantage to using an additional six months of data in the rate setting process by first determining if there was a statistically significant difference between NORs of six months of data compared to a full year worth of data.

$$D^{fy} = \left| NOR_{Total \in \tilde{M}_{ijk}^{fy}}^{fy} \right| - \left| NOR_{Total \in M_{ijk}^{fy}}^{fy} \right| \quad (15)$$

$$H_0: \bar{D} = 0 \quad (16)$$

$$H_1: \bar{D} \neq 0$$

Table 2: 6 Mo. vs. 1 Yr. Absolute Value Hypothesis Test

NOR Absolute Value Hypothesis Test			
Year	6 Mo. Total NOR	1 Yr. Total NOR	Difference
2008	\$643,489,451.80	\$1,020,773,693.93	-\$377,284,242.13
2009	\$640,377,982.47	\$246,567,321.77	\$393,810,660.70
2010	\$231,942,489.02	\$310,688,812.89	-\$78,746,323.87
2011	\$2,263,095,157.86	\$1,807,929,549.71	\$455,165,608.15
2012	\$198,375,070.19	\$261,340,625.57	-\$62,965,555.38
2013	\$160,774,484.36	\$148,425,341.32	\$12,349,143.04
2014	\$45,854,486.45	\$46,430,739.79	-\$576,253.34
Mean:			\$48,821,862.45
Std:			\$288,166,452.19
t-value:			0.448
Df:			6
P-Value:			0.67

The test resulted in failing to reject the null-hypothesis at all reasonable significance level possessing a p-value of 0.67. Thus, it was concluded that more data did not benefit the NOR leading to other avenues of evaluation within the research.

4.3 Rates

This section seeks to analyze the results obtained by implementing the methodology explained by Section 3.4 (B.2.2) of this research. The goal of this analysis

was to identify what rates were set by the documented methodology and which rates were set by implementing an undocumented methodology, to the data provided. As indicated in Section 3.4 (B.2.2), the relative percent deviations of our predicted rates and from the actual, historically-set rates were calculated to prevent minor differences between the data sets from dictating the identification of outliers.

The same statistical process applications implemented throughout the remainder of this research were applied to identifying the outliers in the rate setting methodology, with the difference being that the common practice of using two and three respective standard deviations away from the mean to identify outliers was set aside for the purpose of this analysis. The reason for this decision is that the standard deviations observed were sufficiently large, and caused two standard deviations to have upper and lower control limits of roughly positive and negative 200%. In other words, the upper and lower control limits would fail to identify a predicted rate that is slightly less than two times the actual rate as an outlier. Due to the assumption that deviation in the rates either arise from deviations in the data set utilized to produce rates or an alteration to the methodology, we proceeded by categorizing all predicted rates greater than or equal to 100% away from the actual rate set as an outlier. We believed this threshold was very generous in allowing for deviations caused by the variation in the data set while still portraying the rates for origin-destination/destination-origin and commodity code combinations for which an alteration to the methodology was likely implemented.

Tables 3-5 show the origin-destination/destination-origin and commodity code combinations in which the data determined that a different rate setting approach was likely implemented. These combinations all possessed percent deviations greater than

100%, leading to the conclusion that a different procedure was utilized. That is, instead of applying the prescribed rate setting procedure, an analyst adjusted the rates using their experience-informed judgement (i.e., the *art* of rate setting). Across the calculated rates set for the 2014-2016 fiscal years, approximately 8%, 10%, and 4% of the rates were determined to have likely been set using a different methodology as defined by a percent deviation greater than 100% in magnitude. With the exception of FY16, which can be contributed to the smaller number of rates set for FY16 at the time of this research, the rates deemed to be outliers grow at approximately the same rate as the threshold of declaring an outlier is reduced. Preceding it should also be noted that data was only available for us to calculate between 350-400 rates from each fiscal year compared to the roughly 9,000 rates set by USTC J8 each fiscal year, roughly 5% of the total rates set on a yearly basis.

Table 3: Robust Examination of Rate Outliers

% of rates	FY14	FY15	FY16
±100% deviation	8%	10%	4%
±75% deviation	8%	12%	8%
±50% deviation	16%	17%	16%
±25% deviation	35%	29%	36%

Table 4: Rate Outliers from 2014 Rates

Rate Outliers Computed by 2013 Data as Compared to Historical 2014 Rates		
Origin ID	Destination ID	Relative Deviation
01 13 06	13 01 06	202%
01 13 11	13 01 11	345%
01 13 12	13 01 12	539%
01 13 13	13 01 13	148%
01 17 06	17 01 06	124%
01 20 11	20 01 11	156%
01 21 06	21 01 06	168%
01 23 12	23 01 12	339%
01 46 06	46 01 06	101%
02 15 06	15 02 06	285%
02 23 08	23 02 08	173%
02 23 12	23 02 12	652%
03 23 12	23 03 12	550%
03 30 08	30 03 08	760%
03 40 06	40 03 06	325%
03 40 07	40 03 07	154%
04 23 08	23 04 08	135%
17 20 12	20 17 12	187%
17 21 06	21 17 06	193%
20 46 12	46 20 12	189%
23 51 06	51 23 06	109%
25 26 06	26 25 06	185%
27 34 06	34 27 06	294%
27 34 08	34 27 08	241%
28 29 06	29 28 06	310%
28 29 08	29 28 08	272%
29 49 08	49 29 08	245%
29 50 07	50 29 07	389%
29 52 07	52 29 07	444%
32 33 06	33 32 06	129%
34 52 13	52 34 13	209%
43 43 13	43 43 13	490%
46 51 13	51 46 13	628%

Table 5: Rate Outliers from 2015 Rates

Rate Outliers Computed from 2014 Data as Compared to Historical 2015 Rates		
Origin ID	Destination ID	Relative Error
01 13 08	13 01 08	231%
01 90 11	59 01 13	531%
01 90 12	59 01 13	357%
01 90 13	59 01 13	412%
01 91 07	59 01 13	130%
01 91 11	59 01 13	386%
01 91 12	59 01 13	347%
01 91 13	59 01 13	393%
01 92 06	59 01 13	837%
01 92 12	59 01 13	265%
01 92 13	59 01 13	334%
01 93 06	59 01 13	198%
01 93 07	59 01 13	236%
01 93 08	59 01 13	198%
01 93 12	59 01 13	149%
02 90 06	57 02 13	929%
02 90 07	57 02 13	939%
02 90 08	57 02 13	1504%
02 90 12	57 02 13	608%
02 90 13	57 02 13	621%
02 92 07	57 02 13	315%
02 92 08	57 02 13	197%
02 93 06	57 02 13	179%
02 93 07	57 02 13	185%
02 93 08	57 02 13	114%
02 93 12	57 02 13	276%
02 93 13	57 02 13	289%
03 39 08	39 03 08	788%
03 52 04	52 03 04	806%
03 93 06	58 03 13	903%
03 93 07	58 03 13	650%
03 93 08	58 03 13	434%
03 93 12	58 03 13	279%
03 93 13	58 03 13	226%
04 19 06	19 04 06	116%
17 22 08	22 17 08	101%
17 93 06	57 17 13	177%
22 23 08	23 22 08	108%
23 51 06	51 23 06	102%

Table 6: Rate Outliers from 2016 Rates

Rate Outliers Computed 2015 Data as Compared to Historical 2016 Rates		
Origin ID	Destination ID	Relative Error
01 20 12	20 01 12	119%
01 21 07	21 01 07	122%
01 21 11	21 01 11	105%
01 46 13	46 01 13	134%
02 18 13	18 02 13	153%
02 23 12	23 02 12	256%
02 44 08	44 02 08	138%
02 92 07	92 02 07	166%
03 93 07	93 03 07	171%
04 19 06	19 04 06	1081%
04 26 07	26 04 07	183%
04 27 06	27 04 06	206%
17 20 06	20 17 06	216%
19 17 08	17 19 08	130%
20 01 12	01 20 12	109%
20 19 08	19 20 08	136%
21 01 07	01 21 07	167%
21 23 11	23 21 11	106%
23 02 12	02 23 12	256%
23 22 12	22 23 12	264%
23 93 13	93 23 13	236%

Reviewing Tables 3-5, it is of interesting note that all of the identified outliers possessed positive Relative Error values. This means that the current methodology produced a rate that was over twice as large as the historically set rate for all identified outliers, and never set a rate that was twice as small or smaller in magnitude compared to the historically set rates. Thus indicating that the prescribed method of *art* was to take individual rates and drive the rate computed by the current methodology by an undetermined amount. Past methods utilized to determine where to apply the *art* is not identified by this method, and neither is the amount at which the calculated rates are

reduced. With an average and standard deviation of the outliers determined by relative error being 322% and 257% respectively, it is implied that the percent at which the calculated rate was reduced was dealt with on an individual basis.

4.4 Predicted Rate Contributions to the NOR

This section of the thesis is utilized to outline the results from the methodology depicted in Sections 3.4 (B.1) and 3.4 (B.2.1). After identifying the historical rates likely set using a technique other than the prescribed methodology, we then sought to analyze the behavior of the current methodology to determine the areas in which an alternate methodology *should* be applied in order to minimize the NOR of each fiscal year. This analysis was conducted by implementing a new statistic, as described in Section 3.4 (B), the marginal contribution to the NOR (MNOR).

Before continuing it is important to note that the remainder of this research was conducted by setting aside the origin-destination/destination-origin and commodity code combination 02 23 01 from FY12 cost data, Reefer Breakbulk being shipping from CONUS (Gulf Coast) to the Arabian Gulf. This specific data point possessed an MNOR of \$262,275.02/lb., which otherwise influenced the standard deviation of the MNORs in FY12 and prevented any other outliers from being identified. The decision was made for it to be left out, but it should be noted and considered as a potential source of error in the real world system.

4.4.1 Mean of Means Outlier Results

Following Section 3.4 (B.1), this section discusses the results found in the Mean of Means Outlier Analysis. Table 6, summarizes the results from Equations (8)-(11) in

Section 3.4 (B.1). The average marginal NORs are shown, followed by their respective standard errors, the average of the average marginal NORs, as well as each respective X point which corresponds to the number of standard deviations its respective average marginal NOR fell from the overall mean.

Table 7: Mean of Means Results

Year	Average (6 Mo.)	Standard Deviation (6 Mo.)	Xpt (6 Mo.)
2008	34.276	114.985	0.008
2009	-15.001	206.653	-0.234
2010	47.824	253.175	0.057
2011	-7.405	301.969	-0.135
2012	94.773	418.270	0.147
2013	38.564	229.908	0.023
2014	40.673	327.590	0.022
	$\overline{\overline{MNOR(6 Mo.)}}$		
	33.386		

The more meaningful column in Table 6 is the X point column. This column scales the performance of the current methodology for a given fiscal year to the performance of the methodology throughout all the years examined in this study. Values of three or larger in this column would be severe cause for concern, as it would indicate a year in which on average the methodology vastly differed from its history of performance. In Figure 4, the values from the X point column in Table 6 are depicted.

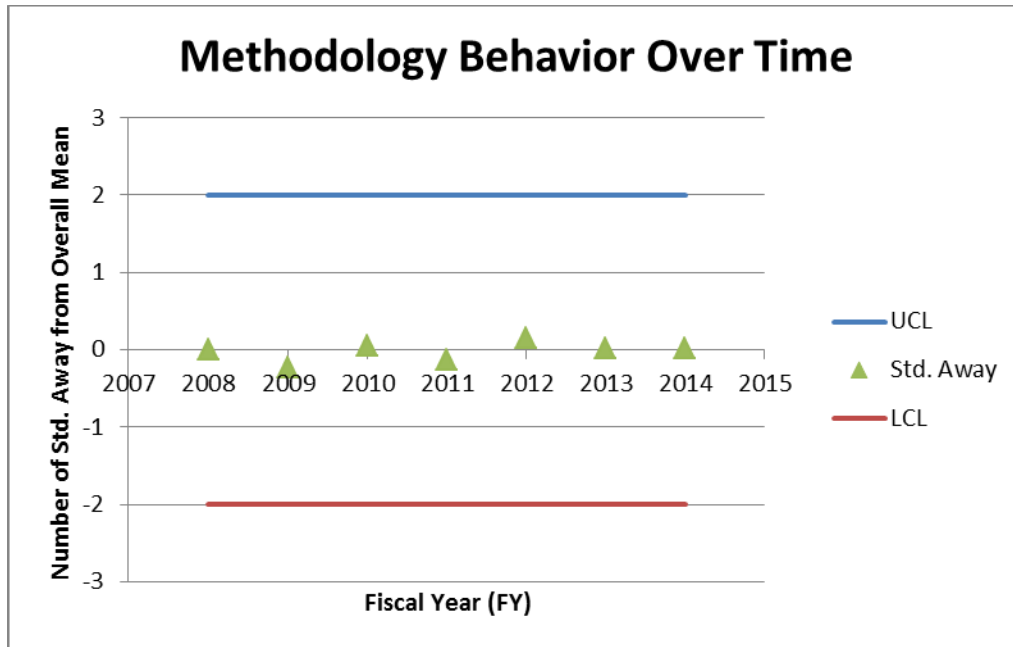


Figure 4: Behavior Over Time

Figure 4 shows that the behavior of the current methodology implemented by USTC J8 is consistent over the duration of the data examine in this study. This is not to be misinterpreted as accurate in terms of the individual NORs, as this outlier methodology does not examine the efficiency of the resulting NORs. This method is just used to look at the behavior of the methodology in terms of the resulting average marginal NORs.

The results gathered indicate that the methodology does, in fact, behave in a consistent manner, which allowed for the remainder of the research to focus on areas which may improve the accuracy of the current methodology.

4.4.2 Individual Outliers (MNOR)

This section seeks to implement the methodology explained in Section 3.4 (B.2.1) while building upon some the results from the previous section. After identifying in Section 4.4.1 that the current methodology yields stable NOR behavior over time, the

individual MNOR outliers were identified in an effort to determine where a budget analyst *should* deviate from the prescribed rate setting methodology.

The marginal NORs (as referenced in Section 4.4.1 and explained in Equation (8) in Section 3.4 (B.2.1)) were calculated and the outliers were identified as the marginal NORs falling outside of three and two standard deviations away from their respective fiscal year means. The number of outliers present within three and two standard deviations respectively from the mean was then analyzed in terms of percentages of the data as a whole. As noted in Figure 2 of Section 3.4 (B.2.1), it was expected to have roughly 1% and 5% of the data fall outside of three and two standard deviations respectively during each fiscal year. To check this assumption of normality, two separate hypothesis tests were conducted on the results. The first hypothesis test, shown below, tests the average percent of outliers falling outside of three standard deviations with the null hypothesis as the mean is greater than or equal to 1% and the alternative hypothesis that the mean is less than 1%. The second hypothesis test was conducted in the same manner for two standard deviations with the only exception that it was conducted for greater than or equal to 5% instead of 1%.

$$H_0: \bar{X} \geq 0.01 \tag{17}$$

$$H_A: \bar{X} < 0.01$$

$$H_0: \bar{X} \geq 0.05 \tag{18}$$

$$H_A: \bar{X} < 0.05$$

Table 8: Data Falling 3 and 2 Standard Deviations Away from the Mean Hypothesis

Test Results

Percent of Data Falling 3σ & 2σ Away from the Mean		
Year	3σ	2σ
2008	2.475%	4.950%
2009	1.240%	2.893%
2010	1.629%	4.235%
2011	1.994%	5.128%
2012	1.937%	3.148%
2013	2.427%	3.641%
2014	0.829%	1.657%
Mean:	0.018	0.037
Std:	0.006	0.012
t-stat:	3.459	-2.883
Df:	6.000	6.000
P-Value:	0.993	0.014

Viewing the results in Table 7, it is clear that on average more than 1% of the data falls outside of three standard deviations from the mean marginal NORs during each fiscal year which is indicated by the p-value of 0.993 for this particular test, which is greater than any acceptable significance level that could be used. During the second hypothesis test, the data rejected the null hypothesis at all significant levels greater than 0.014, as indicated by the p-value. This resulted in accepting the alternative hypothesis at the 0.015 significance level that, on average, less than 5% of the data falls outside of two standard deviations from the mean. Given this result is more closely aligned with the assumption of normality in the data set required for statistical process control, two standard deviations was used to identify outliers in the subsequent analysis.

1) One-Time/Partial Outliers

To examine MNOR outliers the data was sorted into three separate categories: one-time outliers, partially matched outliers, and duplicated outliers. One-time outliers, shown in Table 8, only appeared once as an origin-destination/destination-origin and commodity code combination (TAP-CC). In addition to only appearing once, the origin-destination/destination-origin combinations were not observed in any subsequent fiscal years. Some of these outliers were corrected for by the prescribed rate setting methodology in their following fiscal years, while others failed to be observed in the remaining fiscal years. No further analysis was conducted on these one-time outliers.

Upon eliminating MNOR outliers that occur only once, the next step taken was to look at outliers that shared origin-destination/destination-origin combinations within the respective TAP-CC codes. This was done in order to look at particular origin-destination combinations that consistently produce outliers, but not necessarily outliers that share a corresponding shipping code commodity label. Such outliers could be caused by smaller/larger cost (e.g., port-handling cost, tariffs, security, etc.) at the origin/destination resulting in the unorthodox deviations in the marginal NOR. The resulting outliers identified are outlined by their origin-destination/destination-origin coding and time frame in which they were observed in Table 9.

Table 9: One-time outliers

Year	TAP-CC
2008	01 16 07
2008	02 20 11
2008	02 43 13
2008	04 50 08
2008	23 43 08
2009	01 15 06
2009	01 17 11
2009	04 21 13
2009	23 23 13
2010	01 20 06
2010	01 24 04
2010	01 49 06
2010	03 33 06
2010	03 34 08
2010	23 24 06
2010	27 27 06
2011	02 10 12
2011	04 27 04
2011	23 25 13
2011	50 52 06
2012	03 52 04
2012	19 23 06
2013	01 13 12
2013	01 21 06
2013	27 28 06
2014	03 52 04
2014	04 19 06
2014	02 90 08

Table 10: Partially Matched Outliers

Partially Matched Outliers	
Years	Origin Destination Pair
2008-2013	01 23
2008-2013	01 43
2008-2010	02 21
2010-2011	02 17
2011-2013	02 23
2011-2013	03 23
2010-2011	22 23
2011-2012	23 27
2011-2012	17 23
2011-2013	01 46
2013	46 51

Cross referencing between the data depicted in Tables 8 and 9, problematic individual locations were readily identified due to their recurrent nature in both tables. This led to examining all of the 82 outliers identified across fiscal years 2008-2014 in terms of frequency of occurrence by location. Since the data was provided in origin-destination/destination-origin pairs, it was difficult to distinguish between a location being identified as an outlier due to the origin or destination. This led to the total number of outliers for each respective location accounting for a location being either an origin or a destination. Figure 5, shows the initial results from this analysis. From the 82 observed outliers across fiscal year 2008-2014, locations coded by 23, 01, 02, and 03 comprise 17.68%, 17.07%, 11.59%, and 6.10% on aggregate majority of the outliers identified by MNOR.

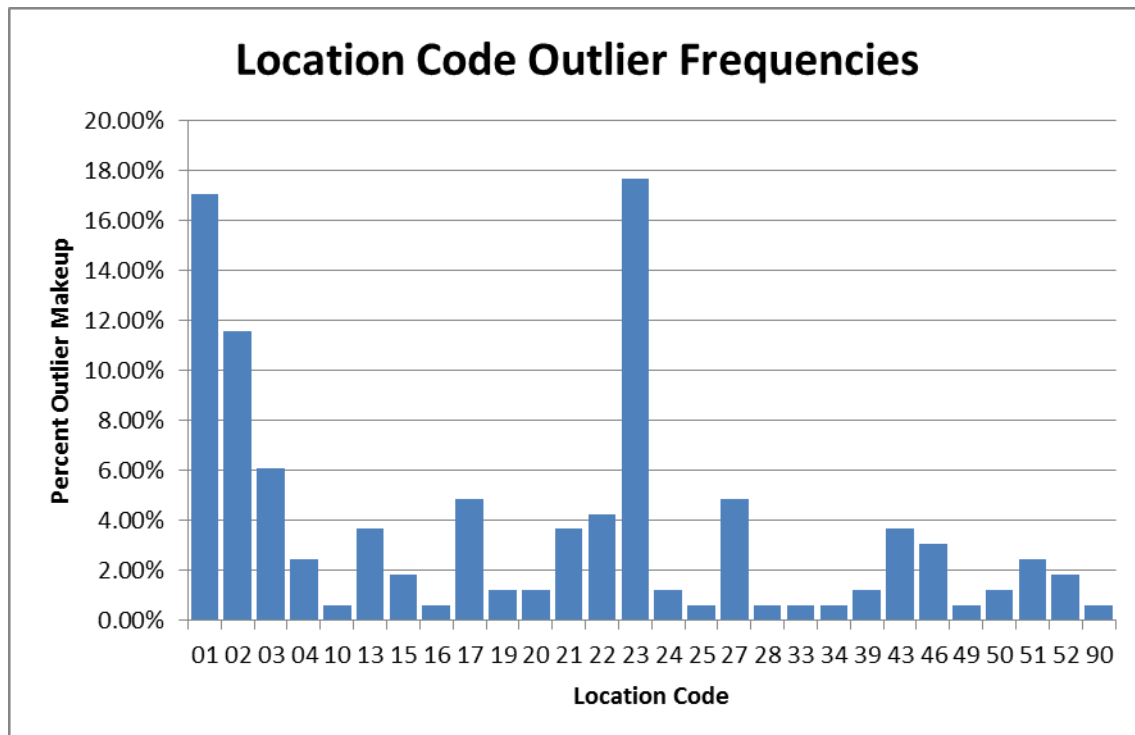


Figure 5: Location Outlier Frequencies

Figure 5 depicts what locations generate the most outliers within the current implemented methodology, but it fails to identify whether these outliers were generated due to the methodology failing to portray the cost behavior of the locations or if insufficient data was the root cause in creating the outlier. To gain more insight into this issue, the proportion of each outlier rate, as determined by their MNOR, generated by less than 30 and less than 10 observations respectively from the previous fiscal year were determined. The results of this analysis are shown in Table. The observations from Figure 5 are supported by Table 10 as indicated by the frequency column; the Arabian Gulf, the East Coast of the United States, the Gulf Coast of the United States, and the California Coast of the United States make up the majority of the outliers in the data. Table 10 indicates, however, that the majority of the outliers found were in large part due to the lack of data available to be utilized by the current methodology. For example, the Arabian Gulf made up roughly 35% of all the outliers, as determined by the MNOR. Of these observations, roughly 66% of them were determined from rates that were generated by using less than 30 observations from the previous fiscal year. Further, approximately 28% of the observations were determined from rates generated by less than 10 observations from the previous fiscal year.

Referencing Table 10, it can be seen that the majority of the outliers can be contributed to the small demand of shipments through each respective location. For the small percentage of outliers produced with larger than 30 observation from the previous fiscal year, it can adequately be said that the current methodology fails to capture the cost behavior of the location via the Central Limit theorem since the current methodology only utilizes the weighted average cost from the previous fiscal year. For the large

proportion of the observed outliers, it can only be said that their respective rates require special attention when setting rates, as the current methodology fails to capture the cost behavior of the location given the current sample sizes available. More data is required to adequately determine whether or not the current methodology accurately portrays the cost behavior of the location.

Table 11: Location Code Frequency Breakdown

Location	Location Code	Frequency	Percent of Outliers w/less than 30 Obs.	Percent of Outliers w/less than 10 Obs.
Arabian Gulf	23	17.68%	32.76%	13.79%
CONUS (East Coast)	01	17.07%	41.07%	28.57%
CONUS (Gulf Coast)	02	11.59%	44.74%	23.68%
CONUS (California Coast)	03	6.10%	25.00%	20.00%
Hawaiian Islands	27	4.88%	37.50%	18.75%
Northern Europe	17	4.88%	31.25%	6.25%
South and East Africa	22	4.27%	50.00%	28.57%
Black Sea	43	3.66%	33.33%	16.67%
West Africa	21	3.66%	50.00%	25.00%
Lesser Antilles Islands	13	3.66%	50.00%	33.33%
Azores	46	3.05%	50.00%	50.00%
CONUS (Northwest Coast)	04	2.44%	50.00%	50.00%
Korea	51	2.44%	37.50%	25.00%
Caribbean (Other)	15	1.83%	50.00%	33.33%
Japan	52	1.83%	16.67%	0.00%
Ryukyu Islands	50	1.22%	50.00%	25.00%
Mediterranean (West)	19	1.22%	50.00%	25.00%
Mediterranean (East)	20	1.22%	50.00%	50.00%
India and Burma	24	1.22%	50.00%	50.00%
South Pacific Islands	39	1.22%	50.00%	50.00%
Cuba (Guantanamo Bay)	16	0.61%	50.00%	50.00%
South East Asia (Other)	49	0.61%	50.00%	50.00%
Thailand	33	0.61%	50.00%	50.00%
New Guinea and Australia	34	0.61%	50.00%	50.00%
West Mexico and Central America	10	0.61%	50.00%	50.00%
Alaska (East)	25	0.61%	50.00%	50.00%
Marshall Islands	28	0.61%	50.00%	0.00%
Afghanistan via Riga, Talinn or Klaipeda linehaul	90	0.61%	50.00%	50.00%

Continuing with individual outlier analysis, the 82 individual outliers were then examined based on their respective shipping code designations attached at the end of their coding. Within this analysis, we were looking to identify particular shipping codes

that skew the MNOR, and to determine whether they were caused by a lack of demand or insufficient representation by the methodology. Looking at Figure 6, shipping codes 06 (General Breakbulk), 13 (General Container), 08 (Special Breakbulk), and 12 (Vehicle Containers) were determined to make up the majority of the outliers. Further, it should be noted that there were no outliers using shipping methods 01 (Reefer Breakbulk) and 02 (Bulk Breakbulk) indicating that the current methodology accurately depicts the cost behavior of these two shipping codes.

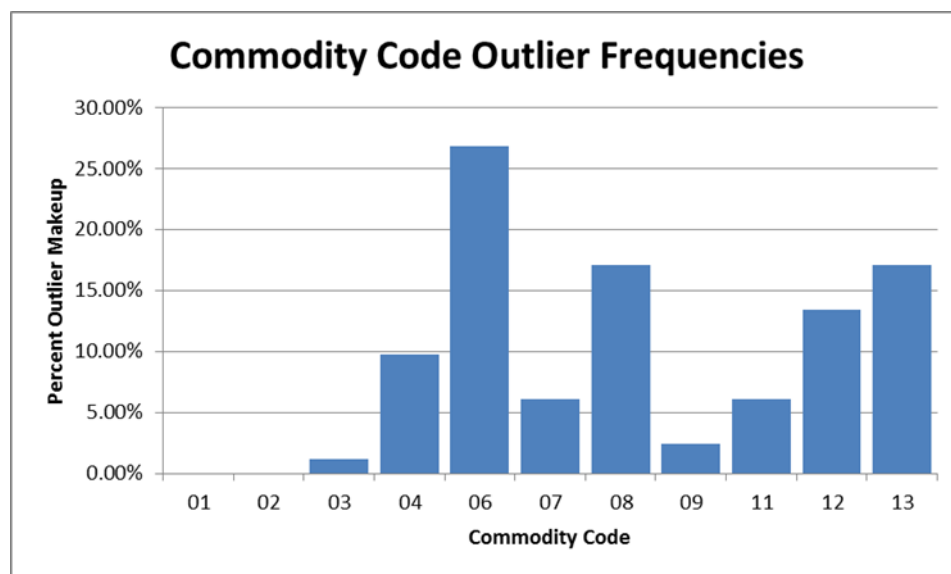


Figure 6: Shipping Code Outlier Frequencies

Referencing Table 11, it is shown that a large majority of the outliers regarding shipping code designation can be explained by the lack of demand from the previous fiscal year for each respective shipping code. General Breakbulk cargoes were the most frequent shipping code observed in the outliers; approximately 91% of the rates calculated that generated an outlier for this type of cargo were set using less than 30 observations from the previous fiscal year, and 50% were set with less than 10

observations. The only shipping code that seemed to be split between outliers that had sufficient and insufficient data to set rates was Hazardous Breakbulk, in which only 50% of its respective outliers were produced from rates that were generated from more than 30 observations from the previous fiscal year. Similar to the location code analysis, it cannot be said whether more data would cause the current methodology to accurately depict the cost behavior of these rates. The small demand for these particular shipping codes, however, cause the need for special attention to be directed toward these particular shipping codes in the rate setting process.

Table 12: Shipping Code Frequency Breakdown

Description	Commodity Code	Frequency	Percent of Outliers w/less than 30 Obs.	Percent of Outliers w/less than 10 Obs.
General Breakbulk	06	26.83%	90.91%	50.00%
Special Breakbulk	08	17.07%	85.71%	64.29%
General Container	13	17.07%	92.86%	57.14%
Vehicles Container	12	13.41%	54.55%	27.27%
Hazardous Breakbulk	04	9.76%	50.00%	25.00%
Trailers Breakbulk	07	6.10%	80.00%	20.00%
Reefer Container	11	6.10%	80.00%	80.00%
Aircraft Breakbulk	09	2.44%	100.00%	100.00%
POV Breakbulk	03	1.22%	100.00%	100.00%
Reefer Breakbulk	01	0.00%		
Bulk Breakbulk	02	0.00%		

2) Recurrent Outliers

Following the analysis of outliers broken down by their respective location and shipping codes, the behavior of recurrent outliers was examined in order to assess the current methodologies ability to self-correct with time. In order to perform this analysis, the methodology from Section 3.4 (B.1) was expanded upon in order to examine each respective recurrent outlier's MNOR for each fiscal year compared to the overall mean

for all marginal contributions to the NOR across fiscal years 2008-2014. This was accomplished by scaling each MNOR in terms of standard deviations away from the overall mean as described by Equation (11) and is modified by substituting each individual MNOR for each respective fiscal year for the average MNOR from each fiscal year. This is an almost identical practice to standardizing the residuals in regression analysis; instead in this particular situation, we substituted the known variance for each given fiscal year. The current outliers are listed in Table 12. For the purpose of this analysis, all of the years in which data was available for the given rates were analyzed in order to analyze the current methodology's ability to adapt to outliers.

Table 13: Recurrent Outliers over Fiscal Years (2008-2014)

Recurrent Outliers	
Origin-Destination/Destination- Origin Commodity Code Comb.	Fiscal Years Outliers were Observed
01 13 06	2012-2013
01 13 11	2012-2014
01 22 13	2009-2010
01 23 04	2011-2012
01 43 08	2010-2011
01 46 08	2011-2012
02 15 06	2012-2013
02 17 06	2010-2011
02 21 13	2008-2010
02 22 13	2008-2009
03 23 12	2008, 2011-2013
03 39 08	2013-2014
22 23 08	2010-2011

The respective outliers were examined in the same fashion depicted by Figure 3 in Section 3.4 (B.2.1), and are shown in Figures 7-11. In Figure 7, origin-destination/destination-origin and commodity codes combinations 01 13 06, 01 13 11,

and 01 22 13 were analyzed. The outlier 01 13 06 appears to be corrected for after oscillating between a negative and positive outlier across fiscal years 2012 and 2013. The outlier 01 13 11 appears to have failed to have been corrected for as it remained an outlier after year 2012. The outlier 01 22 13 seems to have been corrected for after it oscillated between negative and positive outliers between years 2009 and 2010, and it then remained within the bounds of two standard deviations throughout the remainder of the study.

Examining Figure 8, we observe oscillations between negative and positive marginal contributions to the NOR for all three outliers examined. Outlier 01 43 08 appears to oscillate between a negative and positive marginal NOR, but whether it is corrected for cannot be determined, as this rate was not observed again throughout the course of the data set.

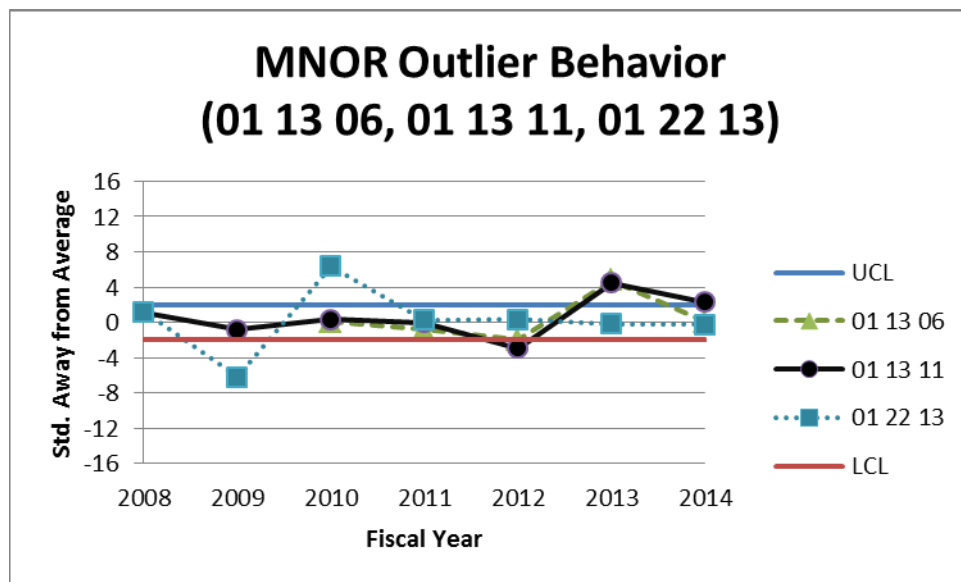


Figure 7: 01 13 06, 01 13 11, 01 22 13 Outlier Behavior

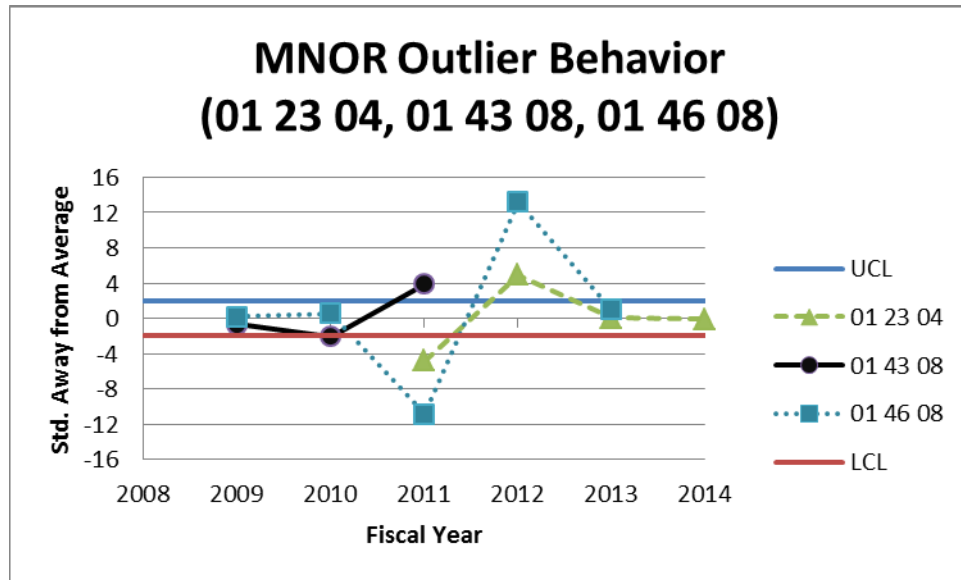


Figure 8: 01 23 04, 01 43, 08, 01 46 08 Outlier Behavior

Figures 9 and 10 display outlier behavior that is oscillating between negative and positive marginal contributions to the NOR before showing indication of being self-corrected as indicated by the smoothing of their respective lines moving forward. Figure 10 offered our first repeating positive marginal contributions to the NOR in outlier 02 22 13. Despite being corrected for in its final two years of observations, this offered a different pattern which can potentially be explained by the two-year running NOR currently implemented by USTC J8 which could have prevented the rate 02 22 13 from changing in a manner to correct after one year. Figure 10 also offers one similar behavior which could be taken as out-of-the-norm due to the lack of observations across the study rate 03 39 08 exhibits a negative to positive change across its two years of observation. Further, rate 03 23 12 is not observed in fiscal year 2010, but it appears to remain negative or close to a negative in its first few years of observation before oscillating and leveling out. The two outliers on Figure 11 appear to follow the common oscillating

trend, but any real insights are difficult to gain due to the lack of observations throughout the study.

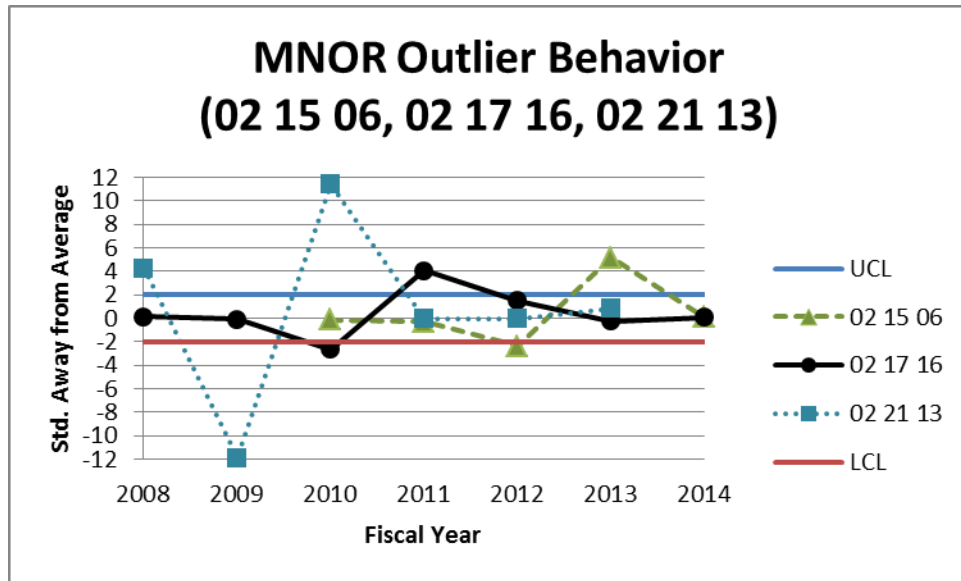


Figure 9: 02 15 06, 02 17, 16, 02 21 13 Outlier Behavior

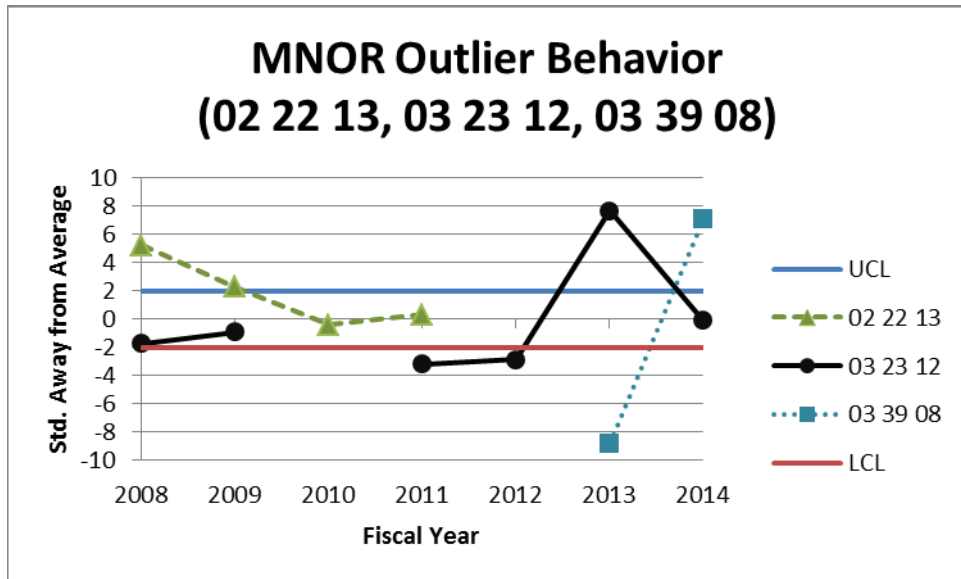


Figure 10: 02 22 13, 03 23 12, 03 39 08 Outlier Behavior

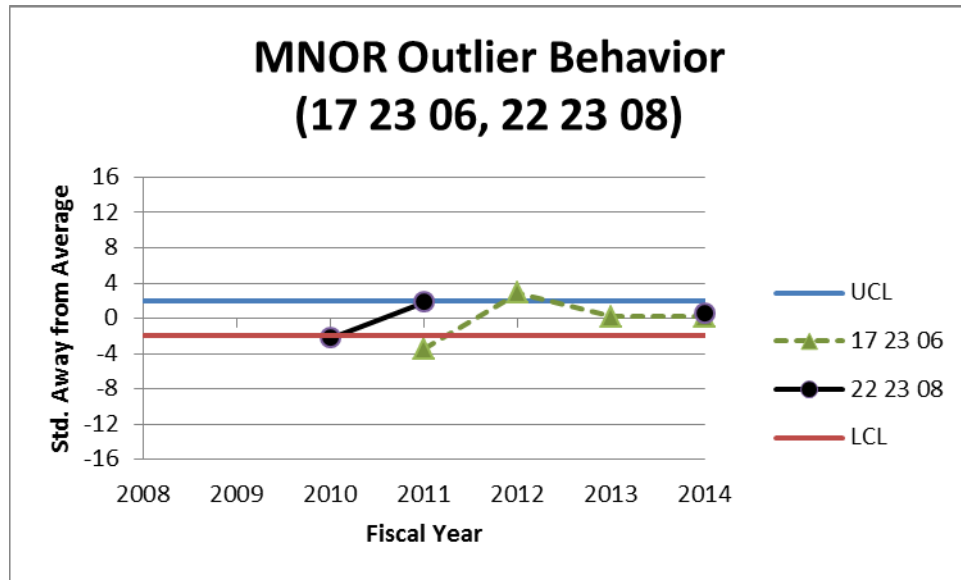


Figure 11: 17 23 06, 22 23 08 Behavior Outlier

The outlier behavior graphs all have a common theme of oscillating between negative and positive marginal contributions to the NOR outliers between their repeating outlier years, with the exception of those outlier rates detailed above. The behavior observed is one in which the methodology appears to over correct itself. The oscillation between negative and positive outliers can be explained by the running NOR implemented by USTC J8 and the use of the AOR factor to reach the two year goal of having a NOR of zero. Given the cost data of a particular rate's sensitivity to the current methodology, it makes sense that the adjustment in the AOR factor made to correct for the negative deficit would cause a positive outlier in the next year due to the rates sensitivity to the methodology. This is exactly what was observed in the majority of the repeating outliers identified over the course of this study, indicating that the current methodology of making up for losses and gains before returning to normal behavior is working in an efficient manner.

4.4.3 The Composite Method Results

Upon completion of outlier analysis, Section 3.4 (A) was carried out by analyzing the desired Composite AOR Factor with the Composite AOR Factor implemented. The absolute value of the difference between the two factors for each respective fiscal year was taken in order to examine the magnitude of the error. A generalized hypothesis test was set up in excel to examine the absolute value of the differences. The null hypothesis was the mean absolute difference was greater than or equal to some variable, X , and the alternative hypothesis was that the mean was less than X . The variant X was then controlled until we were indifferent between failing to reject and rejecting the null hypothesis (e.g., a *p-value* equal to 0.05). As shown in Table 13, we were indifferent between accepting the null hypothesis that mean was greater than or equal to 26.86%, and the mean was less than 26.86%. Therefore, any value less than 26.86% would fail to reject the null, and any value above would reject the null hypothesis at the 5% significance level.

Any value less than 26.86% failed to reject the null hypothesis at the 5% significance level. Therefore, for our data set, it can be justifiably be said that on average the AOR factor and the Composite Rate Adjustment contributed to approximately 25% of the error between the rate set and the actual cost for each fiscal year across 2008-2014, provided that the refresh rates provided were an accurate depiction of the inflation within the environment of each respective fiscal year.

Table 14: Composite Method Hypothesis Test

Year	Composite AOR Rate Implemented	Desired Composite AOR Rate w/Outliers	Abs. Diff. Implemented & Rate w/Outliers	Desired Composite AOR Rate w/o Outliers	ABS. Diff. Implemented & Rate w/o Outliers
2008	25.69%	4.19%	21.50%	13.59%	12.10%
2009	4.25%	44.15%	39.90%	3.12%	1.13%
2010	18.03%	-0.57%	18.60%	1.29%	16.74%
2011	8.93%	36.31%	27.38%	36.65%	27.72%
2012	16.86%	-14.03%	30.89%	-12.90%	29.76%
2013	0.33%	-28.54%	28.87%	-28.19%	28.52%
2014	15.33%	-2.67%	18.00%	-3.05%	18.38%

Average Difference:	26.45%	Average Difference:	19.19%
Standard Deviation:	7.80%	Standard Deviation:	10.45%
Standard Error:	2.95%	Standard Error:	3.95%
t-value:	-0.14	t-value:	-1.94
Df:	6	Df:	6
P-Value:	0.447	P-Value:	0.050

Hypothesis Threshold:	26.86%
-----------------------	--------

4.5 Summary

This chapter presented the results observed when the methodology described in Chapter 3 was implemented. Beginning with Section 4.2 the current process was evaluated, and it was found that utilizing a full year of data in the rate setting process did not improve the methodology. Following Section 4.2, the predicted rates calculated from the prescribed methodology were compared to the actual rates set. There was only enough data readily available to reproduce approximately 5% of the total number of rates set for fiscal years 2014-2016. Within this 5%, roughly 8-10% of the calculated rates had a large enough percent deviation from the actual rate set to conclude that they were produced from a different methodology than the one documented and utilized in this research. Section 4.4.1 showed that the current methodology performs with consistent behavior over the time frame, 2008-2014, of the study. Section 4.4.2 breaks down outliers on an individual basis. It is shown that the lack of data in the rate setting process plays a contributing role in producing outliers in regards to the MNOR. Also, a consistent

behavior of repeating outliers oscillating between negative and positive outlier MNORs before self-correcting was observed, which is expected given the current methodology. Finally, Section 4.4.3 shows the results of what the desired Composite AOR factor should be, provided that the assumption that the given refresh rate is an accurate estimation of the inflation of the cost environment in which the rates are being set for holds. It is shown at a 5% significance level that on average the Composite AOR factor implemented is 25% off. The results shown in this chapter support what was believed to be observed due to the make-up of the current methodology as well as provide insight into areas in which individual attention should be placed in order to reduce the magnitude of the total NOR for each fiscal year.

V. Conclusions and Recommendations

5.1 Overview

This chapter summarizes the results and analyses vis-à-vis the objectives as presented in Section 1.3 of this research. The resulting compilation is sequentially presented to address conclusions, significance of the research, recommendations for action, and recommendations for future research.

5.2 Conclusions

Throughout the course of this research many conclusions were drawn from the provided data. The comparison of the use of six months of data versus a full year of data was examined in order to determine whether expanding the data set utilized provided a significant decrease in the resulting NORs. The extent to which the methodology was deviated from over the course of the rate setting process was analyzed, and areas in which strict utilization of the current methodology failed to adequately describe the cost environment in which it was acting were examined in terms of outliers.

The first component of this research analyzed the current methodology of utilizing six months of data versus a full year of data as they relate to the successive years' NORs. The purpose of this endeavor was to examine the impact calculating the weighted average cost for each origin-destination/destination-origin and commodity code combination from a larger data sample had on the resulting total NORs. A statistically significant decrease in the total NOR would have provided leverage for policy change with regard to the data size utilized in the current methodology. In fact, at the 5% significance level the utilization of a larger data sample did not have a significant impact

on the resulting total NORs. These results do not support the hypothesis that utilizing only six months of cost data in the rate setting process has a negative impact on rate accuracy.

The next component of this research identified the frequency with which the current rate setting methodology was likely not used over fiscal years 2014-2016, presumably in favor of a budget analyst's experience-informed judgement. This research endeavor supported the initial assumption that the current methodology was not strictly utilized in setting all the rates in the rate setting process. Although we were only able to duplicate approximately 5% of the rates set across fiscal years 2014-2016, we found that respectively 8%, 10%, and 4% of the rates set via the current methodology were likely set using a differing technique since the outliers possessed a relative deviation greater than or equal to 100%.

The third component of this research assessed the effectiveness of the current methodology to set rates as measured by the marginal net operating results (MNORs) of the various combinations of origin-destination/destination-origin and commodity code combinations, both in the aggregate and with respect to outliers. The purpose of an aggregate examination was to determine whether the rate setting methodology is working as a process, whereas the purpose of the outlier analysis was to help identify (a) whether the methodology works with respect to problematic combinations and (b) to determine rates for which the current methodology should be deviated from. The behavior of the current methodology over the course of fiscal years 2008-2014 was determined to be consistent. In other words, the current methodology did not affect notable deviations or patterns of deviations in terms of the average MNOR over the course of the study. Of

note, the data also showed that a large proportion of the outliers existed for rates set by data which failed to meet the Central Limit Theorem's recommended sample size of thirty observations. Over fiscal years 2008-2014, 80.49% of the outliers were set using less than thirty observations; however, of all rates set with less than thirty observations only 5.16% were deemed outliers. Thus, it is suggested that a small sample size only affects the accuracy of a small proportion of the rates set.

The final research trail component analyzed the impact of the combination of the AOR Factor and the Composite Rate Adjustment on the resulting NORs over fiscal years 2008-2014. This research was conducted under the assumption that the Refresh Rate was an accurate representation of the inflation within the cost environment; thus, the Refresh Rates provided were utilized and not considered to contribute to error induced by the current methodology. We found that, in combination, the AOR Factor and the Composite Rate Adjustment on average were off 25% provided a resulting total NOR of \$0 was desired over the course of this research.

In summary, this research shows that the current utilization of six months of data in the rate setting process is adequate. The current rate-setting methodology has not been historically utilized for all rates set. The behavior of the methodology is concluded to be consistent over time and small data samples tends to cause outliers, but not all small data samples cause outliers. Finally, the data concludes that a significant portion of error currently is induced as a result of the combination of the AOR factor and the Composite Rate Adjustment utilized each fiscal year.

5.2 Research Significance

This research does not propose new process or overhaul an existing process. Instead, it identifies areas in which the current rate-setting methodology implemented by USTC J8 could be improved upon with additional research, areas in which the current methodology has not been utilized, and areas in which the resulting total NORs could be improved upon by deviating from the currently implemented methodology.

The evaluation of the current practice concluded that introducing an additional six months of data in the implementation of the current practice would not improve the magnitude of the resulting total NOR. This conclusion is significant because it indicates that the utilization of an additional six months of cost data will not improve the current rate-setting methodology's accuracy in terms total NORs, which eliminates an area of focus for USTC J8 moving forward.

The outliers found when analyzing the rates set utilizing the current rate-setting methodology to the actual rates set for fiscal years 2014-2016 as well as the individual outliers analyzed provide the customer, USTC J8, provide insight into where the current methodology has been deviated from in the rate setting process. The sample data utilized within this research restricted the results of this research; however, the data allowed for a portion of the deviations and individual outliers to be identified as well as allowed for a process of identification of outliers to be explained and demonstrated via practice.

The results of the Composite Method were dependent upon the assumption that the Refresh Rate was not a contributing source of error in the current methodology. It would be difficult to make a case that this assumption would strictly hold, as the Refresh Rate is an estimation of the inflation in the transportation realm in which the research was

conducted. It is much more likely that the error comes from a compounding effect of combining three estimated adjustment parameters with their own respective errors into one error which induces a larger variance. Despite the potential shortcomings with this assumption, the results from this analysis still conclude that the combination of the AOR factor and the Composite Rate Adjustment produce an error sufficiently large to warrant additional investigation.

5.3 Recommendations for Action/Future Research

We recommend USTC J8 continue utilizing six months of cost data in the rate setting process. Further we recommend utilizing the methods demonstrated of outlier analysis in order to gain insight into where additional attention, other than that given within the current methodology, should be applied to properly set rates to fit the cost environment in which they are acting.

Based off the findings in the Composite Rate analysis, we recommend additional research into the processes of respectively setting the Refresh Rate, AOR Factor, and Composite Rate Adjustment. We understand that each of these rates serves a purpose within the current methodology, but we believe that the error found in the resulting total NORs could be greatly reduced by reducing the combined variability induced by these respective variables. We propose that goal programming, stochastic programming, and/or robust optimization might be worth investigating as a method to (a) minimize the NOR, (b) minimize the variance induced by the three rate adjustment factors, and/or (c) minimize the NOR over a range of uncertain outcomes with regard to both forecasted

demand over the set of origin-destination/destination-origin and commodity code combinations as well as inflationary outcomes.

5.4 Summary

In summary, Chapter 5 formulates the results shown in Chapter 4 of this research into words and recommends actions that can be utilized by the decision makers at USTC J8. Beginning with the first component of the research, six months of data was concluded to be an adequate amount of data utilized in comparison with a full-year worth of data. Further, a process to determine outliers, both in terms of deviations from the methodology taken in previous years as well as rates that should be addressed separately, was demonstrated with the data provided for this research giving USTC J8 a method to better set rates. Finally, this research was finished by concluding that the AOR Factor and Composite Rate Adjustment are significant contributors to the magnitude of the resulting total NORs each year.

Appendix A: Script to compute the Weighted Average Costs via Equation (1)

```
'This code was implmented for both six months of data and a full year of data.
'The six month weighted averages were calculated from a data set with all
'observations occuring after June 31st removed. The full year weighted averages
'were calculated using the full data set from each fiscal year.
Sub Average_Builder()

'This code dimensionalizes the variables utilized in this code.
Dim wb1 As Workbook
Dim wb2 As Workbook
Dim DataArray(1 To 264, 1 To 2) As Variant
Dim i As Variant
Dim X As Variant
Dim j As Variant
Dim Obs As Variant
Dim WeightSum As Variant
Dim TotalCost As Variant
Dim AvgW As Variant
Dim Y As Variant

'Turns all alerts off in excel so that a sheet can be deleted without Excel
'automatically asking the user if they are sure they want to delete the page.
'Without turning alerts off the code would stop after each loop through the for
'loop as a sheet is added and deleted within the for loop for data manipulation
'purposes.
Application.DisplayAlerts = False

'Sets the open workbook as wb1
Set wb1 = ActiveWorkbook

'This code loops through all of the origin-destination/destination-origin and
'commodity code combinations and saves them into the DataArray. The for loop
'ranges from 1 to 264 because for the particular year this code was implemented
'data was available to set rates based on 264 origin-destination/destination-
'origin and commodity code combinations.
For i = 1 To 264
DataArray(i, 1) = wb1.Sheets("Sheet1").Range("A" & i + 1)
DataArray(i, 2) = wb1.Sheets("Sheet1").Range("B" & i + 1)
Next i

'Prompts the user to open the cost data file in which they want to calculate the
'average costs from.
FileToOpen = Application.GetOpenFilename _
(Title:="Please choose a Report to Parse", _
FileFilter:="Report Files *.xlsx (*.xlsx),")

'If the user fails to identify a file to open this code will notify the user of
'their error and will exit the sub.
If FileToOpen = False Then
MsgBox "No File Specified.", vbExclamation, "ERROR"
Exit Sub
Else

'Sets the user specified workbook as wb2
Set wb2 = Workbooks.Open(Filename:=FileToOpen)
wb2.Activate

'Gives the variable X the value of the length of column F. Column F holds the
'origin-destination commodity code combinations for each individual data point.
X = Range("F" & Rows.Count).End(xlUp).Row
```

```

'i is looped through 264 times because there were 264 origin-destination/destination
'-origin and commodity code combinations in the given year the average was calculated
'for.
For i = 1 To 264

'Sets the total weight shipped for each origin-destination/destination-origin and commodity
'code combination equal to zero.
WeightSum = 0

'Sets the total cost for each origin-destination/destination-origin and commodity code
'combination equal to zero.
TotalCost = 0

'Selects all the data in the cost data worksheet for filtering. The sheet "FY08" has to be
'changed for each year (e.g., Fiscal Year 2009 will be indicated by "FY09" and so forth).
wb2.Sheets("FY08").Range("A1:O" & X).Select

'Initiates the filter command within Microsoft Excel.
Selection.AutoFilter

'Filters the data to display the origin-destination/destination-origin and commodity code
'pair indexed by (i,1) & (i,2). If the column containing the origin-destination/destination-
'origin and commodity code combinations is not in column F, the AutoFilter:=6 will need to be adjusted
'to match the corresponding column. (e.g., if the data was in column G the code would read
'"AutoFilter Field :=7").
ActiveSheet.Range("$A$1:$O" & X).AutoFilter Field:=6, Criteria1:="=" & DataArray(i, 1) _
, Operator:=xlOr, Criteria2:="=" & DataArray(i, 2)

'Copies the filtered data and paste into a new worksheet within Excel to enable the required
'calculations to be performed.
Selection.Copy
Sheets.Add After:=ActiveSheet
ActiveSheet.Paste

'This Y variable is assigned the number of data points copied over into the new worksheet.
'Y will always be at least one since the headers are always copied.
Y = Range("F" & Rows.Count).End(xlUp).Row

'This if statement checks to see if data is available for the origin-destination/destination-origin
'and commodity code combination indexed by (i,1) & (i,2).
If Y > 1 Then

'If data is available this for loop then loops through all the observations for the origin-destination
'/destination-origin and commodity code combination indexed by (i,1) & (i,2).
For j = 2 To Y

'This WeightSum takes the sum of the total weight shipped via the specific code and adds the next
'observation's weight shipped.
WeightSum = WeightSum + ActiveSheet.Range("B" & j)

'This TotalCost takes the total cost of shipments across the specific code and adds the next
'observation's total cost.
TotalCost = TotalCost + ActiveSheet.Range("G" & j)

Next j

'Sets the number of observations for the specific origin-destination/destination-origin and
'commodity code combination equal to the total length of the copied and pasted data minus one such
'that the headers are not accounted for.
Obs = Y - 1

```

```

'Sets the average weighted cost to the sum of the total cost divided by the total weight shipped via
'the specific origin-destination/destination-origin and commodity code combination.
    AvgW = TotalCost / WeightSum

Else

'If no observations are present, Y=1, then the number of observations and the average weighted cost
'are set to zero.
    Obs = 0
    AvgW = 0
End If

'Activates wb1, where all the calculated values are inputted.
    wb1.Activate

'Writes the weighted average cost (AvgW) and the number of observations (Obs) to the respective columns
'C and D in the corresponding row as their origin-destination/destination-origin and commodity code combination.
'These columns were adjusted to E and F when the full year data set was run.
    wb1.Sheets("Sheet1").Range("C" & i + 1) = AvgW
    wb1.Sheets("Sheet1").Range("D" & i + 1) = Obs

'Takes the sheet that the data was copied into and deletes it to reduce the amount of memory used.
    wb2.Activate
    ActiveSheet.Delete

'Reactivates the worksheet, and clears the filter just applied. The sheet "FY08" has to be
'changed for each year (e.g., Fiscal Year 2009 will be indicated by "FY09" and so forth).
    wb2.Activate
    wb2.Sheets("FY08").Select
    Selection.AutoFilter
Next i

End If

    wb2.Close

'Turns the Display Alerts back on within Excel.
Application.DisplayAlerts = True

'Saves the updated workbook represented by wb1 and closes the workbook.
wb1.Close SaveChanges:=True

End Sub

```

Appendix B: A Script to Compute NORs via Equation (3)

```
'This code was implemented to determine the resulting NORs using six months and a full year
'of cost data in the rate setting process. The weighted average cost were determined using
'the Average_Builder() code. These averages were then multiplied by the provided Refresh Rate,
'AOR Factor, and Composite Rate Adjustment provided to set the rate.
Sub NOR_Calculator()

'The following code dimensionalizes the variables utilized within the code.
Dim wb1 As Workbook
Dim wb2 As Workbook
Dim DataArray(1 To 264, 1 To 4)
Dim i As Variant
Dim X As Variant
Dim j As Variant
Dim NOR6W As Variant
Dim NOR1W As Variant

'Turns all display alerts off within Excel in order to prevent the code from stopping during each
'iteration when a sheet is deleted to preserve memory space.
Application.DisplayAlerts = False

'Sets the workbook that is currently open as wb1.
Set wb1 = ActiveWorkbook

'This for loop builds the DataArray that holds each origin-destination/destination-origin and
'commodity code combination and rates produced from six months and a full year of data
'respectively. The active workbook, wb1, contains the origin-destination and commodity code
'combination in column A, the destination-origin and commodity code combination in column B,
'the rate set using six months of data in column G, and the rate set using a full year of data
'in column H. All hard coded for loops and the dimensions of the DataArray must be adjusted to
'match the number of different origin-destination/destination-origin and commodity code combinations
'available in each data set.
For i = 1 To 264
DataArray(i, 1) = wb1.Sheets("Sheet1").Range("A" & i + 1)
DataArray(i, 2) = wb1.Sheets("Sheet1").Range("B" & i + 1)
DataArray(i, 3) = wb1.Sheets("Sheet1").Range("G" & i + 1)
DataArray(i, 4) = wb1.Sheets("Sheet1").Range("H" & i + 1)
Next i

'Prompts the user to open the desired data to calculate the NORs from. This particular code used
'rates generated from fiscal year 2008 data; therefore, the user should grab the full year of cost
'data from fiscal year 2009 to generate NORs.
FileToOpen = Application.GetOpenFilename _
(Title:="Please choose a Report to Parse", _
FileFilter:="Report Files *.xlsx (*.xlsx),")

'Notifies the user of an error if a file is not properly selected and exits the code.
If FileToOpen = False Then
MsgBox "No File Specified.", vbExclamation, "ERROR"
Exit Sub
Else

'Sets wb2 to the workbook specified by the user.
Set wb2 = Workbooks.Open(Filename:=FileToOpen)
wb2.Activate

'Sets X to the length of the column containing all origin-destination/destination-origin and commodity
'code combinations. This row was used to count the length of the data as it was one column that was
'consistently populated with data throughout.
X = Range("F" & Rows.Count).End(xlUp).Row
```

```

'This for loop is utilized to loop through all 264 origin-destination/destination-origin and commodity
'pairs.
    For i = 1 To 264

'This code selects all of the available data in the sheet FY09 and prepares it for filtering. The sheet
' "FY09" must be updated for each fiscal year (e.g., if you are analyzing the performance of fiscal year
'2010 rates you must update the sheet to "FY10").
        wb2.Sheets("FY09").Range("A1:N" & X).Select
        Selection.AutoFilter

'Filters the data in wb2 according to field 6, which holds the origin-destination/destination-origin and
'commodity code combinations, by the two criterium DataArray(i,1) and DataArray(i,2), which hold the respective
'origin-destination and destination-origin and commodity code combinations.
        ActiveSheet.Range("$A$1:$N" & X).AutoFilter Field:=6, Criteria1:="" & DataArray(i, 1) _
            , Operator:=xlOr, Criteria2:="" & DataArray(i, 2)

'Selects and copies the filtered data and paste the data into a new sheet within wb2.
        Selection.Copy
        Sheets.Add After:=ActiveSheet
        ActiveSheet.Paste

'Sets Y to the length of the row containing all the origin-destination/destination-origin and commodity
'code combinations.
        Y = Range("F" & Rows.Count).End(xlUp).Row

'Sets the total NORs for each origin-destination/destination-origin and commodity code combination equal to zero at
'beginning of its respective analysis for both its rate calculated utilizing six months of data and its rate
'calculated using a full year of data.
        NOR6W = 0
        NOR1W = 0

'If statement to check and see if data is observed for the particular origin-destination/destination-origin
'and commodity code combination. Y will always be at least one as the headers are always copied.
        If Y > 1 Then

'This for loop loops through all the observations for a given origin-destination/destination-origin and
'commodity code combination and calculates a given observation's NOR contribution and adds it to the total
'NOR for a given origin-destination/destination-origin and commodity code combination. Column B contains the total
'weight of a shipment. This is multiplied by the rate set, (i,3) represents the rate set using six months of data
'and (i,4) represents the rate set using a full year of data, and the total cost, contained in column G, for a
'shipment is then subtracted out to form the NOR for a given observation.
            For j = 2 To Y
                NOR6W = NOR6W + ((ActiveSheet.Range("B" & j) * DataArray(i, 3)) - ActiveSheet.Range("G" & j))
                NOR1W = NOR1W + ((ActiveSheet.Range("B" & j) * DataArray(i, 4)) - ActiveSheet.Range("G" & j))
            Next j

'Writes the resulting NORs from the rate set with six months of data and the rate set with a full year of data
'for each respective origin-destination/destination-origin and commodity code combination to columns I and J
'of the corresponding row.
            wb1.Sheets("Sheet1").Range("I" & i + 1) = NOR6W
            wb1.Sheets("Sheet1").Range("J" & i + 1) = NOR1W

'Activates wb2 and deletes the sheet in which the data was copied to in order to reduce the required memory\
'to run this code.
            wb2.Activate
            ActiveSheet.Delete

'This code clears the filter just performed on all the cost data.
            wb2.Activate
            wb2.Sheets("FY09").Select
            Selection.AutoFilter

        End If

    Next i

End If

'Closes wb2 and does not save any changes performed on your cost data. This prevents an error in the code from
'manipulating your data and causing future errors.
wb2.Close

'Turns back on display alerts within Microsoft Excel.
Application.DisplayAlerts = True

'Saves the updates to wb1 and closes the workbook.
wb1.Close SaveChanges:=True

End Sub

```

Appendix C: A Script to Compute MNORs via Equation (7)

```
'This code calculates the marginal contribution each measurement ton had on its respective NOR.
Sub MNOR_Calculator()
Dim wb1 As Workbook
Dim wb2 As Workbook
Dim DataArray(1 To 264, 1 To 3)
Dim i As Variant
Dim X As Variant
Dim j As Variant
Dim Weight As Variant

'Turns the display alerts off in Microsoft Excel.
Application.DisplayAlerts = False

'Sets the active workbook to wb1
Set wb1 = ActiveWorkbook

'Builds an array that stores each origin-destination/destination-origin and commodity code combination
'and its respective NOR as found in previous code.
For i = 1 To 264
DataArray(i, 1) = wb1.Sheets("Sheet1").Range("A" & i + 1)
DataArray(i, 2) = wb1.Sheets("Sheet1").Range("B" & i + 1)
DataArray(i, 3) = wb1.Sheets("Sheet1").Range("I" & i + 1)
Next i

'Gets file the user wants to find the MNOR from.
FileToOpen = Application.GetOpenFilename _
(Title:="Please choose a Report to Parse", _
FileFilter:="Report Files *.xlsx (*.xlsx),")

'Checks to see if the user properly selected a file. Displays an error and exits the sub if not.
If FileToOpen = False Then
MsgBox "No File Specified.", vbExclamation, "ERROR"
Exit Sub
Else

'Sets wb2 as the user selected file.
Set wb2 = Workbooks.Open(Filename:=FileToOpen)
wb2.Activate

'Sets X to the length of the column containing all origin-destination/destination-origin and commodity
'code combinations. This row was used to count the length of the data as it was one column that was
'consistently populated with data throughout.
X = Range("F" & Rows.Count).End(xlUp).Row

'This for loop is utilized to loop through all 264 origin-destination/destination-origin and commodity
'code combinations.
For i = 1 To 264

'This code selects all of the available data in the sheet FY09 and prepares it for filtering. The sheet
' "FY09" must be updated for each fiscal year (e.g., if you are analyzing the performance of fiscal year
'2010 rates you must update the sheet to "FY10").
wb2.Sheets("FY09").Range("A1:N" & X).Select
Selection.AutoFilter

'Filters the data in wb2 according to field 6, which holds the origin-destination/destination-origin and
'commodity code combinations, by the two criterium DataArray(i,1) and DataArray(i,2), which hold the respective
'origin-destination and destination-origin and commodity code combinations.
ActiveSheet.Range("$A$1:$N" & X).AutoFilter Field:=6, Criteria1:="=" & DataArray(i, 1) _
, Operator:=xlOr, Criteria2:="=" & DataArray(i, 2)
```

```

'Selects and copies the filtered data and paste the data into a new sheet within wb2.
    Selection.Copy
    Sheets.Add After:=ActiveSheet
    ActiveSheet.Paste

'Sets Y to the length of the row containing all the origin-destination/destination-origin and commodity
'code combinations.
    Y = Range("F" & Rows.Count).End(xlUp).Row

'Sets the total weight shipped for a given origin-destination/destination-origin and commodity code combination
'equal to zero.
    Weight = 0

'Checks to see if the provided origin-destination/destination-origin and commodity code combination is observed.
'If Y > 1 then at least one observation is recorded.
    If Y > 1 Then

'This for loop runs through all the observations for a specific origin-destination/destination-origin and
'commodity code combination.
        For j = 2 To Y

'Takes the weight shipped for a specific observation and adds it to the total weight shipped via the origin
'-destination/destination-origin and commodity code combination being analyzed.
            Weight = Weight + ActiveSheet.Range("B" & j)

        Next j

'Sets the marginal contribution to the NOR equal to the total NOR for a specific origin-destination/destination
'-origin and commodity code combination divided by the total weight shipped via that combination.
        MNOR6W = DataArray(i, 3) / Weight

'Takes the marginal contribution to the NOR and writes it into column K in the corresponding row to the respective
'origin-destination/destination-origin and commodity code combination.
        wb1.Sheets("Sheet1").Range("K" & i + 1) = MNOR6W

'Deletes the sheet that the filtered data was copied into to preserve memory space.
        wb2.Activate
        ActiveSheet.Delete

'Erases the filter applied to the data.
        wb2.Activate
        wb2.Sheets("FY09").Select
        Selection.AutoFilter

    Else

'Deletes the sheet the data was copied into and erases the filter if there were no observations present.
        wb2.Activate
        ActiveSheet.Delete

        wb2.Activate
        wb2.Sheets("FY09").Select
        Selection.AutoFilter

    End If

Next i

End If

'Closes wb2 and does not save any changes that are applied to the workbook
wb2.Close

'Turns Excel's display alerts back on
Application.DisplayAlerts = True

'Saves the changes to the originally opened workbook and closes the workbook.
wb1.Close SaveChanges:=True

End Sub

```

Appendix D: A Script to Compute the Composite AOR Factor via Equation (6)

```
'This code calculates the desired combination of the AOR factor and Composite Rate Adjustment which
'would yield a total NOR equal to zero.
Sub Combined_AOR_Calculator()

'The following code dimensionalizes the variables utilized within the code.
Dim wb1 As Workbook
Dim wb2 As Workbook
Dim aRR As Variant
Dim aAOR As Variant
Dim TCNY As Variant
Dim TRNY As Variant
Dim X As Variant
Dim DataArray(1 To 264, 1 To 3) As Variant
Dim i As Variant
Dim j As Variant

'This is the Refresh Rate, it is hard coded and need to be changed for each fiscal year's analysis.
aRR = 0.05

'Turns Excel's display alerts off which prevent the code from being stopped by a save request message
'when sheets are deleted within the code.
Application.DisplayAlerts = False

'Sets wb1 to the open workbook.
Set wb1 = ActiveWorkbook

'This for loop loops through each origin-destination/destination-origin and commodity code combination
'and grabs its two respective combination codes and its associated average weighted cost.
For i = 1 To 264
    DataArray(i, 1) = wb1.Sheets("Sheet1").Range("A" & i + 1)
    DataArray(i, 2) = wb1.Sheets("Sheet1").Range("B" & i + 1)
    DataArray(i, 3) = wb1.Sheets("Sheet1").Range("C" & i + 1)
Next i

'Gets the file to evaluate from the user. If the user if using the average weighted cost for fiscal year
'2008 to analyze then the cost data for fiscaly year 2009 should be opened.
FileToOpen = Application.GetOpenFilename _
    (Title:="Please choose a Report to Parse", _
    FileFilter:="Report Files *.xlsx (*.xlsx),")

'Notifies the user if they make an error selecting the file and exits the code.
If FileToOpen = False Then
    MsgBox "No File Specified.", vbExclamation, "ERROR"
    Exit Sub
Else

'Sets wb2 to the one opened by the user
    Set wb2 = Workbooks.Open(Filename:=FileToOpen)
    wb2.Activate

'Sets X to the length of the column containing all origin-destination/destination-origin and commodity
'code combinations. This row was used to count the length of the data as it was one column that was
'consistently populated with data throughout.
    X = Range("F" & Rows.Count).End(xlUp).Row

'This for loop is utilized to loop through all 264 origin-destination/destination-origin and commodity
'code combinations.
    For i = 1 To 264
```

```

'This code selects all of the available data in the sheet FY09 and prepares it for filtering. The sheet
' "FY09" must be updated for each fiscal year (e.g., if you are analyzing the performance of fiscal year
'2010 rates you must update the sheet to "FY10").
wb2.Sheets("FY09").Range("A1:N" & X).Select
Selection.AutoFilter

'Filters the data in wb2 according to field 6, which holds the origin-destination/destination-origin and
'commodity code combinations, by the two criterium DataArray(i,1) and DataArray(i,2), which hold the respective
'origin-destination and destination-origin and commodity code combinations.
ActiveSheet.Range("$A$1:$N" & X).AutoFilter Field:=6, Criteria1:="" & DataArray(i, 1) _
, Operator:=xlOr, Criteria2:="" & DataArray(i, 2)

'Selects and copies the filtered data and paste the data into a new sheet within wb2.
Selection.Copy
Sheets.Add After:=ActiveSheet
ActiveSheet.Paste

'Sets Y to the length of the row containing all the origin-destination/destination-origin and commodity
'code combinations.
Y = Range("F" & Rows.Count).End(xlUp).Row

'This statement checks to see if an observation has been made for the given origin-destination/destination
'-origin and commodity code combination.
If Y > 1 Then

'This for loop runs through each observation of each origin-destination/destination-origin and commodity code
'combination.
For j = 2 To Y

'The variable TCNY, stands for Total Cost New Year, is updated by adding the total cost of each observation. When all
'the different origin-destination/destination-origin and commodity code combinations have been filtered through TCNY
'will represent the total cost of all the shipments shipped for the given fiscal year.
TCNY = TCNY + ActiveSheet.Range("G" & j)

'The variable TRNY, stands for Total Rate New Year, is updated by adding the amount charged of each observation. The
'amount charged is represented by the weight of an observation multiplied by the weighted average cost and the Refresh Rate.
'When all the different origin-destination/destination-origin and commodity code combinations have been filtered through
'TRNY will represent the total compensation for all shipments minus the AOR Factor and Composite Rate Adjustment.
TRNY = TRNY + (ActiveSheet.Range("B" & j) * DataArray(i, 3) * (1 + aRR))
Next j

End If

'Deletes the sheet that the filtered data was copied into.
wb2.Activate
ActiveSheet.Delete

'Erases the filter that was just applied to the data so that a new filter can be applied.
wb2.Activate
wb2.Sheets("FY09").Select
Selection.AutoFilter

Next i

End If

'Closes wb2
wb2.Close

wb1.Activate

'Sets the combined AOR factor to the total cost divided by the total compensation minus one.
aAOR = (TCNY / TRNY) - 1

'Writes the combined AOR Factor into the cell "R12"
wb1.Sheets("Sheet1").Range("R12") = aAOR

'Turns Excel's display alerts back on.
Application.DisplayAlerts = True

End Sub

```

Appendix E: Location Table of the Set *I* and *J*

Location Code	Location Description
01	CONUS (East Coast)
02	CONUS (Gulf Coast)
03	CONUS (California Coast)
04	CONUS (Northwest Coast)
05	Canada (Newfoundland)
06	Canada (Labrador)
07	Canada (Pine Tree)
08	Greenland (Thule)
09	Iceland
10	West Mexico and Central America
11	Panama (Caribbean Coast)
12	Bermuda Islands
13	Lesser Antilles Islands
14	Puerto Rico
15	Caribbean (Other)
16	Cuba (Guantanamo Bay)
17	Northern Europe
18	British Isles
19	Mediterranean (West)
20	Mediterranean (East)
21	West Africa
22	South and East Africa
23	Arabian Gulf
24	India and Burma
25	Alaska (East)
26	Alaska (West)
27	Hawaiian Islands
28	Marshall Islands
29	Marianas Islands
30	Taiwan
31	Bonin Islands
32	Philippines
33	Thailand
34	New Guinea and Australia
35	Great Lakes Area
37	Alaska (Aleutian Islands)
38	North Central Pacific Islands

39	South Pacific Islands
40	South West Pacific Islands
42	Scandinavia
43	Black Sea
44	South America (West Coast)
45	South America (East Coast)
46	Azores
47	Antarctica
48	Vietnam
49	South East Asia (Other)
50	Ryukyu Islands
51	Korea
52	Japan
53	Mississippi River
54	Rhine River
55	Cambodia
56	Panama (Pacific Coast)
57	Indian Ocean
58	North East Asia (Other)
59	Russia
90	Baltic Ports/Afghanistan via Baltic Ports
91	Izmir/Iskenderun/Mersin/Afghanistan via Turkey
92	Poti/Afghanistan via Poti
93	Pakistan/Afghanistan via Pakistan
94	AFG via Russia
98	Baku

Appendix F: QuadChart



RATE SETTING ANALYSIS: A STATISTICAL APPROACH TO OUTLIERS IN THE RATE SETTING PROCESS WITHIN THE UNITED STATES TRANSPORTATION COMMAND



2Lt Warren McGriff

Advisor: LTC Brian J. Lunday, PhD

Reader: Dr. Jeffery D. Weil, PhD

Department of Operational Sciences (ENS)

Air Force Institute of Technology

Sponsor: USTC J8

Background:

- USTC J8 sets transportation rates for DoD customers
- Rates must be set prior to each Fiscal Year (FY)
- Rates, costs, and workload (combined) yield a Net Operating Result (NOR)
- The goal is for the rates to result in a NOR of \$0 each FY
- Current Practice (UP to FY10):
 - Aggregate all rates by origin-destination/destination-origin and commodity code combinations
 - Weighted averages are then computed and multiplied by three factors: the Refresh Rate, the Accumulated Operating Result, and the Composite Rate Adjustment to set the rates for the upcoming FY

Motivation:

- Inaccuracies in the rate setting process lead to improper cost recoupment by USTC

Research Objectives:

- Evaluate the current practice's performance using six months of historical data compared to the utilization of a full-year of data
- Examine the impact of errors within the factors: Refresh Rate, Accumulated Operating Result, and the Composite Rate Adjustment
- Utilize a prescribed methodology to identify and examine the impact of outliers

Literature Review Topics:

- Transportation Rate History
- SAIE Interviews
- Forecasting Techniques
- Statistical Process Control (SPC)

Methodology & Results - Objective 1

- utilized a t-test as per Equation (1) to examine whether there was a statistically significant difference in the respective total NORs deviation from the target goal of NOR=\$0 when using six months of data compared to a full year of data in the rate setting process
- Net Operating Result (NOR): the difference between the revenue generated by a rate and the actual cost of transportation
- No statistically significant difference was identified ($p=0.67$ level)

$$D^2 = \frac{NOR_{Prescribed}^2 - NOR_{Actual}^2}{NOR_{Prescribed}^2 + NOR_{Actual}^2} \quad (1)$$

$$H_0: D = 0$$

$$D^2_{obs} = \frac{\sum_{i=1}^n \sum_{j=1}^m (NOR_{Prescribed}^2 - NOR_{Actual}^2)^2}{\sum_{i=1}^n \sum_{j=1}^m (NOR_{Prescribed}^2 + NOR_{Actual}^2)^2} = 1 \quad (2)$$

Methodology & Results - Objective 2

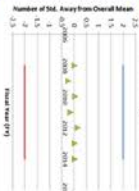
- Using historical data from FY08-FY15, solved for the optimal combination of the Combined the Accumulated Operating Result factor (AOR) and the Composite Rate Adjustment Factor (CRA), as per Equation (2)
- Identified historical combinations of these factors as contributing to ~25% of the error in achieved NOR at the $\alpha=0.05$ level.



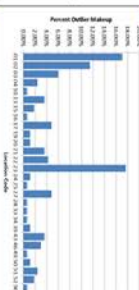
Methodology & Results - Objective 3

- Adapted SPC to identify outliers for rates set by the current practice as MNOR outliers as well as outliers according to historical rates
- Marginal NOR (MNOR): the contribution per measurement to the NOR
- The current practice possessed consistent behavior over time.
- Selected locations & commodities yielded outliers more frequently.
- The current practice corrects for MNOR outliers over time.
- Past rate setting practices have deviated from the prescribed practice using the "art of rate setting", as evidenced by the percent deviation from rates set and historical rates

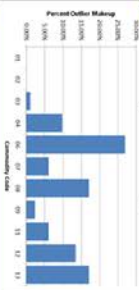
A. Methodology Behavior Over Time



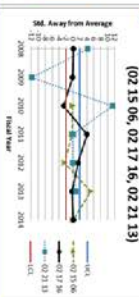
B. Location Code Outlier Frequencies



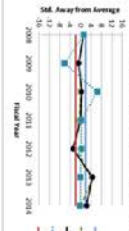
C. Commodity Code Outlier Frequencies



D. MNOR Outlier Behavior



E. MNOR Outlier Behavior



F. Proportion of Historical Rates Deviation

Percent of Rates	FY14	FY15	FY16
±100% Deviation	8%	10%	4%
±75% Deviation	8%	12%	8%
±50% Deviation	16%	17%	16%
±25% Deviation	35%	29%	36%

Conclusions:

- Implementing an additional six months of data does not improve the current practice
- The combination of the AOR and CRA factors contributes a significant source of error with regard to the deviation from the goal of a zero-balance NORs
- The Current rate setting practice corrects for outliers over time
- Manual deviation from prescribed rate setting practices should be used sparingly and only with sound justification.

Bibliography

- Allen, B. (1977). The Demand for Freight Transportation: A Micro Approach. *Transportation Research* , 9-14.
- Brooks, M. R., & Button, K. J. (1996). The Determinants of Shipping Rates: A North Atlantic Case Study. *Transport Logistics*, 21-30.
- Bryan, I. (1974). Regression Analysis of Ocean Liner Freight Rates on Some Canadian Export Routes. *Journal of Transport Economics and Policy*, 161-171.
- Chambers, J. C., Mullick, S. K., & Smith, D. D. (1971). How to Choose the Right Forecasting Technique. *Harvard Business Review*, 200-205.
- Command, U. S. (2005). *Dictionary of Military and Associated Terms* .
- Forkenbrock, D. J. (2001). Comparison of External Cost of Rail and Truck Freight Transportation. *Transportation Research Part A: Policy and Practice*, 321-337.
- Harvey, A., & Todd, P. J. (1983). Forecasting Economic Time Series with Structural and Box-Jenkins Models: A Case Study. *The American Statistical Association*, 299-307.
- Hill, G., & Woodworth, W. (1980). Automatic Box-Jenkins Forecasting . *The Journal of Operations Research* , 413-422.
- Hsu, C. C., & Chen, C. Y. (2003). Applications of Improved Grey Prediction Model for Power Demand Forecasting. *Energy Conservation and Management*, 2241-2249.
- Hummels, D. (2007). Transportation Costs and International Trade in the Second Era of Globalization. *Journal of Economic Perspectives*, 131-154.
- Jr., B. C. (2011). *Forecasting Enlisted Attrition in the United States Marine Corps by Grade and Years of Service*. Monterey California: Naval Post Graduate School.
- Moss, T. D. (2015, June). Rate Setting in the Trucking Transportation Realm. (W. McGriff, Interviewer)
- Muth, J. F. (1960). Optimal Properties of Exponentially Weighted Forecasts. *Journal of the American Statistical Association*, 299-306.

Schneerson, D. (1976). The Structure of Liner Freight Rates A Comparative Route Study.
Journal of Transport Economics and Policy , 52-66.

StatSoft. (1984). Quality Control Chart. Texas, USA.

REPORT DOCUMENTATION PAGE				Form Approved OMB No. 074-0188	
The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of the collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.					
1. REPORT DATE (DD-MM-YYYY) 24-03-2016		2. REPORT TYPE Master's Thesis		3. DATES COVERED (From – To) September 2014 - March 2016	
TITLE AND SUBTITLE Rate Setting Analysis: A Statistical Approach to Outlier Analysis In The Rate Setting Process Within The United States Transportation Command				5a. CONTRACT NUMBER	
				5b. GRANT NUMBER	
				5c. PROGRAM ELEMENT NUMBER	
				5d. PROJECT NUMBER	
				5e. TASK NUMBER	
6. AUTHOR(S) McGriff, Warren B. , 2 nd Lieutenant, USAF				5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAMES(S) AND ADDRESS(S) Air Force Institute of Technology Graduate School of Engineering and Management (AFIT/EN) 2950 Hobson Way, Building 640 WPAFB OH 45433-8865				8. PERFORMING ORGANIZATION REPORT NUMBER AFIT-ENS-MS-16-M-118	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) United States Transportation Command J8 Scott Air Force Base Scott AFB, Illinois 62225 618-229-5006 timothy.landvogt.1@us.af.mil ATTN: Timothy Landvogt				10. SPONSOR/MONITOR'S ACRONYM(S) USTC/J8	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION/AVAILABILITY STATEMENT DISTRIBUTION STATEMENT A. APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.					
13. SUPPLEMENTARY NOTES This material is declared a work of the U.S. Government and is not subject to copyright protection in the United States.					
14. ABSTRACT This research sought to identify areas in which the current rate setting methodology can be improved. We initially examined the use of six months of historical cost data versus a full year of data to set rates, concluding that there is not a statistically significant difference with respect to their relative effect on the NOR; USTC should proceed with their current practice. The research also identified outliers, first with regard to likelihood of historical rates not being set by the prescribed process and second with regard to whether the rates set by the prescribed process would be an outlier in terms of the marginal contribution to the net operating result. We found that approximately 8%, 10%, and 4% of the rates in FY14–FY16 were likely set using budget analyst experience in lieu of the prescribed method, for the most part imposing a reduction in the prescribed rates. Adapting classical Statistical Process Control (SPC) methods, we found that the prescribed rate setting method does work in aggregate but can induce recurrent outlier rates. However, a pattern in these outlier rates remains elusive – some are self-correcting – but the demonstrated methodology is shown to be useful for identifying outlier rates that do merit budget analyst experience-informed judgment for rate setting. The final component of this research examined the combination of two factors used in the current methodology to adjust current average weighted costs to set future rates: the Accumulated Operating Result and Composite Rate Adjustment factors. Using historical data from FY08–FY15, we calculate the optimal combined factor values for each respective fiscal year to achieve an NOR equal to \$0. In doing so, we concluded that the combination of these two factors contributed to approximately 25% of the induced error in NOR. We suggest a more detailed examination of these rate computations for additional analysis.					
15. SUBJECT TERMS Statistical Process Control					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT UU	18. NUMBER OF PAGES 95	19a. NAME OF RESPONSIBLE PERSON LTC Brian Lunday, AFIT/ENS
a. REPORT U	b. ABSTRACT U	c. THIS PAGE U			19b. TELEPHONE NUMBER (Include area code) (937) 255-3636 (brian.lunday@afit.edu)

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39-18