3-23-2018

Managing Supply Discrepancies: The Effect of Performance Measurement and Feedback on Order Fulfillment Quality

Michael J. Weber

MANAGING SUPPLY DISCREPANCIES:

THE EFFECT OF PERFORMANCE MEASUREMENT AND FEEDBACK ON

ORDER FULFILLMENT QUALITY

THESIS

Michael J. Weber, Captain, USAF

AFIT-ENS-MS-18-M-168

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.

AFIT-ENS-MS-18-M-168

MANAGING SUPPLY DISCREPANCIES: THE EFFECT OF PERFORMANCE MEASUREMENT AND FEEDBACK ON ORDER FULFILLMENT QUALITY

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 Logistics and Supply Chain Management

Michael J. Weber, MA

Captain, USAF

March 2018

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

AFIT-ENS-MS-18-M-168

MANAGING SUPPLY DISCREPANCIES: THE EFFECT OF PERFORMANCE MEASUREMENT AND FEEDBACK ON ORDER FULFILLMENT QUALITY

Michael J. Weber, MA
Captain, USAF

Committee Membership:

Dr. Daniel Steeneck
Chair

Dr. William Cunningham
Member

AFIT-ENS-MS-18-M-168

Abstract

Extensive research on the impact of shipping and packaging errors in the private sector finds numerous negative outcomes, including reduced customer satisfaction, reduced customer loyalty, and lower profitability. However, little research has been done examining the impact of order fulfillment errors on military operations. The purpose of this research is to quantify the impact of supply discrepancy reports (SDRs) on military aircraft readiness metrics, including cannibalizations, not mission capable supply (NMCS) hours, aircraft availability and MICAP hours. Results show SDRs significantly impact aircraft readiness metrics in seven of the fifteen analyses conducted. Additionally, a quasi-experimental study is implemented at DLA Distribution Susquehanna, Pennsylvania (DDSP) aimed at reducing supply discrepancies using performance measurement and feedback over a seventeen-week period. Cumulative sum (CUSUM) control charts showed a decline in the number of reported SDRs for fifteen consecutive weeks, amounting to the lowest average in over six years. The results of this research suggest that aircraft readiness metrics across the Air Force could show measurable improvement if similar SDR reduction strategies are implemented throughout more DoD suppliers.

Acknowledgments

I would like to express my sincere gratitude to my faculty advisor, Dr. Daniel Steeneck, for his guidance and support throughout the course of this thesis effort. Your valuable insight, feedback, and encouragement over the past year is truly appreciated. I would also like to thank my sponsors of the 635th Supply Chain Operations Wing Director of Logistics, and Commander of DLA Distribution Susquehanna, PA for the support and latitude provided to me in this endeavor.

Michael J. Weber

AFIT-ENS-MS-18-M-168

To my amazing wife:
Your love, support, and encouragement over the past year and a half made it possible to complete this work.

Table of Contents

List of Figures

List of Tables

MANAGING SUPPLY DISCREPANCIES: THE EFFECT OF PERFORMANCE MEASUREMENT AND FEEDBACK ON ORDER FULFILLMENT QUALITY

I. Introduction

One of the primary roles of the Air Force supply chain is weapon system sustainment—which involves years of maintenance and repairs aimed at maximizing equipment lifespans. This spare parts supply chain is similar to commercial service supply chains such as auto repair, where customer satisfaction is largely dependent on the firm's ability to provide fast, reliable service. In consumer markets, poor order fulfillment quality leads to reduced customer satisfaction, decreased loyalty, and ultimately reduced financial performance (Rao, Griffis, & Goldsby, 2011). These metrics are useful in the private sector, because customers are free to change suppliers if service falls below expectations. However, the impact of poor supply chain performance in the Air Force must be assessed differently because in most cases customers do not have this flexibility.

The Defense Logistics Agency (DLA) provides 86 percent of the military's spare parts, in support of over 2,300 weapon systems. Given its size and scope, shipping and packaging errors are to be expected. However, the Air Force alone reports over 20,000 supply-related errors each year due to the huge volume of parts consumed on an annual basis; most of which originate from DLA distribution centers. The Defense Distribution Center Susquehanna, Pennsylvania (DDSP) is the Air Force's largest single supplier and, not surprisingly, accounts for the highest number of errors each year. On a given day, DDSP commits an average of 11 supply discrepancies on Air Force orders, amounting to over 300 errors each month. With an average of over 4,000 daily Air Force orders, this amounts to an error rate of approximately 0.025 percent. For comparison, the average error rate among a sample of over 500 companies surveyed in 2013 was 0.05 percent (Warehousing Education and Research Council, 2013). This suggests that order fulfillment quality at DDSP actually exceeds that of its commercial counterparts. While the rate of error as a percentage of total orders is low, the total frequency of errors is of concern due to the severity of the disruptions that can be caused by discrepant orders. While order fulfillment errors in the private sector can result in unsatisfied customers, errors in DoD orders have the potential to impact military operations. It is therefore important that this problem is investigated to better understand the impact of poor order fulfillment on Air Force readiness, and how supply discrepancies can be effectively managed within DLA.

Supply Discrepancy Report (SDR)

The Supply Discrepancy Report (SDR) is a tool used to report shipping and packaging errors attributable to the shipping activity (including U.S. Government sources and contractors), and to determine the root cause of the discrepancies, affect corrective actions, and prevent recurrence (Defense Logistics Manual 4000.25, 2016). Examples of supply discrepancies include: overages or shortages, incorrect items received, missing parts, misdirected shipments, improper packaging, expired shelf life, damaged goods (not TSP-related), and other supply-related errors. The receiving activity will initiate the SDR when one or more of the above conditions are noted on an inbound shipment, then submit to the responsible shipping activity for corrective action and enter into DLA Transaction Services Web Supply Discrepancy Reporting (WebSDR) tool. Alternatively, the hard copy SF 364 Report of Discrepancy form may be used by exception when access to WebSDR is unavailable (see Appendix A).

While the financial cost of SDRs can be estimated in terms of administrative and holding costs, the actual impact on operations is not well understood (McKinney, 1995). The supply discrepancy reporting program is designed to facilitate evaluation of supplier performance by identifying trends and disseminating reports to DoD component representatives, with the goal of bringing management attention to problems with shipping activities and to prevent recurrence (DLM 4000.25, 2016). Yet, aggregate SDR metrics are not closely monitored within the supply activities where the discrepancies occur, and thus little emphasis has been placed on improving supplier performance within DLA.

At bases with assigned aircraft and an active flying mission, supply discrepancies can disrupt operations by increasing lead times and delaying essential maintenance actions, potentially impacting the wing's overall aircraft availability and its ability to fulfill air tasking orders.

Performance Measurement

A common strategy used by organizations to reduce the occurrence of preventable errors and other negative indicators is the use of metrics to track and report performance. Well-developed metrics provide leaders with relevant and useful information about the organization's performance and enable evidence-based decision making, as opposed to reliance on "gut feelings" from management (Stahl, 2014). In military organizations, where discipline and strict adherence to commander-directed policies are expected, metrics can be used as a tool to elicit desired actions and behaviors. Nearly all operational units within the Air Force track a variety of metrics for either internal use, reporting to higher headquarters (HHQ), or both. These include personnel management related metrics such as physical fitness failure rates, training compliance, or individual medical readiness statistics, as well as operational metrics specific to the unit's mission such as aircraft availability and inventory accuracy. While there is considerable variety in the types and purposes of metrics tracked throughout the Air Force, they generally seek to achieve the same two objectives:

1. To assess performance in a given function or activity deemed essential to the unit or Air Force mission, and

2. To drive behaviors toward achieving a set level of performance in the given function or activity.

Therefore, if it can be shown that supply discrepancies impact Air Force operations in a meaningful way, then the implementation of an order fulfillment quality metric as a performance measure within DLA may drive behaviors that result in improved performance.

Background

The Air Force has long sought to develop effective metrics to evaluate its ability to meet strategic objectives and improve performance. As early as 1956, with the establishment of set standards for aircraft maintenance in AFM 66-1 *Maintenance Management,* Air Force leaders have pursued strategies to improve aircraft in-commission rates, component repair standards, and scheduling objectives (Stahl, 2014). Over the past 60 years, the Air Force has continued to develop its metrics to best direct behaviors toward attaining performance standards deemed essential toward meeting mission requirements. Perhaps the metric of most concern with regard to readiness over this time is aircraft availability, given that the ready employment of aircraft is integral to nearly all Air Force activities. Many of the strategies aimed at improving aircraft availability and developing other efficiencies have involved organizational changes to aircraft maintenance units based on the size, culture, and strategy of the Air Force at the time. These strategies, while well intentioned, often led to poor maintenance practices that negatively impacted the metrics they sought to improve. Reorganization often proved ineffective because the actual behaviors and activities that impact aircraft availability were not given equal consideration and attention (Johnson, 2000). Thus, the performance of work centers and activities that impact aircraft availability are of crucial importance to monitor if measurable improvements are desired.

One function with direct impact on the availability of aircraft within the Air Force is Materiel Management. Effective inventory management and demand forecasting practices are essential to ensuring spare parts are available when and where they are needed to support mission requirements. Such strategies have garnered much attention over the past several decades, leading to the development of advanced readiness-based sparing and multi-echelon inventory models currently used today for computing stockage requirements (Muckstadt, 1973). Yet, equally important as the systems in place to manage inventory are the behaviors and practices of those responsible for their effective operation. Performance measures, then, must assess not only the degree to which the systems are meeting standards with regard to managing inventory, but the degree to which orders are fulfilled effectively.

In 1993, the requirement for performance measurement was codified when the United States enacted the Government Performance and Results Act (GPRA), a law requiring government agencies to establish strategic objectives, implement strategies to improve performance, and conduct regular evaluations of their programs (Vector Research, 1997). In response, the Department of Defense published within its strategic plan a framework of performance goals, measures, and targets to ensure compliance with the GPRA. Formal performance improvement initiatives soon emerged across the DoD.

In 1999, the Logistics Management Institute (LMI) provided a guide to senior DoD leaders titled *Supply Chain Management: A Recommended Performance Measurement Scorecard.* The report asserted that the metrics in use at the time were ineffective in measuring the effectiveness of the DoD supply chain. Instead, LMI proposed a set of balanced performance measures across customer service, cost, readiness, and sustainability performance objectives. Recommended measures included supply chain response time, non-mission capable rates, and *perfect order fulfillment*, defined as an order that is complete, on-time, includes accurate information, and is in expected condition (Klapper et al., 1999). This guide was based on Kaplan and Norton's Balanced Scorecard (BSC) system for performance measurement, a widely-used and highly regarded business tool that emphasizes the use of non-financial and often intangible performance metrics to better achieve organizational goals. The four questions answered by the BSC are (1) How do customers see us? (2) What must we excel at? (3) Can we continue to improve and create value? and (4) How do we look to shareholders? (Kaplan & Norton, 1992).

Since its creation, the BSC has been applied across numerous Air Force MAJCOMs to help achieve organizational goals. HAF/A4 implemented the BSC in 2006 with a primary goal of increasing equipment availability by 20 percent. The logistics strategy to achieve this overarching goal was to "improve the response time to supply chain requirements", measured by three metrics: customer wait time, MICAP hours, and MICAP incidents. Notably, LMI's recommended measure of perfect order fulfillment was excluded from the scorecard. This is problematic because much of what constitutes these metrics is outside of commanders' control, such as availability of repair parts and reliability of transportation service providers. Although improvement across the included metrics would very likely increase aircraft availability, measurement alone is not sufficient to achieve this goal. Thus, to effectively improve these metrics, commanders must target the factors within their control. Specifically, order fulfillment quality, measured by the rate of SDRs within a given timeframe, must be monitored to identify negative performance trends and implement necessary corrective measures.

Problem Statement

The extent to which supply discrepancies impact aircraft readiness is not well understood, and order fulfillment quality is not measured within DLA Distribution. As a result, supply discrepancies occur frequently and supplier performance has not recently been targeted for improvement. This research seeks to quantify the operational impact of SDRs on the Air Force, measured by their impact on MICAP hours, cannibalization rates, NMCS rates, and aircraft availability. In addition, this research seeks to determine the effectiveness of measuring SDR rates as a performance metric within DLA Distribution Susquehanna, PA to help improve order fulfillment quality.

Research Questions

1. To what extent do supply discrepancies impact aircraft readiness?
2. How can performance measurement and feedback help reduce SDRs within Defense Distribution Center Susquehanna, PA?

Investigative Questions

1. What is the effect of SDRs on MICAP hours, NMCS hours, cannibalizations, and aircraft availability?
2. What effect does order fulfillment quality measurement have on performance?
3. What effect does performance feedback have on order fulfillment quality?

Implications

As a result of this research, leaders will have a better understanding of how supply discrepancies impact key maintenance metrics and potentially degrade operations. Additionally, this study will demonstrate the extent to which supply discrepancies can be reduced as a consequence of managing SDRs as a performance metric within DLA. Results of this experiment showed a 35 percent reduction in SDRs at DLA Distribution Susquehanna, PA over a period of 17 weeks. By reducing SDRs across more DLA distribution centers, the aircraft maintenance community will be able to better support flying operations worldwide.

Organization

The following chapters will be organized as follows:

Chapter II provides a review of relevant literature on the impact of poor order fulfillment in consumer markets, and identifies a gap in the literature with regard to the effect of poor order fulfillment in military settings. Additionally, a review of performance measurement theory, development, and practice is provided. Furthermore, the current state of DoD supply chain metrics is discussed in the context of performance management and operational improvement. Finally, the role of feedback in enhancing performance is discussed.

Chapter III details the specific methods used to collect data in this study, the samples chosen for in-depth analysis, procedures used to carry out the study, statistical analyses chosen to test the hypotheses, and the tests used to evaluate the validity of the findings.

Chapter IV provides the results of the analyses discussed in chapter III. This chapter will assess whether the given hypotheses are supported by the data.

Chapter V discusses the results from chapter IV and offers insight into the possible explanations for the findings, implications and limitations of the results, and suggestions for future research related to this topic.

II. Literature Review

This chapter provides a detailed overview of the order fulfillment process, both in commercial companies and in the Air Force, as well as the impact of poor order fulfillment on customer loyalty and profitability. Additionally, an overview of business performance measurement theory is provided, along with the criteria required to develop effective metrics that balance both financial and non-financial information. Next, the current DoD policies surrounding supply chain metrics are discussed, as well as the weaknesses with current practices. Finally, levels of performance feedback are discussed, along with their impact on future effort and attention to learning.

Order Fulfillment

Order fulfillment is a crucial component of the supply chain process, and essential for maintaining positive customer relations. It is a complex process involving a network of interdependent activities. These activities are performed by various functional entities within an organization with the goal of meeting or exceeding customer expectations and maximizing profits (Croxton, 2014; Lin & Shaw, 1998). These processes include generating, filling, delivering, and servicing customer orders in a way that is both efficient and achieves the greatest competitive advantage for the firm.

From a strategic standpoint, effective order fulfillment practices require a deep understanding of customer needs. A firm must then align its capabilities and business processes, both internal and external, to effectively meet customer needs (Croxton, 2014). Additionally, each customer has unique service expectations and must be dealt with accordingly (Christopher, 2005; Thirumalai & Sinha, 2005; Mentzer et al., 2001;

Bienstock et al., 1997). For example, some customers may value price above all, while others are most sensitive to delivery reliability. Still others may desire a relationship approach, "valuing technical support and close supplier liaison." Thus, the challenge is to develop supply chain solutions most appropriate to meeting the needs of each customer segment. This includes having a responsive supply chain with sufficient capacity to handle surges in demand, and the ability to provide a quality customer service experience even when demand exceeds supply (Croxton, 2014). According to Coyle, Bardi, and Langley (2002), care must be taken to ensure the need for customer service is not overshadowed by the goal of increased sales and reduced costs. Thus, the selection of effective metrics to measure performance is essential to strategic order fulfillment, both in terms of meeting customer expectations and ensuring efficient processes that maximize profitability.

Operational Order Fulfillment Process.

At the operational level, order fulfillment is a transactional process that relies on a network of logistics functions. Commonly referred to as "order management", this process involves all activities that occur over the order cycle, or the time from when the order is received by the seller until it is received by the buyer (Coyle et al., 2009). Figure 1 provides an overview of the operational order fulfillment process.

The order cycle begins with the generation of a customer order. Although this activity has largely become automated with the advancement of electronic data interchange (EDI) and vendor-managed inventory (VMI) systems, it is the process that sets in motion the logistics function. Thus, great care must be taken to ensure orders are captured accurately and efficiently because once entered, errors can be very costly to correct (Croxton, 2014).

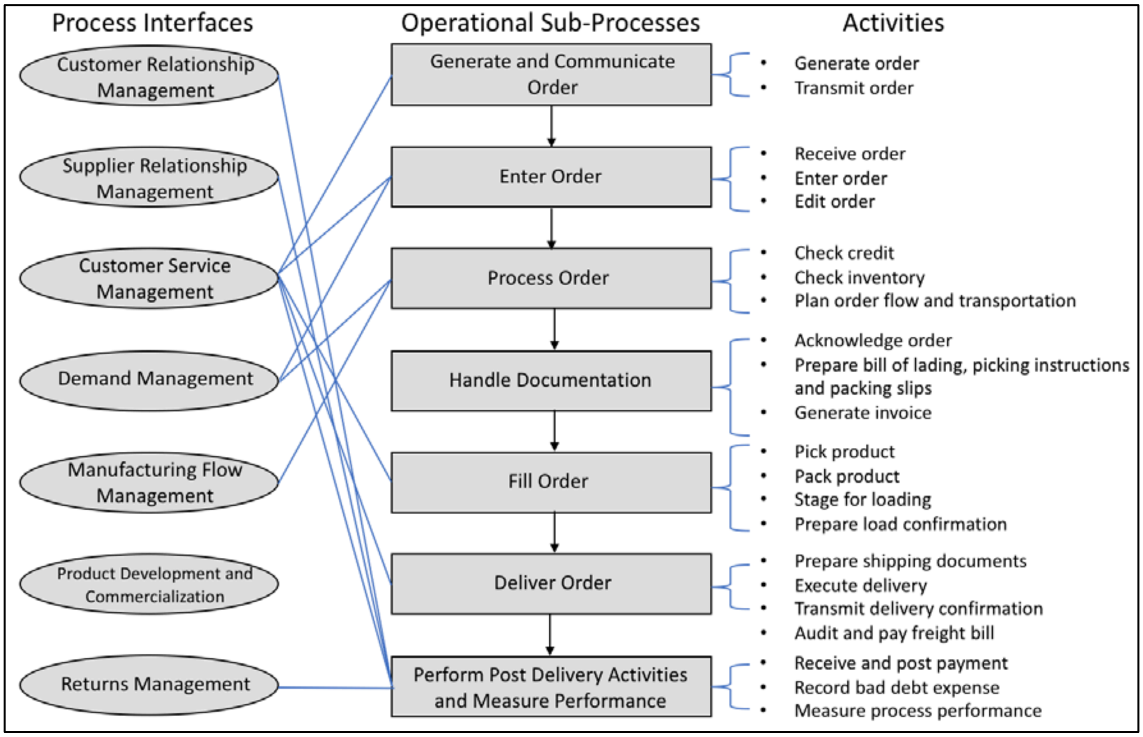


Figure 1. The Operational Order Fulfillment Process (Croxton, 2014)

Once the order has been received, the next step is to process the order. This involves verifying inventory levels for availability of products, and determining how the order will move through the supply chain; a process known as distribution requirements planning (DRP) (Croxton, 2014). If inventory is not available to fill the order, the seller will coordinate with the manufacturer or supplier to schedule delivery (Coyle et al., 2009). In either case, the buyer will typically be provided an expected delivery date. This step also involves selection of the carrier based on the priority of the order, characteristics of the load (i.e. size, weight, freight type, etc.), and shipping costs.

Service performance factors such as delivery reliability, damage record, driver courtesy, and commitment to excellence must also be considered when selecting a carrier (Copacino, 1997). In most cases the interface with the carrier is the only face-to-face interaction the customer will have throughout the order fulfullment process, so it is important that the carrier represents the seller in a positive and professional manner.

After the order has been processed and method of delivery planned, the documentation for the order is generated. These documents often include the bill of lading, cargo manifest, and customer invoice. International orders will also require customs clearance documents (Croxton, 2014). Care must be taken to ensure documents are accurate and complete to prevent discrepancies in picking and fulfilling orders.

Next, the order is picked at the distribution center or warehouse and prepared for delivery. This process includes packing, staging, and arranging the load for shipment (Croxton, 2014). Accuracy and timeliness are critical during order preparation, and often hindered by inefficient or outdated processes. Advancements in warehouse layout and technologies such as RFID and mechanized material handling systems have resulted in significant improvements in the efficiency and timeliness of order fulfillment in recent years (Croxton, 2014). Technology has also enabled dramatic improvements in in-transit visibility (ITV), allowing the customer to track the order throughout the shipping process. Especially important when lead times are longer than average, such as with international shipments, ITV is crucial to ensuring customer satisfaction (Peleg-Gillai & Bhat, 2006).

The final two steps in the order fulfillment process are order shipment and post-delivery activities, such as order verification by the customer and assessment of process performance by the seller. It is at this point that any discrepancies in documentation, quality, quantity, completeness, or accuracy are identified by the customer, and the seller must take action to either credit the buyer for the inadequate material or, if possible, correct the discrepancy. Research conducted by Davis-Sramek, Droge, Mentzer, and Myers (2009) indicates that the level of customer service provided by a seller in resolving order fulfillment errors plays a significant role in the customer's overall rating of the firm. Therefore, it is crucial that efforts are made both to prevent errors from occurring and ensuring errors are corrected effectively when they do occur.

Common metrics used to track performance of the order fulfillment process are order-to-cash cycle time, customer wait time, and perfect order fulfillment. Order-to-cash cycle time measures the total time from when the order is placed by the buyer to the time payment is received by the seller. Customer wait time is the total time from when the order is placed to the time the order is received by the customer. Finally, perfect order fulfillment, refers to the percentage of orders that are delivered error-free (Christopher, 2005; Bowersox, Closs, & Cooper, 2002). According to Copacino (1997), the order fulfillment process is highly variable and inefficient in many companies. Close monitoring of metrics can reveal opportunities for process improvements that can drive cost reductions and drastically shorten order cycle times.

Order Fulfillment Service Quality.

Due to the inherent challenges associated with meeting customer demand in spare parts supply chains, timely and accurate order fulfillment is essential for businesses to remain competitive. Extensive research has examined order fulfillment as a key component of Physical Distribution Service Quality (PDSQ), rated by customers according to timeliness, availability, and condition of orders (Bienstock et al., 1997).

While traditional factors such as price and product quality are still key differentiators in business, there is ample evidence to suggest that PDSQ is a critical determinant of business success. As previously discussed, however, effective order fulfillment is no easy task (Ricker and Kalakota, 1999). Therefore, companies that can develop efficient supply chain processes and successfully meet customer expectations will be in position to have a competitive advantage in the market (Davis-Sramek et al., 2009).

According to Davis-Sramek, Mentzer, and Stank (2008), order fulfillment service quality is the customer's evaluation of all activities associated with the initial purchase until it is fulfilled by the seller and the customer is satisfied. These activities can be broken into relational service quality and technical service quality (also termed operational service quality). Relational service quality refers to the seller's ability to establish a trusting relationship with the buyer (Chiou and Droge, 2006). This concept is especially important in transactions requiring a significant service component, both in business-to-business (B2B) and business-to-customer (B2C) contexts.

In a B2B context, trust and rapport is essential in services marketing and building long-term relationships. More and more companies are outsourcing business processes not considered core competencies, such as logistics, finance and accounting, human resources, and marketing (Lovelock & Wirtz, 2011). Because these functions are essential for success in all organizations, it is important that strong relationships are established. Relational service quality in a B2C situation could involve any high-involvement service or purchase decision such as appliance shopping, home remodeling projects, or the premium cosmetic product market (Chiou and Droge, 2006). Customers of these products and services rightly expect high relational service quality because they are putting their faith in the expertise of service employees—typically at a high price.

Technical order fulfillment service quality, conversely, refers to the customer's perception of the seller's ability to "deliver the right products on time and dependably" (Davis-Sramek et al., 2009). The truest measures of technical service quality are perfect order fulfillment, discussed in the previous section, and customer satisfaction surveys. In general, technical order fulfillment service quality is easier to measure than relational service quality because the seller is immediately made aware if an order is received in an unsatisfactory condition. Relational service quality is often left unmeasured unless either the seller requests feedback via a customer survey, or the buyer provides unsolicited feedback in the form of an online review. Both technical and relational order fulfillment service quality are important because, as shown in Figure 2, research suggests they directly relate to the customer's satisfaction, commitment to the company, and ultimately loyalty behavior. Technical order fulfillment is the focus of this thesis and will thus be termed simply "order fulfillment" for the remaining chapters.

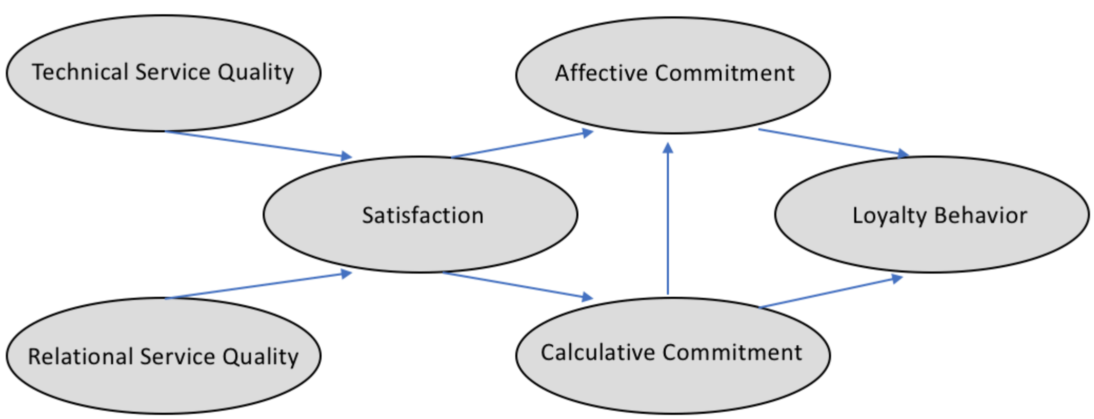


Figure 2. Technical and Relational Service Quality (Davis-Sramek et al., 2009)

Effect of Order Fulfillment on Satisfaction, Commitment, and Loyalty.

As indicated above, effective order fulfillment has been linked with several positive outcomes; including customer satisfaction, commitment, and loyalty.

According to Bowersox, Closs, and Cooper (2002), customer satisfaction is achieved when the supplier's performance meets or exceeds expectations. Thus, customer satisfaction is a direct consequence of effective order fulfillment. The question, then, for any for-profit firm is *How does customer satisfaction increase profitability?* The answer, based on years of research across numerous consumer markets, is that customer satisfaction is directly related to favorable impressions and increased loyalty to a particular company, which in turn drives higher profits.

Davis-Sramek et al. (2009), found that order fulfillment quality was a significant predictor of customer satisfaction, as well as commitment to continue to do business again in the future among retail customers of a large manufacturing company. Similarly, Hewett et al. (2006) and Chandrashekaren et al. (2007) found customer satisfaction was directly related to repurchase intentions and continued patronage in both large service organizations and industrial markets. While plans of continued patronage are positive outcomes of order fulfillment quality, they mean little if customers do not act on their intentions. According to Fornell (1992), firms are likely to abandon efforts to maximize customer satisfaction "unless it can be demonstrated that there are positive economic returns." Fortunately, for companies devoted to providing a superior customer service experience, evidence suggests that repurchase intentions are in fact associated with increased loyalty behavior, and even higher market share.

According to Reichheld and Sasser (1990), high customer satisfaction indicates loyalty, which will in turn lead to increased profitability because loyal customers ensure a steady stream of future revenue. In their highly influential study, Swedish researchers Anderson, Fornell, and Lehmann (1994) compiled customer satisfaction surveys of 77 firms across a wide variety of industries, and compared the responses to each firm's end-of-year return on investment, a long-term measure of economic health (Anderson et al., 1994). Results indicated that customer satisfaction, measured in the first half of the fiscal year, was a significant predictor of year-end economic returns. Similarly, Stank, Goldsby, Vickery, and Savitskie (2003) surveyed customers of large 3PL providers and found that companies with higher service performance, measured by order accuracy and responsiveness, held greater market share than competitors, and were rated higher in customer satisfaction and loyalty. More recently, Griffis et al. (2012) found that order fulfillment, and satisfaction with online retailers, is a predictor of referral behavior among customers. These findings are significant because they suggest that order fulfillment service quality leads not only to customer satisfaction, but also greater purchasing behaviors and potential profitability.

Not surprisingly, the opposite is true of businesses that fail to meet customer expectations. Rao, Griffis, and Goldsby (2011) found that retailers who failed to deliver upon order fulfillment promises experienced reduced future orders, as well as reduced dollar value of subsequent orders. Order fulfillment glitches were also associated with increased order anxiety, measured by proxy using the number of times a customer checked the order status online as an indicator of anxiety (see Figure 3) (Rao et al., 2011).

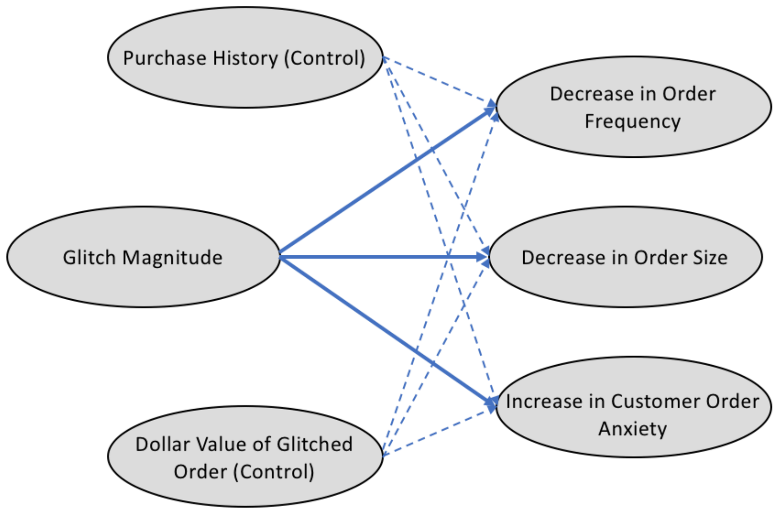


Figure 3. Impact of Order Fulfillment Glitches (Rao et al., 2011)

Additionally, A 2012 survey of six hundred consumers regarding their expectations for delivery of online purchases found that 62 percent of respondents indicated they would be much less likely to shop with a retailer online if their purchase is not delivered within two days of the date promised. Further, 29 percent of respondents stated they would permanently cease shopping with the retailer altogether if they received an incorrect delivery (Voxware, 2012). These findings suggest that customer opinions about a retailer can be greatly impacted by their shopping experience.

Studies have also shown that these negative perceptions can and often do translate into lost sales and lowered financial performance. According to Hendricks and Singhal (2005), delayed shipments can hurt the credibility and reputation of a firm, leading to reduced customer loyalty and a loss in net sales. In many instances, the impact on the firm's credibility may require an increase in advertising and public relations expenses, and can increase the cost of raising capital due to increased wariness of investors.

The negative outcomes associated with ineffective order fulfillment practices discussed above are meaningful to commercial firms because their objective is to maximize profits. The following section will discuss the process of order fulfillment in the Air Force, and why traditional measures of performance are not relevant in a military context.

Air Force Order Fulfillment.

Order fulfillment in the Air Force is handled in largely the same way as the private sector. The major differences are primarily in the documentation and systems used to generate and manage transactions. The Integrated Logistics Support-Supply (ILS-S) is the overarching term used to describe the systems used by retail materiel management operations, and is comprised of the Standard Base Supply System (SBSS), the Enterprise Solution-Supply (ES-S), and the Air Force Supply Centralized Database (AFSCDB) (AFH-123, 2013). SBSS is the legacy base level inventory accounting and order management system in the Air Force, while ES-S is the information technology system that provides transaction processing, order management, shipment tracking, asset management, and data visibility functions. AFSCDB combines all data from SBSS accounts into a database, providing global data access and visibility (AFMAN 23-122, 2016).

Orders placed using Military Standard Requisitioning and Issue Procedures (MILSTRIP) at Air Force installations may be entered directly from an external maintenance IT system such as the Integrated Maintenance Data System (IMDS) or G081, or requested from the Materiel Management activity and entered into ILS-S using the DD Form 1348. If the requested material is not in stock at base level, ILS-S will generate an automatic requisition to the source of supply (e.g. DLA, GSA, Boeing, etc.) under the Uniform Military Movement and Issue Priority System (UMMIPS) (AFMAN 23-122, 2016). UMMIPS establishes Time-Definite Delivery (TDD) standards based on customer designated priorities. A Required Delivery Data (RDD) of less than 8 days for CONUS customers and 21 days for OCONUS customers indicates the requirement for expedited shipping, generally due to a not-mission capable supply (NMCS) condition.

For DLA managed items, a retail service order (SO) is created and managed through DLA's Enterprise Business Solution (EBS) system. For items in stock, a material release order (MRO) will be processed and the item will be designated for shipment to the customer. Material is then validated for kind, count, and condition (KCC) and prepared for shipment in accordance with DLA standard operating procedures (SOPs). Finally, transportation documentation is generated and the material is delivered to its destination via either organic transportation or commercial carriers. Damaged, incomplete, or inaccurate shipments are investigated and monitored to resolution in accordance with DLA Distribution policy and procedures (DLAI 4140.08, 2015).

While order fulfillment in consumer markets has been studied exhaustively, very little research has been done on the impact of bad order fulfillment in the Air Force. Furthermore, the impact of inaccurate shipments cannot be assessed using traditional measures of customer satisfaction, loyalty, and profitability, because these are not relevant Air Force metrics. What little research that has been done has focused simply on the financial costs associated with discrepant shipments.

Bray (1990) quantified two costs resulting from discrepant shipments – administrative costs and holding costs. Administrative costs include SDR processing, investigation, and resolution, and were estimated to average 519 dollars per shipment (adjusted for inflation). Holding costs result from the storage and handling of discrepant items, and from “the lost opportunity of investment for money ‘tied up’ in these supplies.”, and were estimated to cost 3.22 percent of the contract value for a typical DLA item. A key weakness of this study, as identified by the author, is that it failed to quantify the “readiness degradation” resulting from discrepant shipments. The present study seeks to fill this critical gap in literature by quantifying the readiness degradation resulting from supply discrepancies in terms of aircraft availability, MICAP hours, not-mission capable supply rates, and cannibalization rates.

Performance Measurement

Supply chain management has grown increasingly complex over the last several decades as manufacturing and logistics companies have evolved into global, integrated organizations. As a result, companies are becoming increasingly reliant on information systems to drive down costs and increase efficiency (Akyuz & Erkan, 2010). In addition, inter-organizational collaboration has become widely accepted as a means to allow for a greater flow of ideas, information, and people, resulting in greater innovation and reduced risk (Chesbrough & Garman, 2009). With greater integration, however, comes greater challenges in measuring performance (Bitici et al., 2011).

In an effort to better meet customer expectations, firms have adopted the use of performance measurement systems across all facets of operations. Business performance measurement can be defined as a system utilizing a multi-dimensional set of metrics to quantify the efficiency and effectiveness of actions for the planning and management of a business (Monica et al., 2007; Bourne et al., 2003; Neely et al., 2000). According to Gunasekaran and Kobu (2007), there are eight purposes of a performance measurement system:

1. Identifying success,
2. Identifying if customer needs are met,
3. Better understanding of processes,
4. Identifying bottlenecks, waste, problems, and improvement opportunities,
5. Providing factual decisions,
6. Enabling progress,
7. Tracking progress, and
8. Facilitating a more open and transparent communication and co-operation.

It follows that performance measurement initiatives should provide insight into one or more of these areas in order to "supply the right information to the right decision-maker" (Andersson et al., 1989).

Unfortunately, many performance measurement systems have failed to provide benefit because companies have been unable to develop the metrics needed to maximize efficiency and effectiveness (Akyuz & Erkan, 2010). In many cases, contextual and processual issues stand in the way of successful implementation. Problems such as a lack of time or resources, lack of management involvement, lack of "buy-in" from employees, or lack of vision and strategy in developing metrics can doom performance measurement initiatives from the start (Bourne et al., 2002). Even under ideal conditions, however, performance measurement systems will not prove beneficial unless the metrics are developed thoughtfully and correctly.

According to Gunasekaran and Kobu (2007) there are five primary problems with performance measurement systems:

1. Incompleteness and inconsistencies in metrics,
2. Failure to develop a balanced set of financial and non-financial measures,
3. Too many metrics, making it difficult to discern between the important and the trivial,
4. Failure to connect strategy to measures, and
5. Being too inward looking.

Given the propensity for problems and difficulty in establishing effective performance measurement systems, it is important that companies take a strategic approach to developing the right metrics for the organization. The following sections will discuss important factors that must be considered when establishing performance metrics, as well as the weaknesses in current DoD metrics that this study seeks to correct.

Establishing Effective Metrics.

A comprehensive review of the existing literature on performance measurement has identified numerous characteristics of effective measurement systems (Akyuz & Erkan, 2010; Wauters, 2009; Kobu, 2007; Gunaskekaren, 2004). This review will focus on the most commonly identified characteristics that make up quality metrics for use in measuring performance.

Link Measures with Strategy.

The tendency of most firms that utilize performance metrics is to measure all quantifiable business aspects without regard for the actual impact these metrics may have on performance. It is important for leaders instead to view performance measurement as a strategic tool to help reach organizational objectives (Neely, 2002). Companies with carefully crafted performance measures have typically done so with the expressed objective of achieving organizational goals, because measurement without purpose is typically a waste of time and resources (Monica et al., 2007). Thus, prior to implementing a performance measurement system, leaders must first determine *what* to measure and *how* those measures achieve strategic alignment (Bitici et al., 2012).

Balance Financial and Non-financial Measures.

Traditional performance measures up until the 1990s almost exclusively focused on financial outcomes such as revenue, market share, and return on investment because they are clear indicators of how well the organization is meeting corporate objectives (Otley, 2002). However, because financial measures are lagging indicators they provide little insight into *why* the organization is or is not meeting its goals (Kumar et al., 2013). Thus, specific leading measures are needed to fully understand the drivers of performance (Otley, 2002). Measuring both financial and non-financial metrics is key in order to relate the operative drivers with the financial results (cause and effects) (Gairdelli, Saccani, & Songini, 2007).

Valid and Reliable.

Validity, according to Noe et al. (2016) refers to the extent to which a measure assesses all the relevant aspects of performance. A measure is considered deficient if it does not measure all aspects of performance, and contaminated if it measures irrelevant aspects of performance. The common phrase "What you measure is what you get" could not hold more true with regard to performance measurement, so it is important that the metrics chosen to assess performance are in fact representative of the desired outcome

(Hauser & Katz, 1998). For example, if employees are evaluated based on absenteeism, then the goal will become minimization of absences instead of the desired outcome of increased productivity.

A reliable measure, conversely, is one that is consistent and free from random error. Reliability is a necessary, but not sufficient condition for validity. In other words, a measure may provide consistent data, but if the information is consistently inaccurate then it cannot be a valid measure (Brennan, 2001). Determining reliability, by definition, requires at least two instances of a given measure. Therefore, some degree of history must be established before a measure can be determined reliable and thus by extension, valid (Brennan, 2001).

Controllable.

According to Copacino (1997), performance measures must be controllable and important to the functions being measured if the goal of measurement is to improve in some meaningful way. Control is achieved by having "appropriate standards of performance relative to the established metrics to indicate when the logistics system requires modification or attention." (Bowersox et al., 2002). Having sufficient control enables manages to act when logistics systems fall below standards by identifying causes and making adjustments to bring the system back into compliance. This is an important point because if the function or individual responsible for a performance measure do not have control, then no action can be taken to improve the system. For example, using on-time deliveries as performance measure for a sales department would be inappropriate because the department has no control over deliveries. Although measures must be controllable by individuals or departments, the goal is not to "control" employees.

Rather, measures should be used in the following way:

> Performance measurement systems need to be developed not as a system that only serves higher management needs and serve only to control employees' behaviour (coercive formalization) but should serve to support employees do their work better (by providing feedback, identifying problems, revealing improvement opportunities and, help prioritizing action) enabling formalization. (Wouters and Wilderom, 2008).

The goal, therefore, is not to control behavior, but to provide information as a means for improvement and growth that is mutually beneficial to the individual and the firm. Goh (2012) argues that employees should also be permitted to have an active involvement in the formulation of performance measures. Engaging employees will have a positive effect by giving them a sense of ownership in the performance measurement system and help drive performance improvements.

DoD Supply Chain Metrics.

Department of Defense (DoD) supply chain metrics are utilized to monitor DoD supply chain performance based the following five attributes outlined in DoDM 4140.01-V10 (2017):

1. **Materiel Readiness:** The ability of the supply chain to provide materiel needed to support weapon systems in undertaking and sustaining their assigned missions.

2. **Responsiveness:** The ability of the supply chain to respond to customer materiel requests according to priority.

3. **Reliability:** The ability to deliver required materiel support at a time and destination specified by the customer.

4. **Cost:** The amount of supply chain resources required to deliver a specific outcome.

5. **Planning and Precision:** The ability of the supply chain to accurately anticipate customer requirements and plan, coordinate, and execute accordingly.

Thus, all DoD supply chain metrics are in place for the purpose of managing or improving one or more of the above attributes. Moreover, Military Departments and DLA are encouraged to balance performance measures across these five attributes in order to best meet the strategic needs of customers and facilitate performance improvement initiatives (DoDM 4140.01-V10, 2017). This form of strategic performance management, balancing both financial and non-financial measures to improve supply chain performance, is analogous to the balanced scorecard methodology popularized by Kaplan and Norton (1992):

> The balanced scorecard includes financial measures that tell the results of actions already taken, and compliments the financial measures with operational measures on customer satisfaction, internal process, and the organization's innovation and improvement activities—operational measures that are the drivers of future financial performance.

This idea of an all-inclusive view of organizational performance is clearly visible in the structure of DoD supply chain metrics. Although the DoD is not a for-profit entity, financial measures are still valuable for the purposes of reducing costs, improving resource utilization, and operating more efficiently. The balanced scorecard has proven successful across countless organizations since its inception, and is therefore a quality framework for performance management within the DoD (Cooper, Ezzamel, & Qu, 2016).

DoD Supply Chain Metrics Hierarchy.

DoD supply chain metrics fall into one of three categories: Enterprise, Functional, and Program/Project (see Figure 4). Enterprise metrics are cross-functional in nature and describe the overall effectiveness and efficiency of the DoD supply chain. At this level, the focus is on mission results, and information is used to "choose policy directions and make mission decisions." (Vector Research, 1997). Functional metrics support enterprise level metrics by measuring a major process's internal performance (DODM 4140.01 V-10, 2017). At this level, the focus is on unit results, where information is used to manage and improve operations. Finally, Program/Process level metrics are diagnostic in nature and subordinate to functional level metrics. At this level, activity and task information is used to make tactical decisions and execute management directions (Vector Research, 1997).

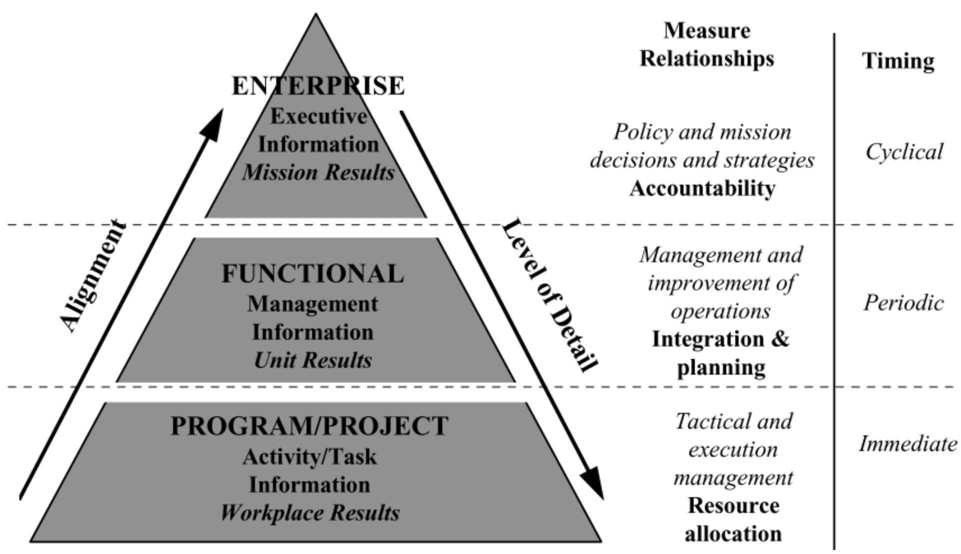


Figure 4. DoD Levels of Performance Measurement (DoD SCM Guide, 2016)

DoD Order Fulfillment Metrics.

Order fulfillment within the DoD is measured under the attribute of Reliability. Within this framework, the specific metric tracked is "Wholesale perfect order fulfillment" (POF) (see Table 1), defined as "The percentage of orders delivered on time with the correct quantity, in the right condition, and with proper documentation." (DoD Supply Chain Metrics Guide, 2016). Individual metrics comprising POF include on-time fill percentage, right quantity percentage, sufficient quality percentage, and proper documentation percentage.

Table 1. DoD Supply Chain Metrics Framework (DoD SCM Guide, 2016)

	Enterprise level metrics		Functional level metrics	
Attribute	**Outcome metrics**	**Diagnostic metrics**	**Inventory Management**	**Distribution Management**
Materiel readiness	♦ Not mission capable (NMC) rates	♦ NMC supply (NMCS) backorders		
Reliability	♦ TDD compliance	♦ Wholesale perfect order fulfillment (POF) ♦ Wholesale supply availability ♦ Materiel denial rates		♦ DLA wholesale supply availability ♦ DLA backorders
Responsive-ness	♦ Customer wait time (organizational level)	♦ Logistics response time (LRT) ♦ Response time effectiveness		♦ DLA logistics response time (LRT)
Cost	♦ Log cost baseline ♦ Value of secondary item inventory	♦ Inventory segmentation of no demand items ♦ Tiered inventory turns ♦ Supply management costs ♦ Supply management cost changes	♦ ERS as a percentage of total inventory ♦ Economic benefit of ERS ♦ CRS as a percentage of total inventory ♦ Secondary item stockage costs and stockage footprint ♦ Inventory dollars with 0–10+ years of no demand ♦ PRS reviewed and sent to disposal	♦ DLA value of inventory ♦ Lateral redistribution ♦ Procurement offset ♦ Lateral Redistribution ♦ Procurement Offset

On time.

An order is considered on time if the logistics response time (LRT), or total time to complete the order from initiation to completion is within the TDD standard.

Right Quantity.

A delivery has the correct quantity if its Materiel Receipt Acknowledgement (MRA) discrepancy code is not "F".

Sufficient Quality.

A delivery has sufficient quality if no discrepancies have been submitted, to include SDRs, TDRs, PQDRs, or a discrepant receipt that does not meet the report criteria for a discrepancy submission.

Proper Documentation.

A delivery has the proper documentation if its MRA discrepancy code is not "B", indicating there is no record of requisition.

Notably, the DoD's measurement of POF is at the enterprise level, using the Logistics Metrics Analysis and Reporting System (LMARS), rather than the functional level within DLA and the military services. According to the DoD Supply Chain Metrics Guide (2016), this is because DLA manages quality and timeliness issues separately, and quality issues are handled at the individual level rather than the aggregate level. In other words, DLA handles discrepant shipments on a case-by-case basis, but does not actively track order fulfillment performance using measures such as POF or order fulfillment quality.

Since DoD order fulfillment performance is measured at the enterprise level with the purpose of "policy and mission decisions", rather than the functional level with the purpose of “management and improvement of operations”, it can be reasonably concluded that order fulfillment performance is not currently measured at the appropriate level to facilitate operational performance improvement. Furthermore, order fulfillment is not controllable at the enterprise level, and therefore the current policy violates an important attribute of effective metric design previously discussed.

Performance Feedback

While measurement is a key aspect of performance management, simply tracking metrics is insufficient to drive operational improvements. An additional component crucial to effective management is performance feedback. Effective performance feedback is a process by which information regarding one’s performance is provided by a peer or supervisor for the purpose of correcting behavior, or letting the individual know where his or her performance stands relative to a given standard (Hattie & Timperley, 2011). Although aggregate measurement of performance metrics provides an indication of overall group or unit performance, the feedback is not specific to the individuals responsible for driving performance. As a result, the feedback may be “confounded by the perceptions of relevance to oneself or to other group members” (Hattie & Timperley, 2007). Thus, individual performance feedback is necessary to ensure group members understand how their performance compares to the overall unit performance.

Levels of Feedback.

According to Hattie and Timperley (2011), Performance feedback can be delivered at one of four levels. The first level, task feedback, pertains to individual task performance, and aims to provide information regarding the outcomes of specific tasks, such as whether work was performed correctly or incorrectly. Feedback at this level is appropriate for simple or routine tasks, and may include information such as the percentage of orders packaged correctly (Balzer, Doherty, & O'Connor, 1989). The second level involves feedback related to the process used to complete the task. Feedback at this level seeks to provide instructions or guidance to increase understanding and improve performance in complex tasks or tasks requiring some degree of skill, such as writing a paper. Third, feedback can be given for the purpose of enhancing self-regulation. Feedback at this level aims to improve the individual's ability to self-evaluate his or her work and increase confidence in performing the task. Fourth, feedback can be directed at the "self" rather than specific behaviors or task outcomes. Feedback at this level is often subjective and unrelated to actual task performance, and includes statements such as "You are a great employee" or "You don't seem to care about your work".

Each of the four levels of feedback are distinct and serve a unique purpose with regard to performance management. For the purposes of this study, task feedback is of primary concern given that the order fulfillment process is comprised of a series of routine tasks, and does not require significant skill to complete correctly.

Task Feedback.

Task feedback is often referred to as "corrective feedback" because it involves information related to the degree with which tasks are being performed correctly. In most cases this form of feedback is provided as a result of errors or performance discrepancies, though it can be used to provide feedback related to positive task accomplishment as well (Hattie & Timperley, 2011). Numerous meta-analyses have demonstrated the efficacy of corrective feedback in improving learning, motor skill acquisition, and task performance (Walberg, 1982; Lysakowski & Walberg, 1982; Tenenbaum & Goldring, 1989). While useful, task feedback has been shown to become less effective as feedback complexity increases (Balzer, Doherty, & O'Connor, 1989). In a study examining the effects of varying feedback complexity, researchers provided students with reading passages and multiple-choice questions. For each incorrect response, students were first provided only the correct answer. In subsequent trials, incorrect answers were discussed, along with each of the four other responses. Each passage was re-read and used to explain why the selected choice was incorrect. Researchers discovered that the feedback that provided only the correct answers resulted in higher subsequent task performance than the complex feedback (Kulhavy et al., 1985). These findings suggest that task feedback is more effective when directed at outcomes (e.g. correct or incorrect) rather than the process used to complete the task.

The demonstrated success of corrective feedback in improving performance may be explained in part using control theory, which posits that behavior is regulated by a person's internal control mechanisms to maintain some preset standard (Carver & Scheier, 1981). When a performance discrepancy is identified as a result of external feedback, the person is motivated to reduce the discrepancy. This system is referred to as a negative-feedback loop because it seeks to reduce or correct a sensed "error" in order to maintain stability (Kluger & DeNisi, 1996; Carver & Scheier, 1981). While goal setting theory suggests that people are motivated to achieve a goal, control theory argues that people are instead motivated to eliminate discrepancies. In fact, research has shown that corrective feedback indicating that performance does not meet standards results in increased effort, whereas feedback indicating that performance meets standards results in reduced or equivalent performance (Kernan & Lord, 1991). This phenomenon, however, is only shown when the standard performance level is the highest possible achievement level, such as a pass/fail task. When feedback is positive but an opportunity exists for increased achievement or self-enhancement, future performance also tends to increase (Kluger & DeNisi, 1996). Thus, both positive and negative feedback may provide opportunity for enhanced performance when performance above a set standard is achievable. Mediating the relationship between feedback type and performance is the resulting level of effort, as illustrated below in Figure 5. When increased effort fails to result in increased performance, people often resort to a task-specific plan to develop greater understanding. This approach subsequently leads to “deeper processing, better retention, and hence a possible learning effect.” (Kluger & DeNisi, 1996).

While feedback from an external agent has been shown to increase learning, it is not without flaws. According to Frese and Zapf (1994), regular feedback may cause the employee to use it as a crutch, and thus expend less effort learning the task. Direct task feedback through discovery, conversely, has been shown to have a superior effect on learning compared to feedback from an external source. In the context of order fulfillment within DLA however, tasks such as packaging, completing documentation, and shipping offer little if any feedback in terms of correctness because errors are often not discovered until shipments are received by the customer. Moreover, customer complaints are typically handled by separate departments and warehouse employees are not made aware of mistakes.

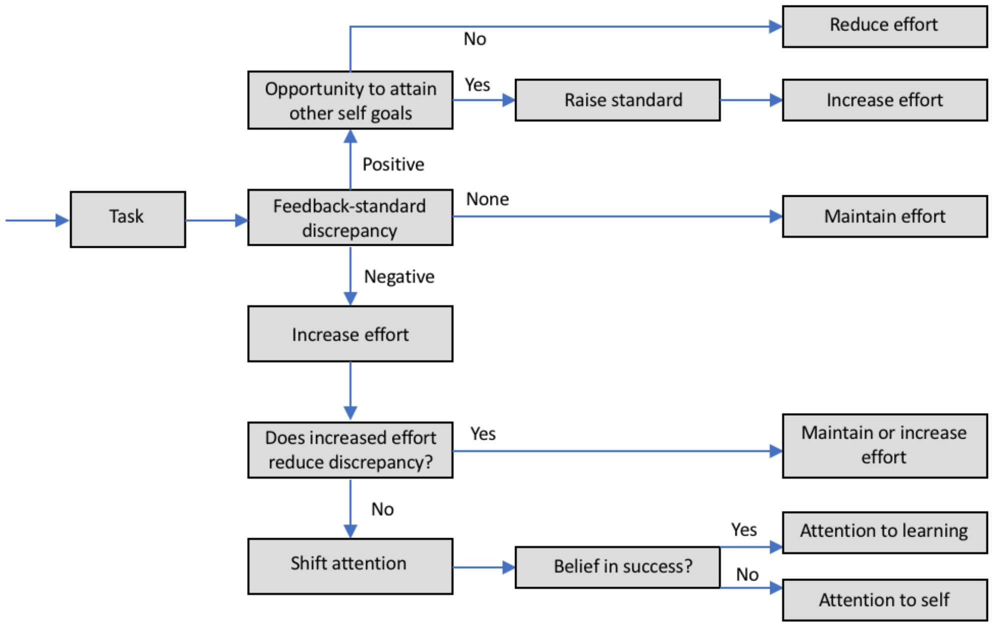


Figure 5. The effect of feedback on task-motivation and performance (Kluger & DeNisi, 1996)

Therefore, feedback from external agents (i.e. supervisors) is necessary for employees to become aware of performance discrepancies. Such feedback should result in increased employee effort and focus when fulfilling orders.

Hypotheses

Given the demonstrated impact of shipping errors in the private sector, and DoD's current management policy for order fulfillment metrics and the demonstrated utility of performance feedback, the following hypotheses are proposed:

Hypothesis 1.

Supply discrepancy reports (SDRs) resulting from shipping and packaging errors are associated with reduced Air Force aircraft availability, increased MICAP hours, increased NMCS rates, and increased cannibalization rates.

Hypothesis 2.

Management of order fulfillment quality at the functional level (i.e. DLA) by providing performance feedback to employees will result in significantly reduced supply discrepancy reports within the Air Force.

III. Methodology

This chapter discusses the research design and methodology used to answer the research questions. Specific variables of interest are examined as well as data collection procedures, instruments, and statistical analyses conducted. Finally, potential threats to validity due to the research design are addressed, as well as the measures in place to reduce these threats.

Research Design

This research utilized a quantitative design to answer the following research questions:

1. To what extent do supply discrepancies impact aircraft readiness?
2. How can performance management help reduce SDRs within Defense Distribution Center Susquehanna, PA?

Quantitative designs are used to examine relationships between two or more variables within a population using various statistical analysis techniques (Babbie, 2010). Quantitative research can be either descriptive in nature or experimental, and data is collected using tools such as surveys, polls, or by manipulating existing data. In the present study, both descriptive and quasi-experimental design methodologies were applied to answer the questions because the two research questions address two distinct problems.

Supply Discrepancy Impact on Readiness

Descriptive Research Design.

Descriptive research designs seek to find out "what is" by analyzing existing data to describe events or draw inferences (Borg & Gall, 1989). These studies can be either qualitative or quantitative, and the method often involves collection of data that can be tabulated along a continuum in numerical form, such as test scores (Knupfer & McLellan, 1996). Descriptive studies often report summary measures of central tendency such as the mean, median, and mode, as well as inferential statistics using methods such as correlation and regression. A descriptive design is used to determine the impact of supply discrepancies on aircraft readiness because SDR data and aircraft maintenance data are readily available, and the variables are not manipulated in any way.

Variables.

The primary variable of interest in this study is order fulfillment quality, measured using supply discrepancy reports. Research question 1 seeks to examine the relationship between order fulfillment quality and four key maintenance metrics: Aircraft Availability, Non-mission Capable Supply Rate, Cannibalization Rate, and MICAP hours. These metrics were selected because of their logical association with supply chain performance, as well as their importance to commanders and relevance to aircraft readiness.

Sample.

Because research question 1 seeks to determine the impact of SDRs on aircraft readiness, the sample consists of operational aircraft maintenance units across the Air Force. The data obtained is then aggregated by MAJCOM and includes Air Combat

Command, Air Mobility Command, and Pacific Air Forces. This sample was collected to ensure adequate representation of Air Force mission sets and geographical locations. At the request of the 635th Supply Chain Operations Wing, Nellis Air Force Base and Kadena Air Base are also sampled individually to assess the impact of SDRs at these specific locations.

Procedures.

Research question 1 is investigated by first collecting SDR data, MICAP data, and aircraft maintenance data from DoD Information Systems. DLA Transaction Services Web Supply Discrepancy Reporting (WebSDR) is used to collect SDR data from each of the sample bases discussed above. DoD WebSDR provides web-based SDR management, enabling new report submission, correction/modification, cancellation, follow-up, requests for reconsideration, and SDR replies. In addition, WebSDR provides users with access to comprehensive reports and custom queries for trend analysis and identification of problems with shipping activities.

Using the query tool in WebSDR, five years of SDR data (2012-2016) is obtained for each of the bases included in this study. SDRs are queried from the "submitter view", meaning that the SDRs were submitted from the sampled bases as a result of discrepancies originating at other sources of supply (e.g. DLA, Boeing, depot, etc.). WebSDR queries are reported via the web-based interface, then exported to Microsoft Excel for analysis. SDR data includes comprehensive information surrounding each SDR occurrence over the sample period, including: submitter, shipper, document number, NSN, discrepancy codes, and dollar value, among others (see Appendix B).

Aircraft and MICAP data are collected using the Logistics Installation and Mission Support – Enterprise View (LIMS-EV) system. LIMS-EV is a Business Intelligence system providing integrated Air Force logistics data from over 60 standalone systems. The system synthesizes data from these systems to provide accurate and well-organized data in the form of dashboards for the purposes of data analysis and reporting (Petcoff, 2010). LIMS-EV offers over a dozen unique applications providing data ranging from supply chain metrics, to vehicle maintenance data. The applications used in the present study were LIMS-EV Weapons System View and LIMS-EV Enterprise Dashboard. LIMS-EV Weapons System View provides comprehensive data on Air Force weapon systems, including aircraft maintenance and operations metrics. LIMS-EV Enterprise Dashboard provides a one-stop shop for Air Force logistics and supply chain data, and is the source for MICAP data in this study.

Like the SDR data, aircraft data is obtained over a five-year period from 2012-2016 for each of the sampled bases. Data includes the following metrics: aircraft availability percentage, non-mission capable supply hours, and number of cannibalizations. Data is again reported via the LIMS-EV web interface, then exported to Microsoft Excel. Prior to analysis, the data underwent preprocessing to ensure the information was consistent with the SDR data and formatted properly for analysis. First, monthly aircraft data was aggregated across all applicable weapon systems at a given installation. For example, Kadena AB operates six unique aircraft mission-design series (MDS), and data is listed in monthly increments for each individual weapon system (see Table 2). Data for each weapon system were combined to provide a single metric capturing the total aircraft availability, NMCS rates, and cannibalizations for each month (see Table 3).

Table 2. Disaggregated Aircraft Data for December 2016 at Kadena AB

Date	*MD*	*MDS*	*Available %*	*NMCS Hrs*	*Canns*
Dec 2016	HH-60	HH060G	75.01	0.00	1
Dec 2016	F-15	F015D	78.14	66.18	1
Dec 2016	KC-135	KC135R	64.24	835.55	8
Dec 2016	E-3	E003C	94.79	0.00	0
Dec 2016	E-3	E003B	38.10	379.68	2
Dec 2016	F-15	F015C	67.11	842.85	26

Table 3. Aircraft Data After Aggregation for July-December 2016

Date	*Total Aircraft*	*Available %*	*NMCS Hrs*	*Canns*
Dec 2016	81	67.54	2,124.26	38
Nov 2016	81	63.46	1,607.09	32
Oct 2016	81	57.70	3,241.89	71
Sep 2016	81	61.72	3,392.08	73
Aug 2016	80	57.24	3,732.68	75
Jul 2016	81	59.23	4,729.67	44

Finally, MICAP data is collected from 2012-2016 for the sampled bases. Data obtained from LIMS-EV Enterprise Viewer provided information surrounding each MICAP over the sample period, including: cause code, document number, MICAP hours, source of supply, NSN, and termination code, among others (see Appendix C). Again, data is reported via the LIMS-EV web interface, then exported to Microsoft Excel. Prior to analysis, MICAP document numbers were matched with SDR document numbers to determine which MICAPs had reported discrepancies.

Data Analysis.

Data obtained from LIMS-EV and WebSDR is analyzed using basic statistical techniques. Independent samples t-tests (with equal variance assumed) are conducted to compare the difference in MICAP hours between MICAPs with reported SDRs, and MICAPs without SDRs across all bases included in the sample. The formula used for t-tests is

$$t = \frac{\bar{x}_1 - \bar{x}_2}{S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \quad , \qquad (1)$$

where $\bar{x}_i$ is the *sample mean*, S_p is *pooled standard deviation*, and n_i is the sample size.

Additionally, linear regression is utilized to assess the relationship between SDRs and aircraft availability, NMCS rates, and cannibalizations. Results of these tests are used to measure the relationship between SDRs and reduced aircraft readiness, and provide justification for the quasi-experiment used to answer research question 2. The formula used for linear regression is

$$y = \alpha + \beta x \quad , \qquad (2)$$

where α is the *y-intercept (constant)* and β is the *slope.*

Threats to Validity.

The primary threat to internal validity with regard to the methodology for research question 1 is the potential for confounding variables. There are dozens of factors that contribute to metrics such as MICAP hours, aircraft availability, and NMCS rates, so it is possible that any relationship found between SDRs and these variables could be due to some other factor not accounted for in the analysis. Possible confounding variables could include other supply chain discrepancies such as Transportation Discrepancy Reports (TDRs) or Product Quality Discrepancy Reports (PQDRs). Additionally, it is possible that aircraft metrics were reported using differing methodologies across bases. Some have questioned the soundness of aircraft maintenance metrics due to the propensity for actions that artificially improve numbers, an act known as "chasing metrics" (Stahl, 2014). Therefore, differences in measurement practices could affect the validity of potential findings.

Effect of Performance Measurement and Feedback on Order Fulfillment Quality

Quasi-experimental Design.

Quasi-experimental designs are similar to true experiments in that they impose an intervention or "treatment" on a sample population, and compare results to a pre-test and/or control group to test for differences. Quasi-experimental designs, however, are typically conducted in real world settings instead of laboratory or other strictly controlled settings, and thus do not randomly assigns participants to intervention and control groups. Instead, groups are selected based on practicality, convenience, self-selection (participants choose), or administrator selection (i.e. by officials, supervisors, teachers, policy makers, etc.) (White and Sabarwal, 2014). Research question 2 utilizes a quasi-experimental design because an intervention (performance measurement) will be imposed on a selected sample and, after a data collection period, results will be compared to a selected control group as well as pre-intervention data. Thus, the study is an interrupted time series design with comparison group.

Variables.

Research question 2 seeks to determine the impact of performance management on order fulfillment quality. Therefore, performance management is implemented as the independent variable (treatment) for the quasi-experiment, and order fulfillment accuracy (SDR rate) is the dependent variable. Specific performance management initiatives will be discussed in further detail below.

Sample.

DLA Distribution is the targeted sample because it is the Air Force's largest supplier of aircraft parts. The Defense Distribution Center Susquehanna, Pennsylvania (DDSP) is chosen as the intervention group because it is DLA's largest distribution center, and therefore provides the largest single source of data over the short timeframe of the experiment. The Defense Distribution Depot San Joaquin, CA (DDJC) is selected as the control group because it is DLA's second largest distribution center, and thus the most equivalent sample in terms of size, mission, and workload.

Procedures.

Research question 2 is investigated by implementing a quasi-experiment at Defense Distribution Susquehanna, PA (DDSP) focused on the implementation of performance measurement and feedback for order fulfillment discrepancies. Prior to establishing performance measurement initiatives, extensive consultation with DDSP leadership (including a site visit) was conducted to determine the best approach to improve supplier performance and reduce SDRs. As a result of these meetings, it was decided to implement two initiatives to emphasize the importance of order fulfillment quality at DDSP:

1. Provide direct feedback to employees and supervisors responsible for committing supply discrepancies.

2. Establish a Key Performance Indicator (KPI) titled "Order Fulfillment Quality" to be actively monitored by leadership and reported to the organization.

The first initiative was implemented on 7 October 2017 after a review from DDSP leadership and union officials for approval. The feedback form was called an *Internal Customer Discrepancy* (ICD) form, and was an existing form in use at DDSP to document employee discrepancies, although it was not previously in use to document supply discrepancies (see Appendix D). The purpose of the form was not to admonish employees for mistakes, but to increase awareness of SDRs and encourage employees to reflect on their individual performance and role in ensuring excellent order fulfillment.

The second initiative required review and approval from DLA Distribution Headquarters (HQ), and thus took significant time to implement. DLA Distribution KPIs are established at DoD level and individual distribution centers are not authorized to deviate from established KPIs, or create local KPIs for internal use only. After a lengthy review process, the new KPI for Order Fulfillment Quality (OFQ) was implemented on 1 Nov 17.

Data Analysis.

Data obtained over the course of the quasi-experiment at DLA was analyzed using statistical process control (SPC) methods. SPC techniques involved Shewhart Charts and CUSUM control charts to detect changes in the average number of monthly SDRs.

Statistical Process Control.

The SDR time series data was first analyzed using a Shewhart chart to detect changes in process control after the introduction of the experimental conditions. A Shewhart chart is a control chart where each point represents a summary statistic computed from either a sample of measurements, or a collection of measurements from a given time frame (NIST, 2008). Shewhart charts are often used in manufacturing settings for the purpose of statistical process control. In the present study, control charts were utilized to detect changes in the average number of monthly SDRs. The three elements of the chart include a graph of the time series data, a central reference line for the process average, and upper and lower control limits (UCL and LCL). The formula used to calculate control limits (CL) is

$$CL = \bar{X} \pm 3\hat{\sigma} \quad , \tag{3}$$

where $\bar{X}$ is the *process average,* and $\hat{\sigma}$ is the *process standard deviation.*

An advantage of the Shewhart chart is that it plots actual values, and is therefore easily interpreted. A major drawback, however, is that minor shifts in the process will often fall within the control limits and go undetected for longer periods of time. An alternative chart used to detect small shifts is the cumulative sum (CUSUM) control chart. A CUSUM control chart is “a plot of the cumulative differences between successive values and a target value” (Stapenhurst, 2005). Rather than plotting each data point independently, these charts show the accumulation of deviations from current and previous values, and are therefore better suited for detecting small shifts in the mean of a process. When the outcome trend is consistent with the average process value, the plot runs randomly along the baseline at zero (Sibanda & Sibanda, 2007). The process will be deemed *out of control* if the upward or downward drift of cumulative deviations exceeds a set boundary. CUSUM charts include two parameters:

1. (K) – A reference value, or allowable slack, specified in sigma units and typically set to one half of the standard deviation. Deviations from the mean must exceed this value in order to be accumulated.
2. (h) – The decision limit specified in sigma units and typically set to 4σ (Stapenhurst, 2005).

The formulas used to calculate lower and upper cumulative sums (S_{Li} and S_{Hi}) are

$$S_{Li} = Min[0, X_i - (\bar{X} - K) + S_{Li-1}] \text{ , and} \quad (4)$$

$$S_{Hi} = Max[0, X_i - (\bar{X} + K) + S_{Hi-1}] \text{ ,} \quad (5)$$

where X_i is the *process measurement* at the i^{th} sample.

S_{Li} and S_{Hi} values are then plotted on the control chart with the decision limits ($\pm h$) given as dotted lines. Cumulative values exceeding *h* are deemed *out-of-control*, and the cumulative sum is then either reset to zero, set to a fast initial response (FIR) value (*h*/2), or, as is the case in the present study, left unchanged.

In addition to the analysis of DDSP SDR trends, analysis of a control group at DDJC was conducted and compared to the findings at DDSP. This will strengthen the study by reducing potential threats to internal validity such as regression to the mean, history, and maturation.

Threats to Validity.

With regard to research question 2, the most obvious design flaw is the lack of randomization due to the nature of the quasi-experimental design. Because the treatment and control groups could not be randomized, it is possible that they differ in some fundamental way that could cause biased findings. For example, the quality of employee training programs may differ between the two organizations, resulting in improved performance from one group over another. An additional threat to internal validity is the potential for design contamination. Since employees from both groups interact with one another on a regular basis, it is possible that key components of the study were shared between groups. If the control group were to implement either of the initiatives from the treatment group in their organization, it would effectively invalidate the comparison analysis. To mitigate the possibility of contamination, the researcher deliberately requested with the treatment group that the study not be shared with the control group.

IV. Analysis and Results

This chapter provides both descriptive statistics and primary findings for the two research questions under investigation. Recall that research question 1 sought to determine what, if any, relationship existed between SDRs and three aircraft maintenance metrics: aircraft availability, not-mission capable supply hours, and cannibalizations. In addition, a comparison of average MICAP hours was conducted between MICAPs with reported SDRs and those without SDRs to determine if any significant difference in MICAP hours existed between the two groups. Also, recall that research question 2 sought to determine whether the trend of SDRs originating at DLA Distribution Susquehanna could be reduced through a targeted performance management intervention.

Research Question 1

Descriptive Statistics.

From 2012 through 2016, Air Combat Command experienced the highest average number of SDRs, NMCS hours, and cannibalizations of the three MAJCOMs sampled. Air Mobility Command experienced the lowest averages, though the sample consisted of only five AMC bases. Between the two individually assessed bases, Kadena AB experienced nearly three times the average number of SDRs compared to Nellis AFB, and nearly twice as many compared to AMC. Conversely, Nellis AFB experienced a much higher average of NMCS hours and cannibalizations compared to Kadena AB and AMC. Complete aircraft maintenance and SDR data for each sampled location are shown below in Table 4.

Table 4. Aircraft and SDR Data by MAJCOM (2012-2016)

Base/MAJCOM	*Avg. Monthly SDRs*	*Avg. NMCS Hrs*	*Avg. NMCS %*	*Avg. AA %*	*Avg. Canns*
ACC	344	33,213	4.01	67.00	593.40
AMC	63	3,006	2.68	66.55	33.13
PACAF	286	12,456	5.10	68.23	215.87
Nellis AFB	33	5,042	N/A	64.04	93.68
Kadena AB	91	2,998	5.06	66.84	59.25

Of the three MAJCOMs sampled, Air Combat Command had the highest incidence of MICAPs over the five-year period with 121,497. Among the 16 bases within ACC, Nellis AFB held the highest frequency of MICAPs at 37,207 (23.4 percent). Pacific Air Forces had the second-highest incidence of MICAPs over the sample period at 75,998, with the greatest frequency occurring at Kadena AB with 18,079 (23.8 percent). Air Mobility Command held the fewest number of MICAPs from 2012-2016 at 32,807, although this was partially due to the exclusion of some AMC bases such as MacDill AFB, Little Rock AFB, McConnell AFB, Fairchild AFB, and Scott AFB. Of the sampled AMC bases, Dover AFB had the highest frequency of MICAPs at 7,609. Complete MICAP data for all three MAJCOMs are listed below in Tables 5-7.

Table 5. Air Combat Command MICAP Data (2012-2016)

Base	*# MICAPs*	*% of total*
Beale AFB	3,625	2.4
Creech AFB	1,406	0.09
Davis-Monthan AFB	12,652	8.1
Grand Forks AFB	280	0.02
Hill AFB	11,098	7.0
Holloman AFB	5,345	3.5
Langley AFB	6,690	4.3
Moody AFB	10,970	6.9
Mountain Home AFB	11,826	7.5
Nellis AFB	37,207	23.4
Offutt AFB	3,086	1.9
Robins AFB	94	0.01
Seymour-Johnson AFB	21,370	13.6
Shaw AFB	14,276	9.1
Tinker AFB	4,093	2.6
Tyndall AFB	14,686	9.4
Total	158,704	100.0

Table 6. Pacific Air Forces MICAP Data (2012-2016)

Base	*# MICAPs*	*% of total*
Andersen AFB	5,717	7.5
Eielson AFB	5,555	7.3
Elmendorf AFB	8,566	11.3
Hickam AFB	3,486	4.6
Kadena AB	18,079	23.8
Kunsan AB	9,323	12.3
Misawa AB	11,213	14.8
Osan AB	10,835	14.3
Yokota AB	3,224	4.2
Total	75,998	100.0

Table 7. Air Mobility Command MICAP Data (2012-2016)

Base	*# MICAPs*	*% of total*
JB Lewis-McChord	7,467	22.7
JB McGuire-Dix-Lakehurst	4,161	12.7
JB Charleston	6,832	20.8
Dover AFB	7,609	23.2
Travis AFB	6,738	20.6
Total	32,807	100.0

With regard to SDRs occurring on MICAP shipments, the overall frequency of SDRs was very low. Among the 267,509 MICAPs across ACC, AMC, and PACAF, a total of just 3,388 SDRs (1.27 percent) were reported. While seemingly small, this is nearly five times the rate of SDRs occurring on all DDSP shipments over the same time frame, suggesting that SDRs are more likely to occur on MICAP shipments. By contrast, of the total 42,646 SDRs reported from 2012-2016, 7.9 percent occurred on MICAP shipments, on par with the rate of MICAPs across all shipments. Complete SDR data across all MAJCOMs are listed in Table 8 below.

Table 8. SDR and MICAP data by MAJCOM (2012-2016)

Base/MAJCOM	*Total SDRs*	*SDRs on MICAPs*	*% of total SDRs*	*% of total MICAPs with SDRs*
ACC	20,639	1,476	7.2	0.93
AMC	4,845	210	4.3	0.64
PACAF	17,162	1,702	9.9	2.24
Nellis AFB	1,975	308	15.6	0.83
Kadena AB	5,445	649	11.9	3.59
Total	42,646	3,388	7.9	1.27

Primary Results.

To adequately answer the first research question, simple linear regression analyses were conducted for each MAJCOM assessing the relationship between SDRs and the following metrics: aircraft availability, not-mission capable hours, and cannibalizations. Additionally, t-tests were conducted to determine whether differences in MICAP hours existed between MICAP shipments that occurred with and without SDRs. Notable findings are listed below, along with tables detailing the results of each analysis. Results of tests for normality and homoscedasticity are listed in Appendices F through J.

Air Combat Command.

A simple linear regression was calculated to predict aircraft availability based on the number of reported SDRs. A significant regression equation was found ($F(1, 58) = 5.268$, $p < .025$), with an R^2 of .083. In other words, 8.3% of the variance in aircraft availability is explained by SDRs. ACC's predicted aircraft availability is equal to 68.9 + -5.8E-5(SDRs) percent.

Assumptions of normality and homoscedasticity were also assessed to verify validity of the model. Results of a Shapiro-Wilk test found that the residuals failed the normality assumption ($w = .023$), but inspection of the residuals scatterplot suggests that the variance is equal across all values (see Figure 6). Although the normality assumption was not met, the sample size was adequately large and the deviation from normal was not extreme.

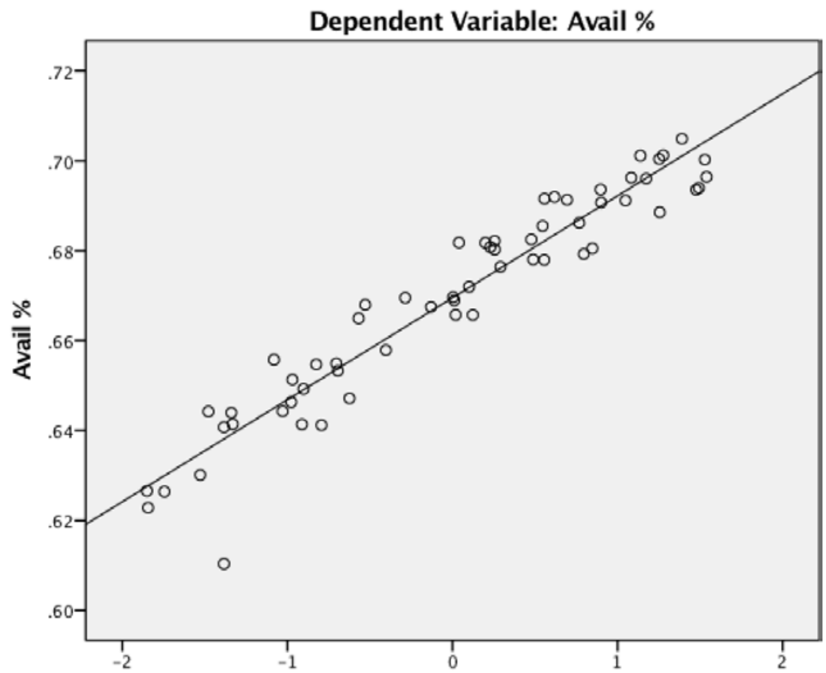


Figure 6. ACC AA regression residuals scatterplot

Results of all three regression analyses calculated for ACC are shown in table 9 below. In addition to aircraft availability, SDRs were found to be a significant predictor of cannibalizations (p =.032). NMCS hours were not found to be significantly impacted by SDRs.

Table 9. ACC SDR Regression Coefficients and Model Results

	SDR coefficients			*Model results*		
Dependent Variables	α	β	$SE(\beta)$	F	p	R^2
Aircraft Availability	-.689	-5.8E-5	.000	5.268	.025*	.083
NMCS Hours	34,714.3	-4.364	4.410	.979	.326	.017
Cannibalizations	520.011	.213	.097	4.808	.032*	.077

*, $p < .05$

Air Mobility Command.

A simple linear regression was calculated to predict NMCS hours based on the number of reported SDRs. A significant regression equation was found ($F(1, 58) = 7.729$, $p < .007$), with an R^2 of .118. Thus, 11.8 percent of the variance in NMCS hours are due to SDRs. AMC's predicted monthly NMCS hours are equal to 2,519 + 7.727(SDRs).

Following the regression analyses, assumptions of normality and homoscedasticity were assessed to verify validity of the model. Results of a Shapiro-Wilk test found that the residuals met the normality assumption (w = .497), and inspection of the residuals scatterplot suggests that the variance is equal across all values (see Figure 7).

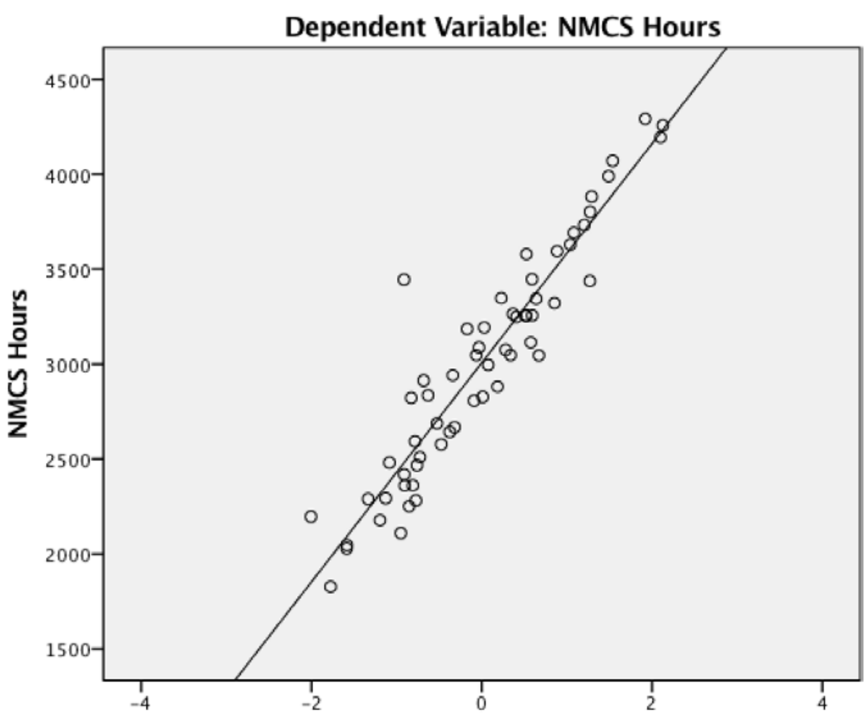


Figure 7. AMC NMCS hours regression residuals scatterplot

Results of all three regression analyses calculated for AMC are shown in Table 10 below. In addition to aircraft availability, SDRs were found to be a significant predictor of cannibalizations (p =.032). Aircraft availability was not found to be significantly impacted by SDRs.

Table 10. AMC SDR Regression Coefficients and Model Results

	SDR Coefficients			*Model Results*		
Dependent Variables	α	β	$SE(\beta)$	F	p	R^2
Aircraft Availability	68.149	-.025	.025	1.015	.318	.017
NMCS Hours	2,519.25	7.727	2.780	7.729	.007*	.118
Cannibalizations	26.138	.111	.051	4.764	.033*	.076

*, $p < .05$

Pacific Air Forces.

A simple linear regression was calculated to predict cannibalizations based on the number of reported SDRs. A significant regression equation was found ($F(1, 58) = 4.455$, $p < .039$), with an R^2 of .071. In other words, 7.1 percent of the variance in cannibalizations can be explained by SDRs. The predicted number of monthly cannibalizations for PACAF is equal to 185 + .141(SDRs).

Following the regression analyses, assumptions of normality and homoscedasticity were assessed to verify validity of the model. Results of a Shapiro-Wilk test found that the residuals met the normality assumption (w = .699), and inspection of the residuals scatterplot suggests that the variance is equal across all values (see Figure 8).

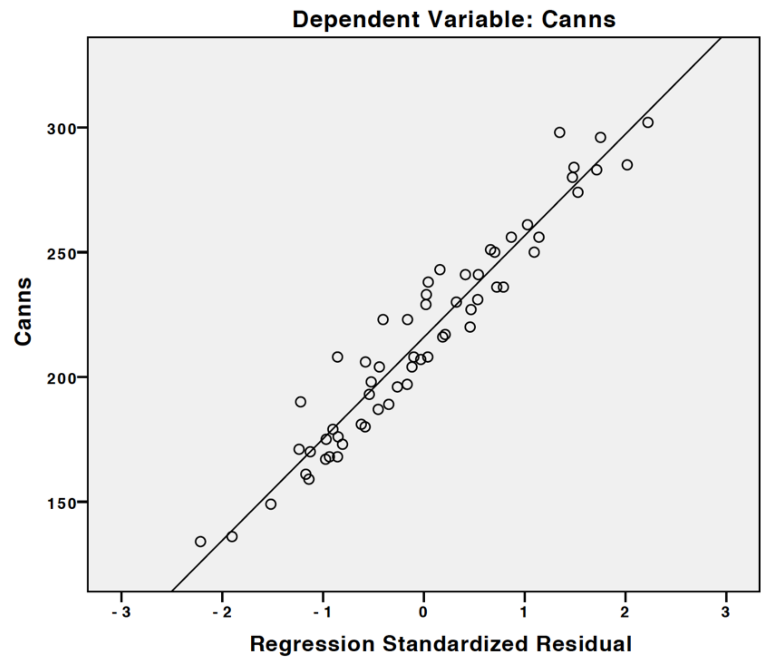


Figure 8. PACAF cannibalizations regression residuals scatterplot

Results of all three regression analyses calculated for PACAF are shown in Table 11 below. Aside from cannibalizations, SDRs were not found to be a significant predictor of aircraft availability (p = .347) or NMCS hours (p =.384).

Table 11. PACAF SDR Regression Coefficients and Model Results

	SDR Coefficients			*Model Results*		
Dependent Variables	α	β	$SE(\beta)$	F	p	R^2
Aircraft Availability	.673	3.297E-5	.000	.899	.347	.015
NMCS Hours	12,863.8	-1.425	2.892	.243	.624	.004
Cannibalizations	185.043	.141	.067	4.455	.039*	.071

*, $p < .05$

Nellis Air Force Base.

A simple linear regression was calculated to predict aircraft availability based on the number of reported SDRs. A non-significant regression equation was found (F(1, 58) = 2.399, p = .127), with an R^2 of .040. Results of this analysis suggest that SDRs are not a significant predictor of aircraft availability at Nellis AFB.

Following the regression analyses, assumptions of normality and homoscedasticity were assessed to verify validity of the model. Results of a Shapiro-Wilk test found that the residuals met the normality assumption (w = .473), and inspection of the residuals scatterplot suggests that the variance is equal across all values (see Figure 9).

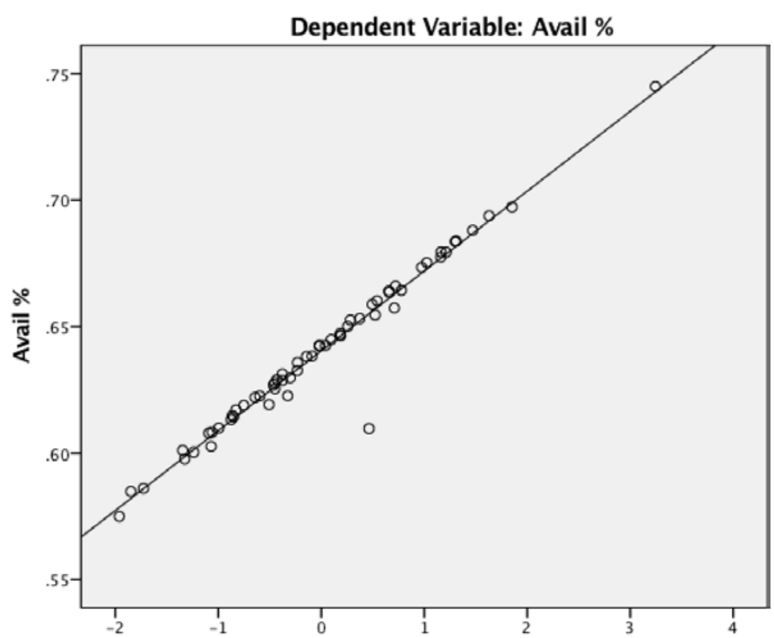


Figure 9. Nellis AFB AA regression residuals scatterplot

Results of all three regression analyses calculated for Nellis AFB are shown in Table 12 below. In addition to aircraft availability, SDRs were not found to be a significant predictor of cannibalizations (p = .362) or NMCS hours (p =.318).

Table 12. Nellis AFB SDR Regression Coefficients and Model Results

	SDR Coefficients			*Model Results*		
Dependent Variables	α	β	$SE(\beta)$	F	p	R^2
Aircraft Availability	.644	.000	.000	2.399	.127	.040
NMCS Hours	5,127.44	-2.558	2.538	1.016	.318	.017
Cannibalizations	92.219	.044	.048	.844	.362	.014

*, $p < .05$

Kadena Air Base.

A simple linear regression was calculated to predict aircraft availability based on the number of reported SDRs. A significant regression equation was found (F(1, 58) = 31.455, p < .000), with an R^2 of .352. In other words, SDRs account for 35.2 percent of the variance in aircraft availability. Kadena's predicted aircraft availability is equal to 71.8 + -.100(SDRs) percent.

Assumptions of normality and homoscedasticity were then assessed to verify the validity of the model. Results of a Shapiro-Wilk test found that the residuals met the normality assumption (w = .998), and inspection of the residuals scatterplot suggests that the variance is equal across all values (see Figure 10).

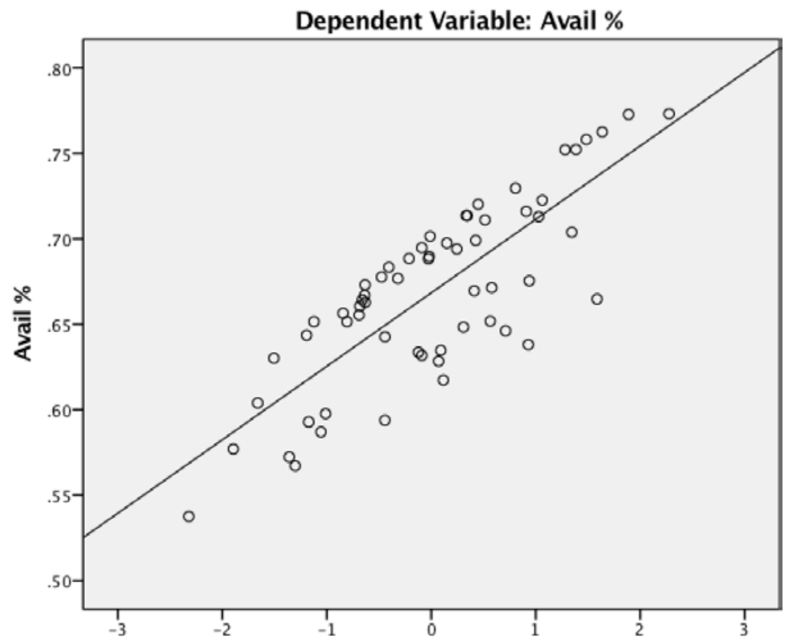


Figure 10. Kadena AB AA regression residuals scatterplot

Results of all three regression analyses calculated for Kadena AB are shown in Table 13 below. In addition to aircraft availability, SDRs were found to be a significant predictor of cannibalizations (p = .036). SDRs were not found to be a significant predictor of NMCS hours (p = .175).

Table 13. Kadena AB SDR Regression Coefficients and Model Results

	SDR Coefficients			*Model Results*		
Dependent Variables	α	β	$SE(\beta)$	F	p	R^2
Aircraft Availability	.718	-.001	.000	31.455	.000*	.352
NMCS Hours	3,197.90	-2.202	1.602	1.890	.175	.032
Cannibalizations	52.295	.077	.036	4.614	.036*	.074

*, $p < .05$

Effect of SDRs on MICAP Hours.

Following the series of regression analyses, independent samples t-tests were calculated using SPSS to determine whether the mean number of monthly MICAP hours at each MAJCOM was different between MICAP shipments with and without an SDR reported. Results indicated that the average MICAP hours per shipment were significantly greater for shipments occurring with an SDR at all MAJCOMs, with the exception of ACC (see Table 14). This finding suggests that SDRs can drastically increase the time taken to fulfill a MICAP order. Tests for normality and homogeneity of variance were also conducted for each test to determine whether the samples met the basic assumptions for analysis. As expected, the MICAP samples were not normally distributed (w = .000), due to the presence of extreme outliers which resulted in the distributions being skewed to the right. Results of Levene tests for equal variance found no difference in variance between the samples (p = .232). Although the normality assumption was violated, the t-test is considered robust against this assumption because the sampling distribution of the test statistic approaches normality with a sufficient sample size, according to the Central Limit Theorem (Edgell and Noon, 1984). Complete results for this analysis are located in Appendices E through I.

Table 14. Difference in MICAP Hours (SDR vs. no SDR)

Base/MAJCOM	*SDR on MICAP?*	*N*	*M*	*SD*	*t*	*p*
ACC	Y	1,342	242.04	846.20	1.086	.277
	N	120,155	221.05	701.00		
AMC	Y	210	106.15	249.27	2.906	.004*
	N	32,597	60.49	226.83		
PACAF	Y	1,623	276.80	782.64	3.250	.001*
	N	74,375	223.50	650.48		
Nellis AFB	Y	323	236.17	539.06	2.469	.014*
	N	36,768	157.14	573.03		
Kadena AB	Y	649	223.37	533.61	2.296	.022*
	N	22,480	180.89	462.54		

*, $p < .05$

Research Question 2

Research question 2 sought to determine whether performance management initiatives implemented at DLA Distribution Susquehanna, Pennsylvania would result in reduced average monthly SDRs over the period of 7 October 2017 to 31 January 2018. To assess the change the impact of the experimental conditions, weekly SDR data from January 2012 to September 2017 was plotted using control charts and compared with SDR data during the experimental period. SDR data from the same time period at DLA Distribution San Joaquin (DDJC) was also plotted using control charts as a comparison group to validate any potential findings. DDSP monthly SDR measurements are listed below in table 15.

Table 15. DDSP Monthly SDR Data (2012-2017)

	Year					
Month	*2012*	*2013*	*2014*	*2015*	*2016*	*2017*
January	386	324	225	255	383	295
February	315	272	225	226	386	225
March	330	340	261	354	402	380
April	321	312	262	308	378	244
May	332	273	218	331	336	279
June	257	313	293	196	345	235
July	259	350	259	322	319	170
August	372	308	295	338	435	365
September	429	331	346	367	375	186
October	376	324	383	350	355	207
November	271	300	235	290	250	181
December	317	353	262	345	298	192

Shewhart Control Chart.

Due to the short timeframe of this study, monthly SDR data was further broken up by week to lengthen the period available for analysis. Weekly SDR data was then examined with a Shewhart control chart created using Microsoft Excel. The mean number of weekly SDRs at DDSP from January 2012 to October 2017 was 71.17 with a standard deviation of 23.05. Due to the high amount of variability in the number of SDRs that occur each week, the LCL for process control was set to 3.4 and the UCL was set to 139.2. Throughout entirety of the observation period, the UCL was exceeded on three separate weeks: 9 Sep 2012; 31 Jul 2016; and 27 Aug 2017. The LCL was not exceeded at any time. Since 3.4 SDRs in a given week is an unrealistic target based on the historical data, the Shewhart chart was not an appropriate tool for assessing the change in weekly SDRs. Close examination of the chart, however, appeared to indicate that weekly SDRs showed a decline beginning the week of 19 November 2017. Data from the year 2017 is shown below in Figure 11, and the complete chart is listed in Appendix J.

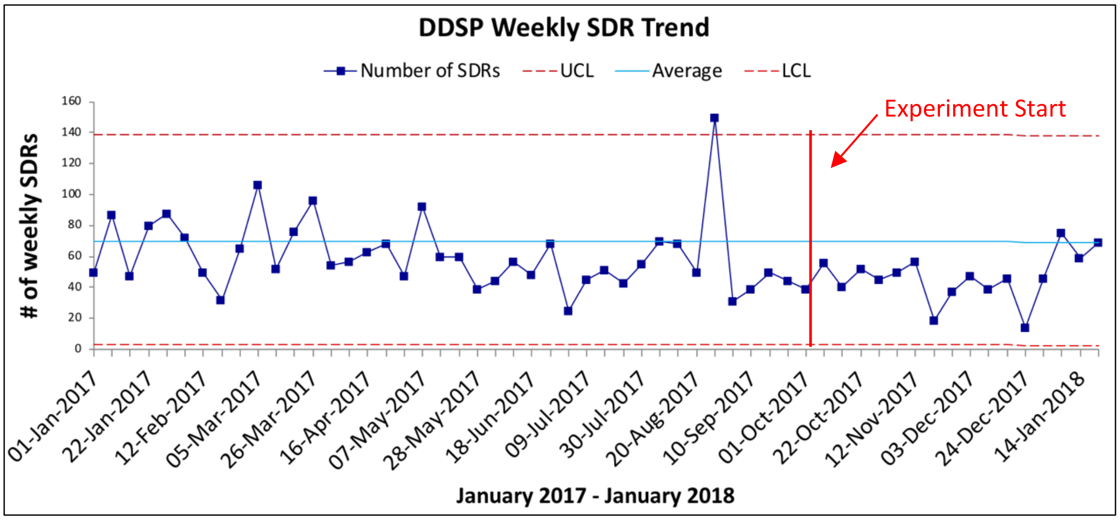


Figure 11. Shewhart Control Chart of Weekly DDSP SDRs

CUSUM Chart.

Due to the inconclusiveness of the Shewhart chart, a CUSUM chart was created next to better detect smaller shifts in the average number of SDRs each week. The first parameter K was set to detect deviations ± 11.53 from the average 71.17 SDRs per week. The second parameter h set the cumulative deviation threshold to ± 92.2. Review of the CUSUM chart revealed that the process exceeded the Upper Cumulative Sum for a period of three weeks in both September and October 2012, as well as a period from 6 March to 19 June 2016 and 31 July to 6 November 2016 (excluding the week of 4

September). It was also found that the process exceeded the Lower Cumulative Sum during the period of 18 May to 1 June 2014, 29 June to 20 July 2014, 16 July to 20 August 2017, and 24 September to 17 December 2017. Data from the year 2017 is shown below in Figure 12, and the complete chart is listed in Appendix K.

Although the CUSUM chart discovered periods deemed out of control prior to the experiment, these periods were short and typically lasted no longer than a few weeks. After the implementation of the SDR performance management experiment at DDSP, SDRs exceeded the lower CUSUM threshold for thirteen consecutive weeks, culminating with a lower cumulative sum value of -348.7. Furthermore, since the start of the experiment the average weekly number of SDRs has dropped to 46.6 from a six-year average of 71.17, and down from 69.2 during the same timeframe the previous year.

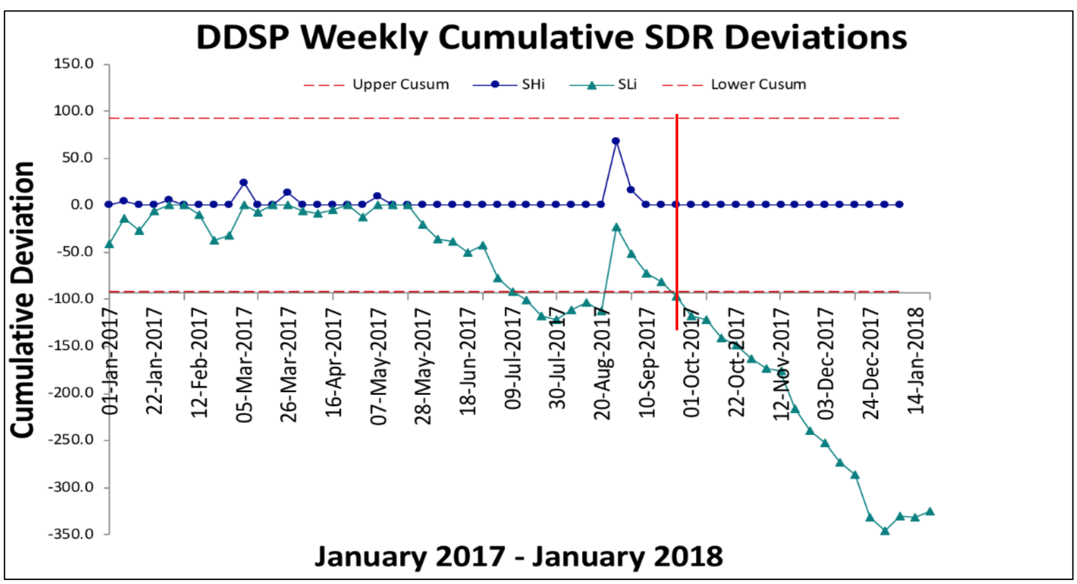


Figure 12. CUSUM Chart of Weekly DDSP SDR Cumulative Deviations

Next, a CUSUM chart was created using SDR data from DDJC to compare with the DDSP chart. The *K* parameter was set to detect deviations ±9.91 from the average 32.09 SDRs per week. The second parameter *h* set the cumulative deviation threshold to ±79.3. Review of the CUSUM chart revealed that the process exceeded the Upper Cumulative Sum threshold from 12 February 2012 to 15 June 2014, but not at any other point following this period. The process exceeded the Lower Cumulative Sum threshold for five consecutive weeks beginning the week of 10 December 2017, but not at any other time prior to this period. 2017 CUSUM data is shown in Figure 13 below, and complete data is listed in Appendix L.

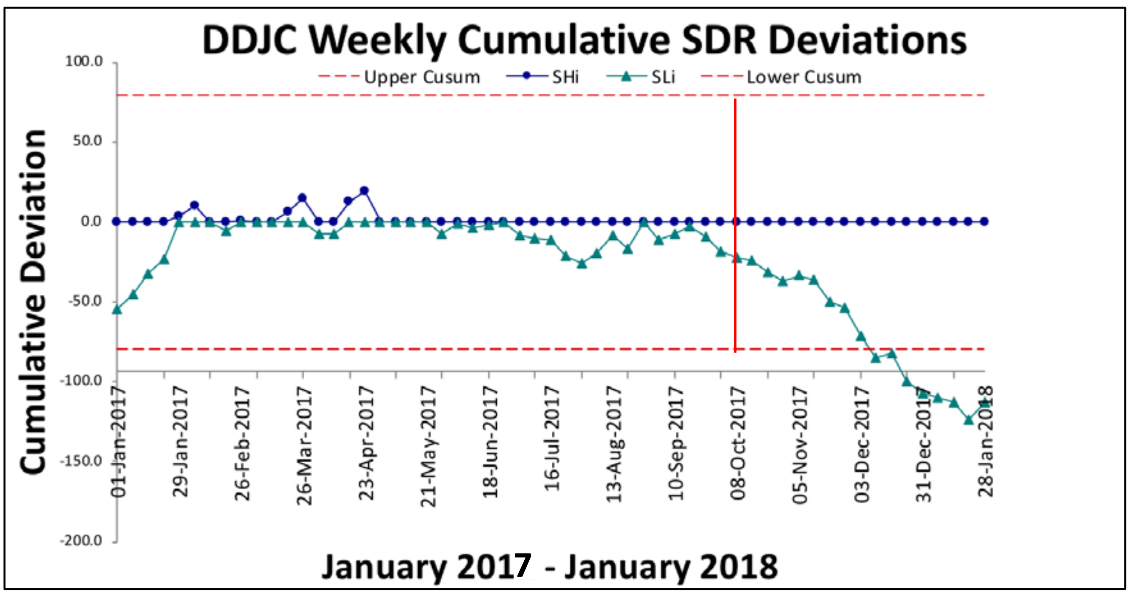


Figure 13. CUSUM Chart of Weekly DDJC SDR Cumulative Deviations

V. Discussion

Shipping and packaging errors are of significant concern to private sector firms due to the proven impact these discrepancies can have on customer satisfaction, loyalty, and profitability. Little research, however, has been done to assess the impact of supply discrepancies in the Air Force, a domain in which SDRs can have significant operational consequences. Thus, the purpose of this study was twofold: first, to determine whether a relationship existed between SDRs and aircraft readiness metrics; and secondly, to assess the impact of performance measurement and employee feedback on SDRs originating from DLA Distribution Susquehanna, PA. The first research question was addressed using simple linear regression to determine whether SDRs were a significant predictor of NMCS hours, cannibalizations, or aircraft availability. Additionally, t-tests were calculated to determine whether significant differences in MICAP hours existed between MICAPs with and without SDRs reported on the shipments.

The second research question was addressed through the implementation of two initiatives at DLA Distribution Susquehanna, PA. The first initiative was directed at employee feedback. Employees responsible for committing SDRs were given direct feedback using a DLA *Internal Customer Discrepancy Form* which detailed the specific discrepancy committed, as well as order information and any additional comments. The intent of these forms was not to be derogatory in nature, but rather to bring greater awareness and mindfulness to employees about common errors in the order fulfillment process. The second initiative was related to formal performance measurement. Prior to this study, aggregate order fulfillment quality metrics were not tracked or reported within

DLA. This initiative sought to bring about greater awareness of the importance of perfect order fulfillment by formally tracking order fulfillment as a performance metric, and communicating the impact of supply discrepancies on military readiness to all employees.

Results of simple linear regression found that of the three metrics tested, SDRs had the greatest impact on cannibalizations. SDRs were significant predictors of cannibalizations across all three sampled MAJCOMs, as well as the individual base sample from Kadena AB. This finding suggests that the occurrence of an SDR on a shipment increases the likelihood that an aircraft part will be cannibalized from another aircraft; a practice akin to "robbing Peter to pay Paul". Additionally, aircraft availability was found to be significantly impacted by SDRs within ACC, and to a much higher degree, Kadena AB. Interestingly, SDRs were found to be a significant predictor of NMCS hours within AMC, but not ACC or Kadena AB where SDRs were predictors of aircraft availability. This is surprising, given that aircraft availability is calculated in part by NMCS rates, and raises the question of whether extraneous variables may have contributed to the findings. Further analysis of this relationship, controlling for extraneous and potentially confounding variables (e.g. NMCM rates) will be required to validate the degree to which SDRs truly affect aircraft availability. Interestingly, all but ACC showed increased MICAP hours resulting from SDRs. This suggests that while SDRs do increase NMCS hours, the number of monthly SDRs that occur at a given location may not be enough to result in a significant impact.

Over the course of the four-month performance management experiment at DDSP, results of the Shewhart chart and CUSUM control charts suggested SDR trends were declining rapidly. Given the high amount of variability in the number of monthly

SDRs and the resulting high standard deviation, SDRs did not exceed the three standard deviation threshold on the Shewhart chart. Close inspection of the chart did however reveal that weekly SDRs fell below the historical average on each of the subsequent weeks after the experiment began. Therefore, the CUSUM chart was developed to quantify the magnitude of the weekly deviations from average. Results of the CUSUM chart showed that weekly SDRs began to drop almost immediately around the start of the experiment. Although the official date of implementation was 7 October, 2017, weekly SDRs dropped below the lower CUSUM threshold beginning the week of 24 September. A possible explanation for this finding could be a form of experiment contamination. Researchers first visited DDSP in late July, and numerous discussions regarding the experiment were held in the weeks leading up to the official implementation. It is possible, therefore, that SDRs began to be scrutinized more closely prior to the formal introduction of ICD forms and performance measurement initiatives.

Interestingly, the DDJC CUSUM chart revealed that SDRs at San Joaquin were also in decline during the study period. Given that both organizations interact on a regular basis, it is possible that the experimental conditions implemented at DDSP were later benchmarked at DDJC. The decline in weekly SDRs did not exceed the lower CUSUM until 10 December, whereas the decline in SDRs at DDSP first exceeded the lower CUSUM on 24 September. The magnitude of the cumulative deviation was also much smaller at DDJC than the deviation that resulted at DDSP. These conditions suggest that any factors contributing to a reduction in SDRs at DDJC would have been implemented several weeks after those implemented at DDSP, and likely to a lesser degree. DDJC leadership confirmed measures had been enacted to reduce SDRs when reached for comment, though specific initiatives were not disclosed. It is also not clear whether DDJC's SDR reduction measures were established in response to the study at DDSP, or if it was an independent initiative.

Implications

Although small, the relationship between SDRs and adverse aircraft metrics was found to be significant in seven of the fifteen regression analyses. Therefore, a measurable reduction in SDRs across Air Force organizations could reasonably improve aircraft readiness. Over the course of the 17-week performance management experiment at DDSP, SDRs dropped by an average of 25 discrepancies per week for a 35 percent reduction. While significant, a reduction in SDRs at DDSP alone is unlikely to result in measurable improvements in Air Force metrics simply because DDSP, while the largest individual supplier, still provides only a fraction of the Air Force's aircraft parts. Thus, in order to create measurable improvements in aircraft readiness, significant improvements in order fulfillment quality across all Air Force suppliers would likely be required. As demonstrated in this study, these improvements can be implemented quickly and at little to no cost. While the overall benefit to aircraft readiness metrics may be small, the ease with which improvements can be made provides a strong case to implement performance management initiatives across additional DLA distribution centers, as well as other DoD suppliers.

Limitations

This research was subject to numerous limitations due to the nature of the experimental design and analysis. Regarding design, this study was a quasi-experiment and thus participants were not randomly assigned, and extraneous variables were not tightly controlled. These conditions made the research susceptible to problems with internal validity including confounding variables and contamination discussed in previous sections. Moreover, we were unable to directly implement and control the experimental initiatives due to geographical separation from DDSP. As a result, the research relied on the oversight of DDSP employees for implementation. The impact of this limitation was minimized by maintaining regular contact with the SDR leadership team at DDSP, and by providing clear guidance and intent throughout the experiment.

An additional limitation was the implementation of ICD feedback forms, which was fraught with challenges. Originally planned for a start date of 1 September 2017, implementation did not begin until 7 October 2017 due to delays in gaining union approval and training supervisors on proper usage. Additionally, there was an unplanned change in supervisors shortly after implementation which resulted in ICD forms being directed to the incorrect individuals for approximately one week. These issues could have potentially weakened the effectiveness of the initiative and limited the overall improvement in SDRs that was found.

From an analysis standpoint, the greatest limitation of this study was the lack of statistical controls for the simple linear regression analyses. Without such controls, it is not possible to draw causal relationships between SDRs and aircraft readiness metrics. The decision to omit these controls was necessary due to time constraints and the large scope of this project. To ensure the statistical relationships found were solely due to SDRs, some possible controls for future consideration should include Transportation Discrepancy Reports (TDRs) and Product Quality Discrepancy Reports (PQDRs).

Suggestions for Future Research

Future research should seek to further investigate the relationship between SDRs and aircraft metrics using controls to draw stronger causal conclusions. At a minimum, other discrepancies such as TDRs and PQDRs should be investigated to first assess their individual impact on aircraft metrics, then to be used as statistical controls for further SDR regression analyses. Additionally, since SDRs impacted some locations more than others, it is important for future research to investigate factors specific to certain locations that may moderate the impact, such as geographic distance from suppliers or types of aircraft assigned.

Although the present study focused only on aggregate SDR totals, future research could investigate the effect of certain SDR types on aircraft readiness—such as shortages, wrong material, and misdirected shipments. This would be useful in determining which SDR types are most detrimental since some types, such as overages, are unlikely to result in negative impact.

Finally, future research could examine the dollar savings that could result from improved order fulfillment quality across DoD suppliers.

Appendix A. SF-364 Report of Discrepancy

REPORT OF DISCREPANCY (ROD)

☐ SHIPPING ☐ PACKAGING

1. DATE OF PREPARATION

2. REPORT NUMBER

3. TO *(Name and address, include ZIP Code)*

4. FROM *(Name and address, include ZIP Code)*

5a. SHIPPER'S NAME

5b. NUMBER AND DATE OF INVOICE

6. TRANSPORTATION DOCUMENT NUMBER *(GBL, Waybill, TCN, etc.)*

7a. SHIPPER'S NUMBER *(Purchase order/Shipment, Contract, etc.)*

7b. OFFICE ADMINISTRATION CONTRACT

8. REQUISTIONER'S NUMBER *(Requisition, Purchase Request, etc)*

9. SHIPMENT, BILLING, AND RECEIPT DATA				10. DISCREPANCY DATA				11. AC[2] TION CODE
NSN/PART NUMBER AND NOMENCLATURE (a)	UNIT OF ISSUE (b)	QUANTITY SHIPPED/ BILLED (c)	QUANTITY RECEIVED (d)	QUAN-TITY (a)	UNIT PRICE (b)	TOTAL COST (c)	CODE[1] (d)	

12. REMARKS *(Continue on separate sheet of paper if necessary)*

[1]DISCREPANCY CODES		[2] ACTION CODES
CONDITION OF MATERIAL C1 — In condition other than that indicated on release/ receipt document C2 — Expired shelf life C3 — Damaged parcel post shipment SUPPLY DOCUMENTATION D1 — Not received D2 — Illegible or mutilated D3 — Incomplete improper or without authority *(Only when receipt cannot be properly process)* MISDIRECTED MATERIAL M1 — Addressed to wrong activity OVERAGE/ DUPLICATE SHIPMENTS O1 — Quantity in excess of that on receipt document O2 — Quantity in excess of that requested *(Other than unit of issue pack)* O3— Quantity duplicates shipment PACKING DISCREPANCY P1 — Improper preservation P2 — Improper packing P3 — Improper marking P4 — Improper unitization	PRODUCT QUALITY DEFICIENCIES Q1 — Deficient material (Applicable to Grant Aid and FMS shipments only) SHORTAGE OF MATERIAL S1 — Quantity less than that on receipt document S2 — Quantity less than that requested *(Other than unit of issue pack)* S3 — Non-receipt of parcel post shipments ITEMS TECHNICAL DATA MARKINGS *(i.e, Name Plates, Log Books, Opening Handbooks, Special Instructions, etc.)* *T1* — Missing T2 — Illegible or mutilated T3 — Precautionary operational markings missing T4 — Inspection data missing or incomplete T5 — Serviceability operating data missing or incomplete T6 — Warranty data missing WRONG ITEM *(Identify requested item as a separate copy in item 9 above)* W1 — Incorrect item received W2 — Unacceptable substitute OTHER DISCREPANCIES Z1 — See remarks	1A— Disposition instructions requested *(Reply on reverse)* 1B— Material being retained *(See remarks)* 1C— Supporting supply documentation requested 1D— Material still required expedite shipment *(Not applicable to FMS)* 1E— Local purchase material to be returned at supplier's expense unless disposition instructions to the contrary are received within 15 days *(Reply on reverse) (Not applicable to FMS)* 1F— Replacement shipment requested *(Not applicable to FMS)* 1G— Reshipment not required. Item to be re-requisitioned. 1H— No action required. Information only 1Z— Other action requested *(See remarks)*

13. FUNDING AND ACCOUNTING DATA

14a. TYPED OR PRINTED NAME, TITLE, AND PHONE NUMBER OF PREPARING OFFICAL

14b. SIGNATURE

15. DISTRIBUTION ADDRESSEES FOR COPIES

STANDARD FORM 364 (REV. 2-80)

Appendix B. Sample WebSDR Report Data

Web Control Number	Major Command Submitter	Submitter	Major Command Action Activity	Action Activity - Current	Owner/ Manager	Shipper	Customer Control Number	Document Number	Suffix	NSN
20120030679	1C	FB4852		AN5	SMS	AN5		FB485212440256		2840014465313
20120130708	1C	FB4852		GSA				FB485220050115		8040008226430
20120181175	1C	FB4852		AQ5	SMS	AQ5		FB485213492371		6695014262107
20120191249	1C	FB4852		AN5	SMS	AN5		FB485213060227		6680010638878
20120251769	1C	FB4852		AQ5	SMS	AQ5		FB485213260449		5305013473677
20120251751	1C	FB4852		AQ5	SMS	AQ5		FB485213200024		1660010188718
20120331015	1C	FB4852	DC	SG2	FGZ	SG2		FB485213190041		1630013352580
20120450724	1C	FB4852	DC	SDD	FLB			FB485211510100		1680013991284
20120650694	1C	FB4852		AQ5	SMS	AQ5		FB485213481445		5331002882036
20120671087	1C	FB4852		AQ5	SMS	AQ5		FB485213390148		1660015449179
20120671144	1C	FB4852		AQ5	SMS	AQ5		FB485213390150		1660015449179
20120671149	1C	FB4852		AN5	SMS	AN5		FB48521277	A	1710010174463
20120671151	1C	FB4852		AN5	SMS	AN5		FB48521160	B	3040013204260
20120671269	1C	FB4852		AQ5	SMS	AQ5		FB485213470074		4820010602543
20120671274	1C	FB4852		AQ5	SMS	AQ5		FB485212844241		5315013191218
20120671276	1C	FB4852		AQ5	SMS	AQ5		FB485213460058		5330011684475
20120671280	1C	FB4852		AQ5	SMS	AQ5		FB485220480097		5340013200707
20120671283	1C	FB4852		AQ5	SMS	AQ5		FB485213493865		5365012350908
20120720405	1C	FB4852		AQ5	SMS	AQ5		FB485220190099		5935013044156
20120720577	1C	FB4852		AN5	SMS	AN5		FB485220380195		5975014622829
20120751138	1C	FB4852	DS	SMS	SMS			FB485220668467		6850014649152
20120751156	1C	FB4852	DC	SDD	FLB			FB485212690397		5895015114600
20120830725	1C	FB4852		AN5	SMS	AN5		FB485220740023		1095001584304
20120830803	1C	FB4852		AQ5	SMS	AQ5		FB485220390031		4240010155194
20120830848	1C	FB4852		AQ5	SMS	AQ5		FB485220100718		4820010602003
20120830902	1C	FB4852		AQ5	SMS	AQ5		FB485213470893		5935013040463
20120860351	1C	FB4852		AQ5	SMS	AQ5		FB48522072	A	5961014239467
20120860517	1C	FB4852		GSA				FB48522059	A	8030012905139
20120880766	1C	FB4852		DL2				FB485220618036		1710000545103
20120890510	1C	FB4852		AQ5	SMS	AQ5		FB485220670059		3040010070113
20120900996	1C	FB4852		DVX				FB485220468096		8030001806222
20120940701	1C	FB4852		AQ5	SMS	AQ5		FB485220240382		9330013113547
20120940703	1C	FB4852		AQ5	SMS	AQ5		FB485213480998		5310004582115
20120940714	1C	FB4852		AQ5	SMS	AQ5		FB485213530072		5340013476583
20120940721	1C	FB4852		AQ5	SMS	AQ5		FB485220590013		1560013557214
20120940725	1C	FB4852		AQ5	SMS	AQ5		FB485220580066		5340013238060
20120970541	1C	FB4852		FLZ	FLZ			FB485220650311		1005010463536
20121000857	1C	FB4852		AQ5	SMS	AQ5		FB485220940248		1005005703737
20121000889	1C	FB4852		F01				FB485220790011		1560013845759
20121030714	1C	FB4852		AQ5	SMS	AN5		FB485211520008		1630013359317
20121030760	1C	FB4852		AQ5	SMS	AQ5		FB485220740331		3120003441495
20121030766	1C	FB4852		GSA				FB485220740661		5120013303822
20121030886	1C	FB4852		AQ5	SMS	AQ5		FB48522074	A	5325011063950
20121030890	1C	FB4852		AN5	SMS	AN5		FB485220890177		5330014652601
20121030911	1C	FB4852		AN5	SMS	AN5		FB485213410085		5945012854514
20121030941	1C	FB4852		AQ5	SMS	AQ5		FB485220091058		6680011107710
20121030943	1C	FB4852		F01				FB485220790008		1560002838125
20121031077	1C	FB4852		AQ5	SMS	AQ5		FB485212300218		5925014249718

Appendix C. Sample MICAP Report Data

Base	MICAPCauseCd	MICAPCondi	MICAPDocumentN	MICAPHou	MDS	NSNLink	SourceofSupply	MICAPTerr
NELLIS AFB	A	E	FB485220278022	99		668501520	FHZ	1
NELLIS AFB	A	F	FB485213268005	415	HH060G	156001391	FLZ	
NELLIS AFB	A	F	FB485220238034	43	KC135R	661099732	FHZ	3
NELLIS AFB	A	F	FB485220248000	13	E003C	661501162	FHZ	1
NELLIS AFB	A	F	FB485220258001	33	E003C	599600982	FHZ	3
NELLIS AFB	A	F	FB485220258029	45	E003C	599801166	FLZ	1
NELLIS AFB	A	F	FB485220258036	90	KC135R	165000676	FHZ	1
NELLIS AFB	A	F	FB485220268000	12	B001B	669501535	FHZ	3
NELLIS AFB	A	G	FB485212788009	704	F016D	156001367	FGZ	1
NELLIS AFB	A	G	FB485212980497	617	F016D	5340P16F2191-18		1
NELLIS AFB	A	G	FB485213258013	75	F015D	100501093	FLZ	
NELLIS AFB	A	G	FB485220058009	94	F016C	165001040	FHZ	3
NELLIS AFB	A	G	FB485220128026	8	A010C	284001051	FHZ	
NELLIS AFB	A	G	FB485220128026	107	A010C	284001051	FHZ	1
NELLIS AFB	A	G	FB485220188017	186	HH060G	161501397	FLZ	3
NELLIS AFB	A	G	FB485220238010	49	B001B	165001184	FHZ	1
NELLIS AFB	A	G	FB485220238015	49	B001B	165001184	FHZ	1
NELLIS AFB	A	G	FB485220258001	1	E003C	599600982	FHZ	
NELLIS AFB	A	G	FB485220268000	1	B001B	669501535	FHZ	
NELLIS AFB	A	G	FB485220288006	75	F015C	482000346	FGZ	3
NELLIS AFB	A	L	FB485220105318	47		596001550	FGZ	1
NELLIS AFB	A	R	FB485211250238	108		598501499	FLZ	0
NELLIS AFB	A	R	FB485220058008	161		598501499	FLZ	3
NELLIS AFB	A	R	FB485220122134	122		599801383	FGZ	1
NELLIS AFB	A	R	FB485220128000	127		598501499	FLZ	1
NELLIS AFB	A	W	FB485220200241	66		283501479	FGZ	0
NELLIS AFB	B	G	FB485220040139	0	F015C	100501093	FLZ	4
NELLIS AFB	B	W	FB485220230144	48		283501561	FGZ	3
NELLIS AFB	B	W	FB485220258022	49		283501561	FGZ	6
NELLIS AFB	C	E	FB485212640221	588		284001325	FHZ	1
NELLIS AFB	C	E	FB485213420089	83		284001479	FHZ	1
NELLIS AFB	C	E	FB485213560445	110		299501519	FHZ	1
NELLIS AFB	C	E	FB485213560451	110		299501519	FHZ	1
NELLIS AFB	C	E	FB485213620110	37		299501519	FHZ	1
NELLIS AFB	C	E	FB485213630296	60		284001371	FHZ	1

Appendix D. Internal Customer Discrepancy Form

Internal Customer Discrepancy Form

Work Area Found		Type of Audit	Select...
Employee Reporting			
Audit Date			
Employee that Made Error			
Supervisor			
Email 1		Email 3	
Email 2			
Discrepancy Origin	Discrepancy Code	PCN/CCN	
	Select...		
REQ QTY	REQ UI	REQ NSN	REQ UNIT COS
ACT QTY	ACT UI	ACT NSN	ACT UNIT COS
RECEIPTED BY		VERIFIED BY	

COMMENTS

1	2	3
To Database	**To Supervisors**	**New Audit**

***Ctrl P to Print**

Employee Signature ______________________________ Date ____________

Supervisor Signature ______________________________ Date ____________

Additional Comments:

For technical support please contact DDSPKnowledgemgmt@dla.mil

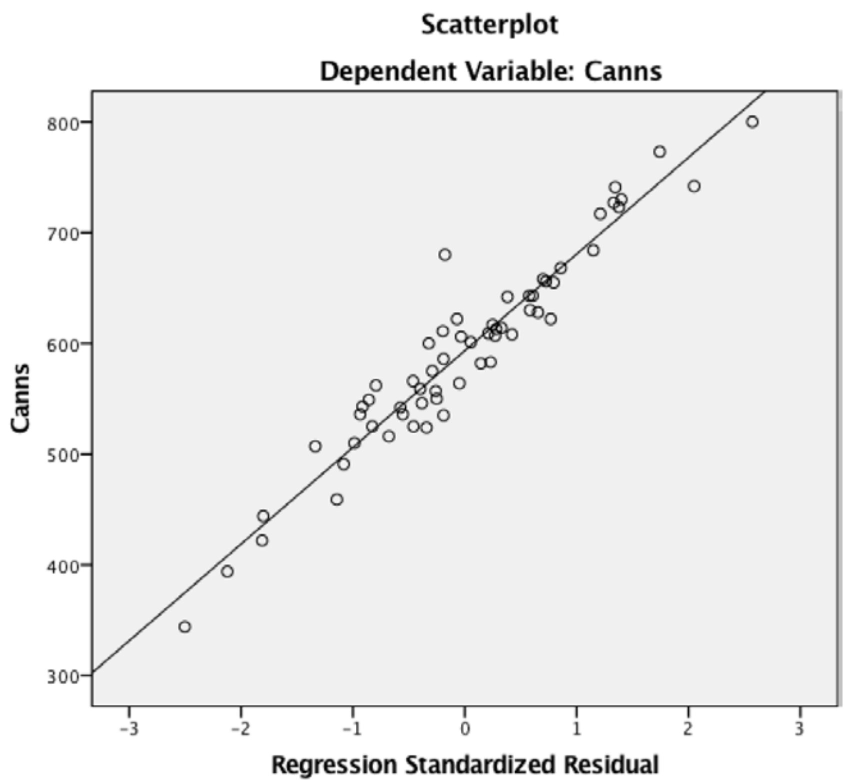
Appendix E. ACC Tests of Normality and Heteroscedasticity

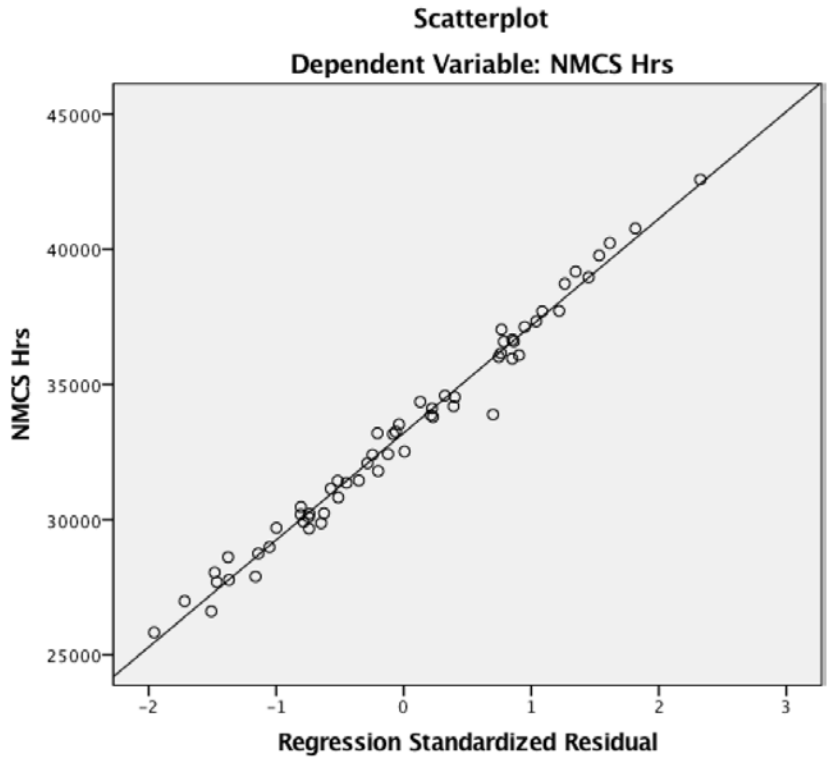
Tests of Normality

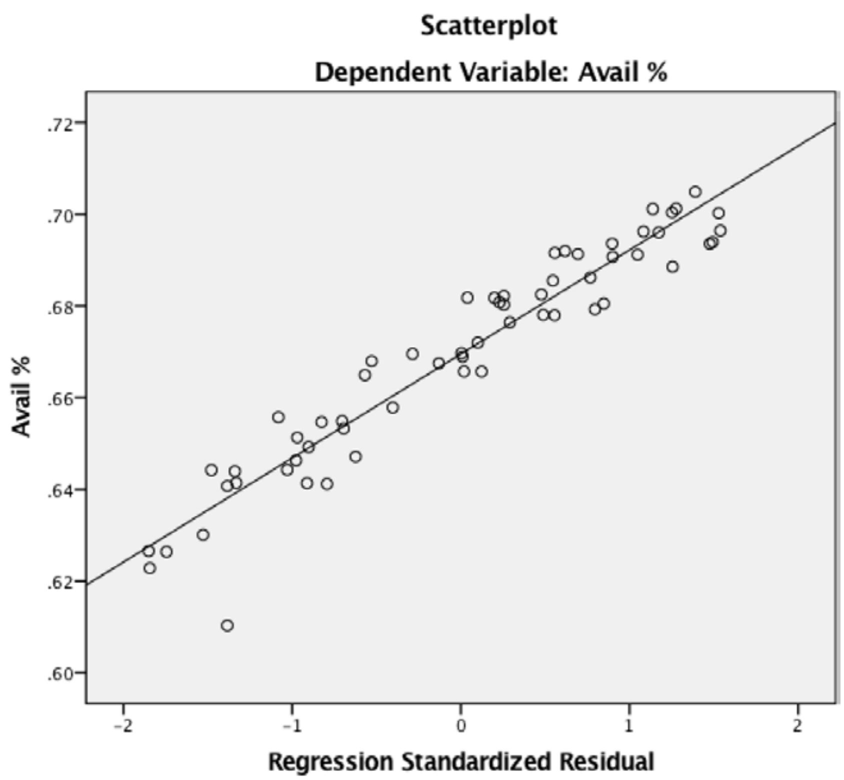
	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Avail %	.123	60	.024	.948	60	.013
NMCS Hrs	.074	60	.200*	.981	60	.452
NMCS %	.077	60	.200*	.980	60	.425
Canns	.071	60	.200*	.985	60	.651
SDRs	.089	60	.200*	.931	60	.002

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction




Scatterplot
Dependent Variable: NMCS Hrs
NMCS Hrs
45000
40000
35000
30000
25000
-2
-1
0
1
2
3
Regression Standardized Residual


Scatterplot
Dependent Variable: Avail %
Avail %
.72
.70
.68
.66
.64
.62
.60
-2
-1
0
1
2
Regression Standardized Residual

Tests of Normality

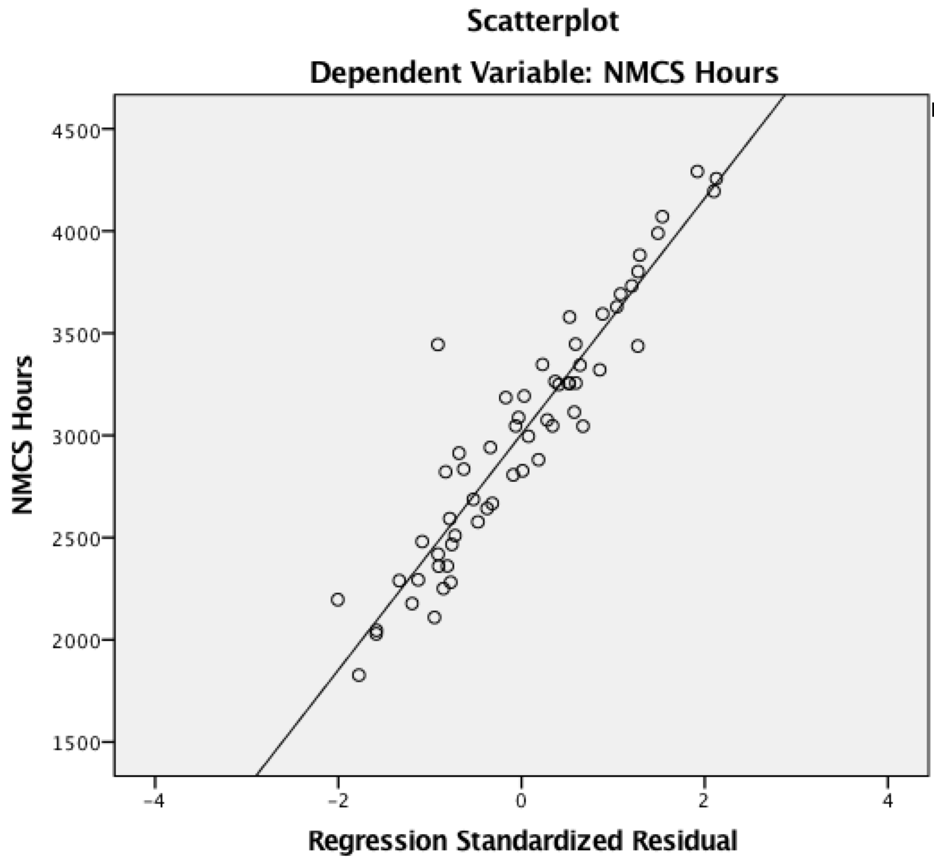
	SDR yes/no	Kolmogorov-Smirnov[a] Statistic	df	Sig.	Shapiro-Wilk Statistic	df	Sig.
MICAPHours	No	.377	120155	.000			
	Yes	.388	1342	.000	.240	1342	.000

a. Lilliefors Significance Correction

Test of Homogeneity of Variance

		Levene Statistic	df1	df2	Sig.
MICAPHours	Based on Mean	1.431	1	121495	.232
	Based on Median	1.540	1	121495	.215
	Based on Median and with adjusted df	1.540	1	121226.290	.215
	Based on trimmed mean	1.635	1	121495	.201

Appendix F. AMC Tests of Normality and Heteroscedasticity

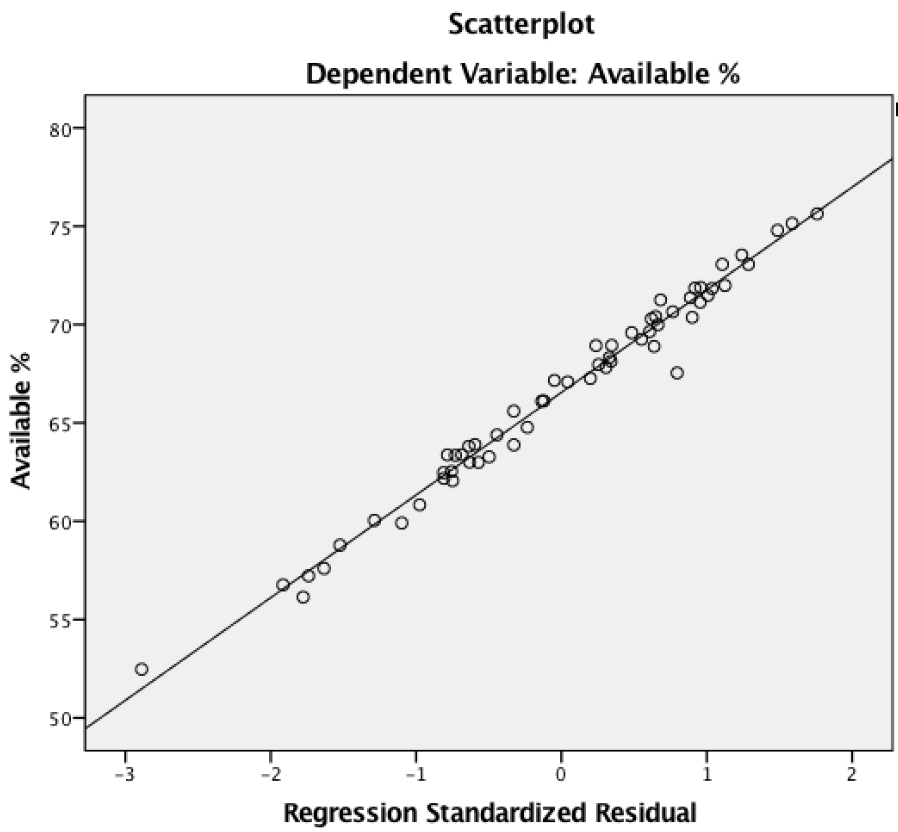


Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Unstandardized Residual	.088	60	.200*	.982	60	.497
Standardized Residual	.088	60	.200*	.982	60	.497

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

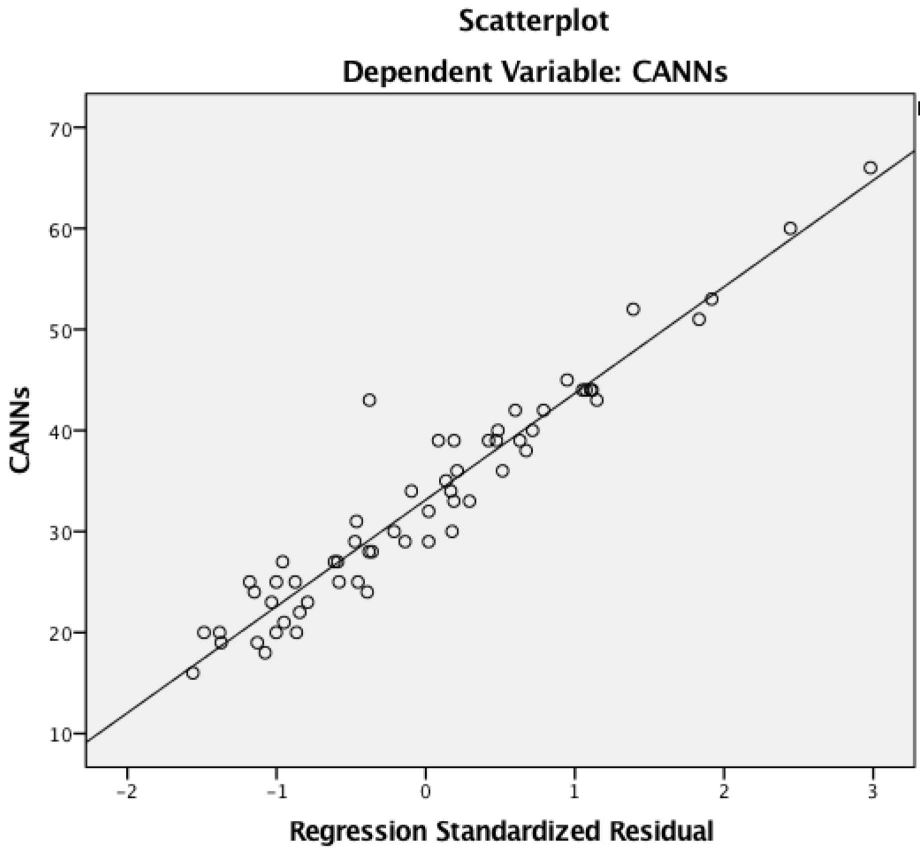


Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Unstandardized Residual	.096	60	.200*	.968	60	.110
Standardized Residual	.096	60	.200*	.968	60	.110

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction



Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Unstandardized Residual	.092	60	.200*	.958	60	.039
Standardized Residual	.092	60	.200*	.958	60	.039

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

Tests of Normality

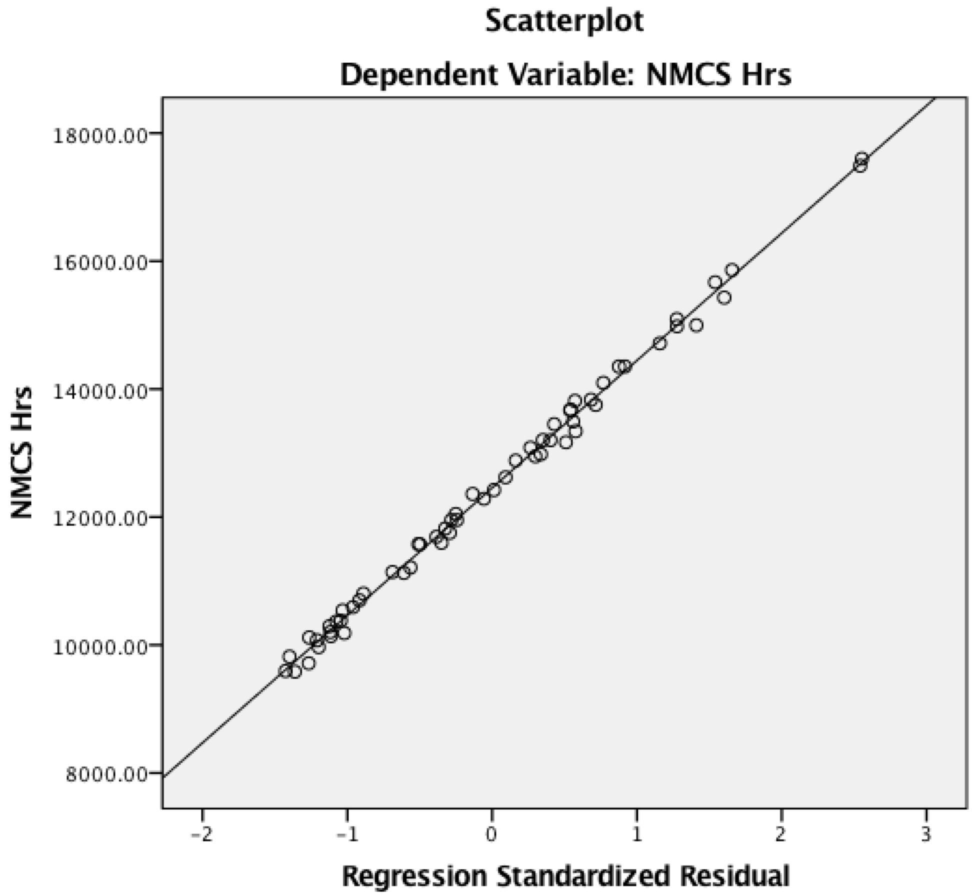
	SDRYesNo	Kolmogorov-Smirnov[a] Statistic	df	Sig.	Shapiro-Wilk Statistic	df	Sig.
MICAPHours	No	.395	32597	.000			
	Yes	.335	210	.000	.391	210	.000

a. Lilliefors Significance Correction

Test of Homogeneity of Variance

		Levene Statistic	df1	df2	Sig.
MICAPHours	Based on Mean	9.938	1	32805	.002
	Based on Median	5.948	1	32805	.015
	Based on Median and with adjusted df	5.948	1	32799.791	.015
	Based on trimmed mean	6.714	1	32805	.010

Appendix G. PACAF Tests of Normality and Heteroscedasticity

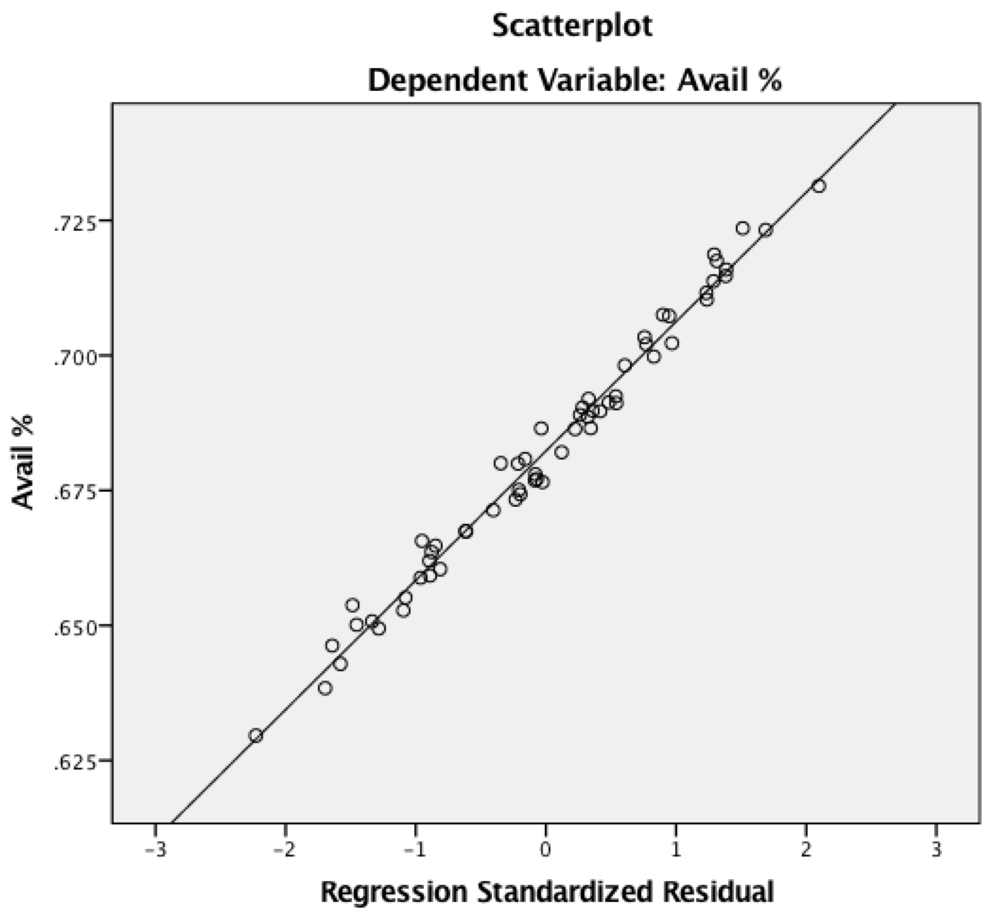


Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Unstandardized Residual	.098	60	.200*	.952	60	.019
Standardized Residual	.098	60	.200*	.952	60	.019

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

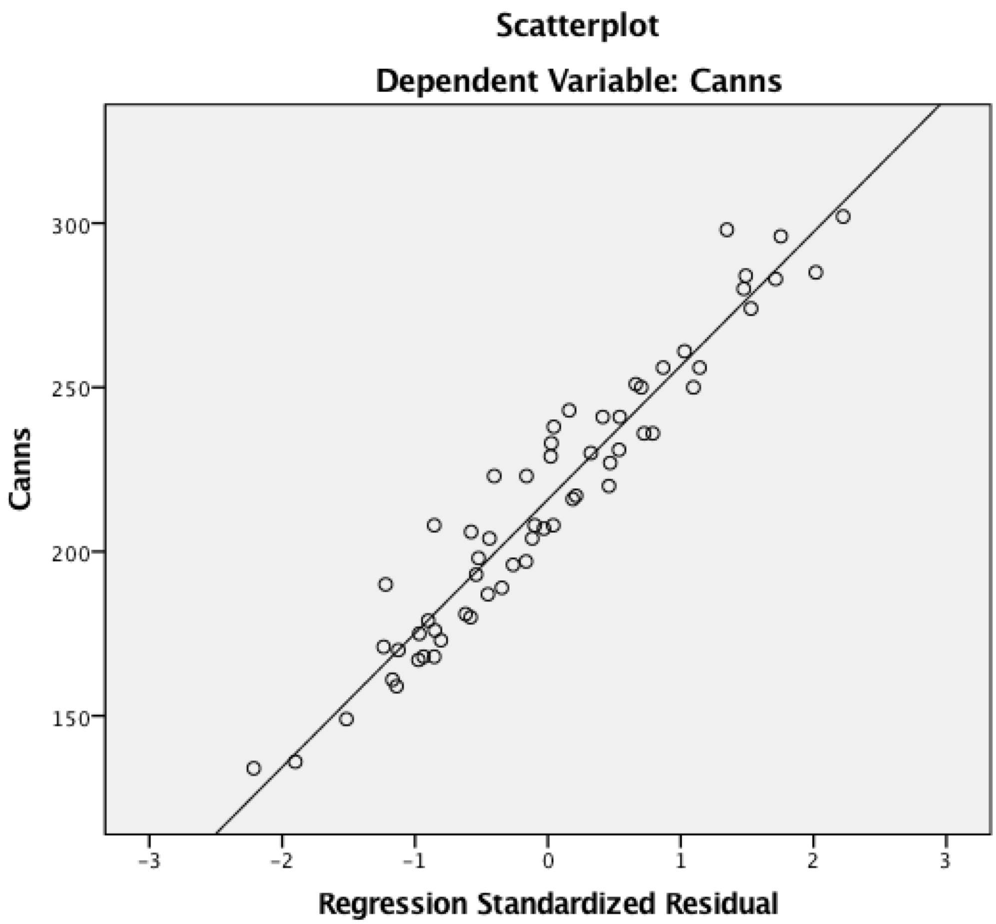


Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Unstandardized Residual	.077	60	.200*	.983	60	.590
Standardized Residual	.077	60	.200*	.983	60	.590

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction



Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Unstandardized Residual	.066	60	.200*	.986	60	.699
Standardized Residual	.066	60	.200*	.986	60	.699

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

Tests of Normality

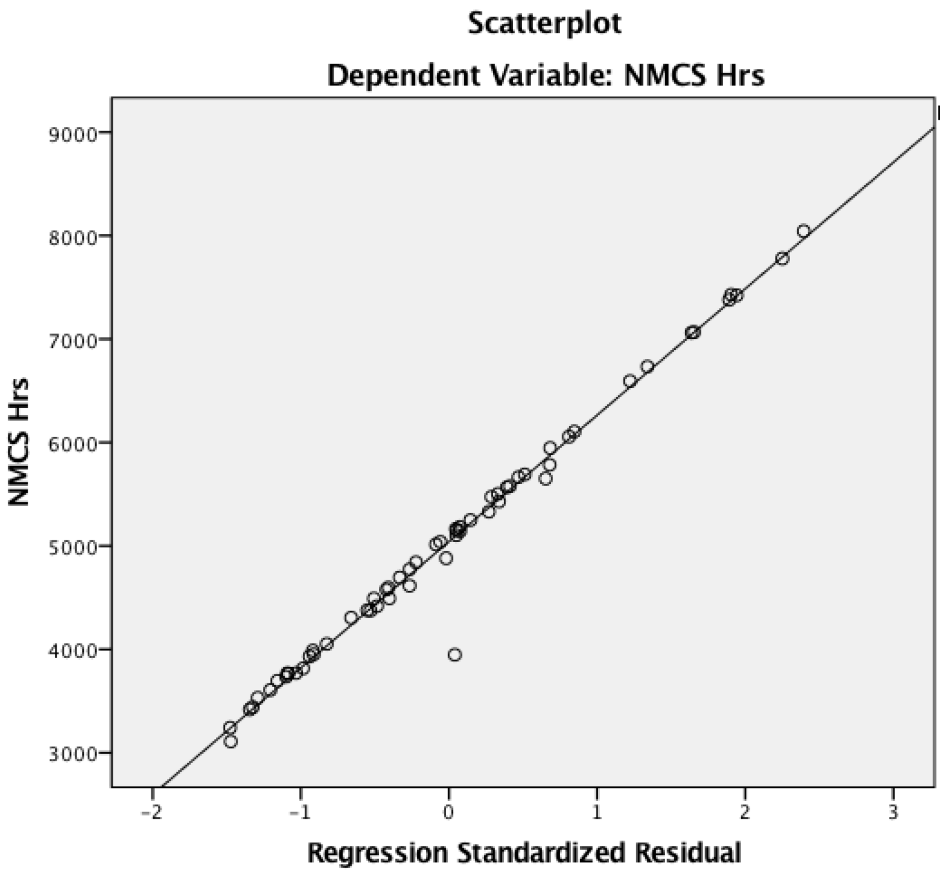
	SDR Yes/No	Kolmogorov-Smirnov[a] Statistic	df	Sig.	Shapiro-Wilk Statistic	df	Sig.
MICAPHours	Yes	.362	1623	.000	.293	1623	.000
	No	.366	74375	.000			
MICAPIncidents	Yes	.368	1623	.000	.499	1623	.000
	No	.395	74375	.000			

a. Lilliefors Significance Correction

Test of Homogeneity of Variance

		Levene Statistic	df1	df2	Sig.
MICAPHours	Based on Mean	12.340	1	75996	.000
	Based on Median	7.482	1	75996	.006
	Based on Median and with adjusted df	7.482	1	75686.431	.006
	Based on trimmed mean	8.997	1	75996	.003

Appendix H. Nellis AFB Tests of Normality and Heteroscedasticity

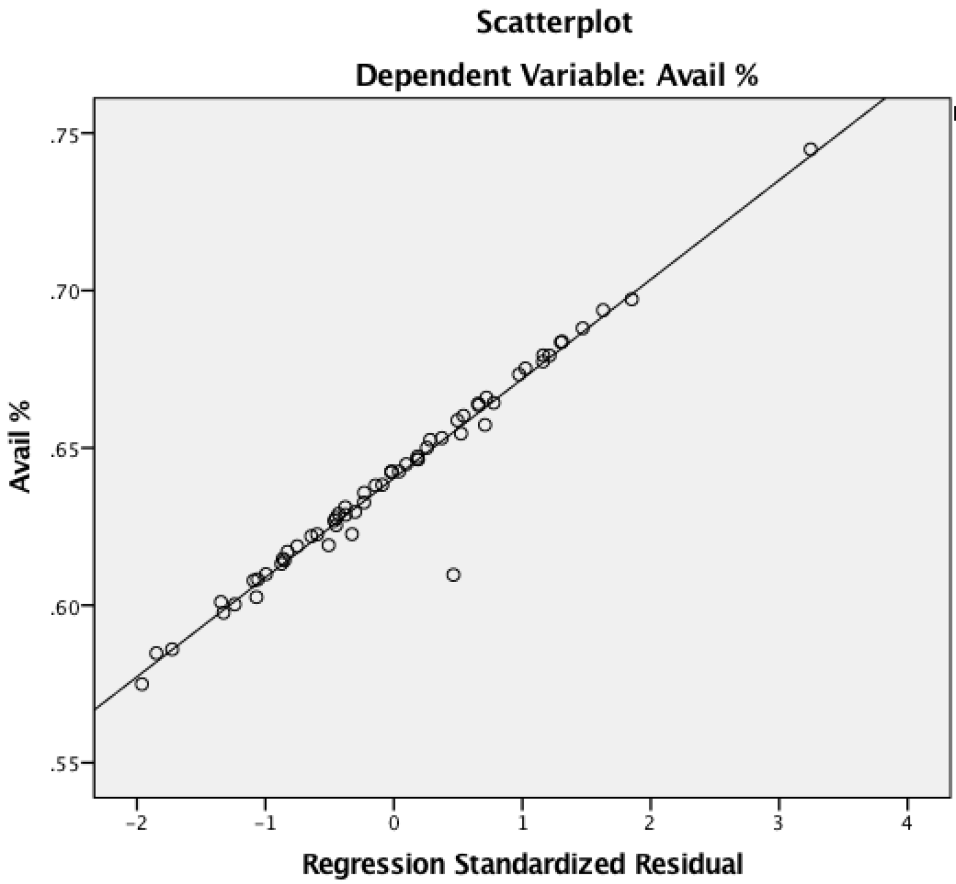


Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Unstandardized Residual	.087	60	.200*	.948	60	.013
Standardized Residual	.087	60	.200*	.948	60	.013

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

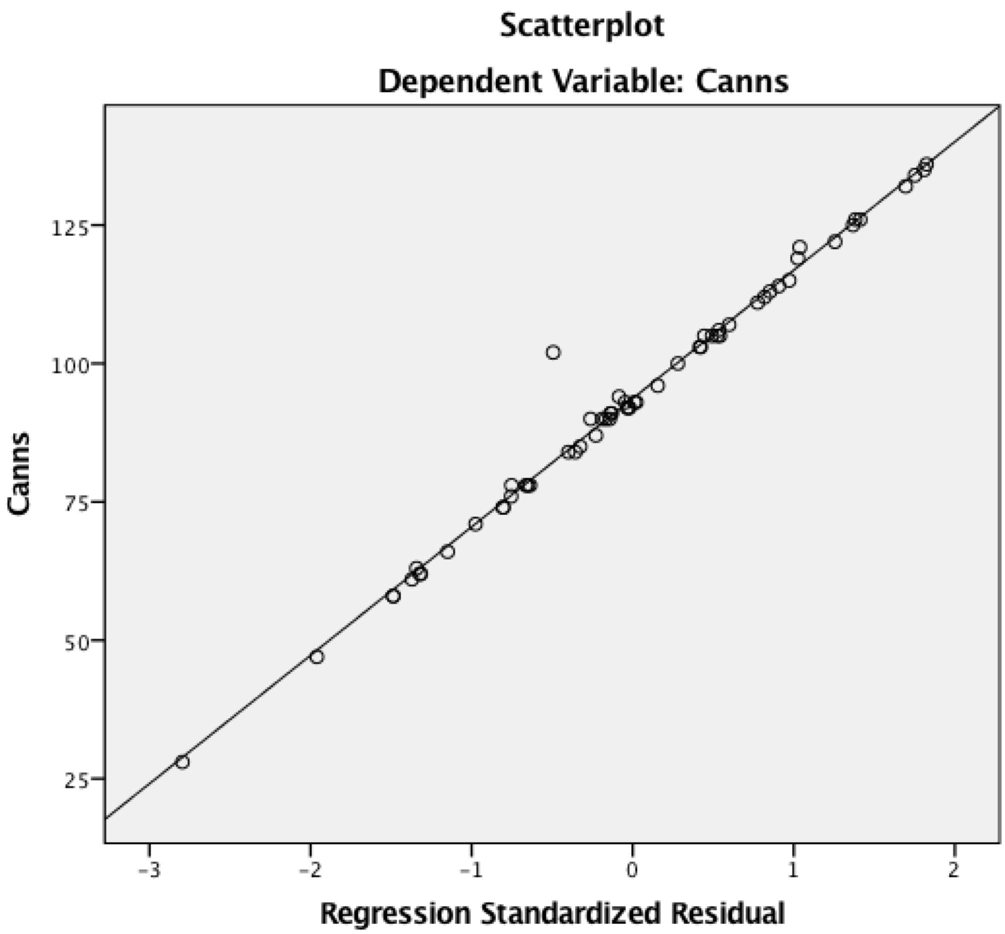


Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Unstandardized Residual	.059	60	.200*	.981	60	.473
Standardized Residual	.059	60	.200*	.981	60	.473

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction



Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Unstandardized Residual	.074	60	.200*	.984	60	.611
Standardized Residual	.074	60	.200*	.984	60	.611

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

Tests of Normality

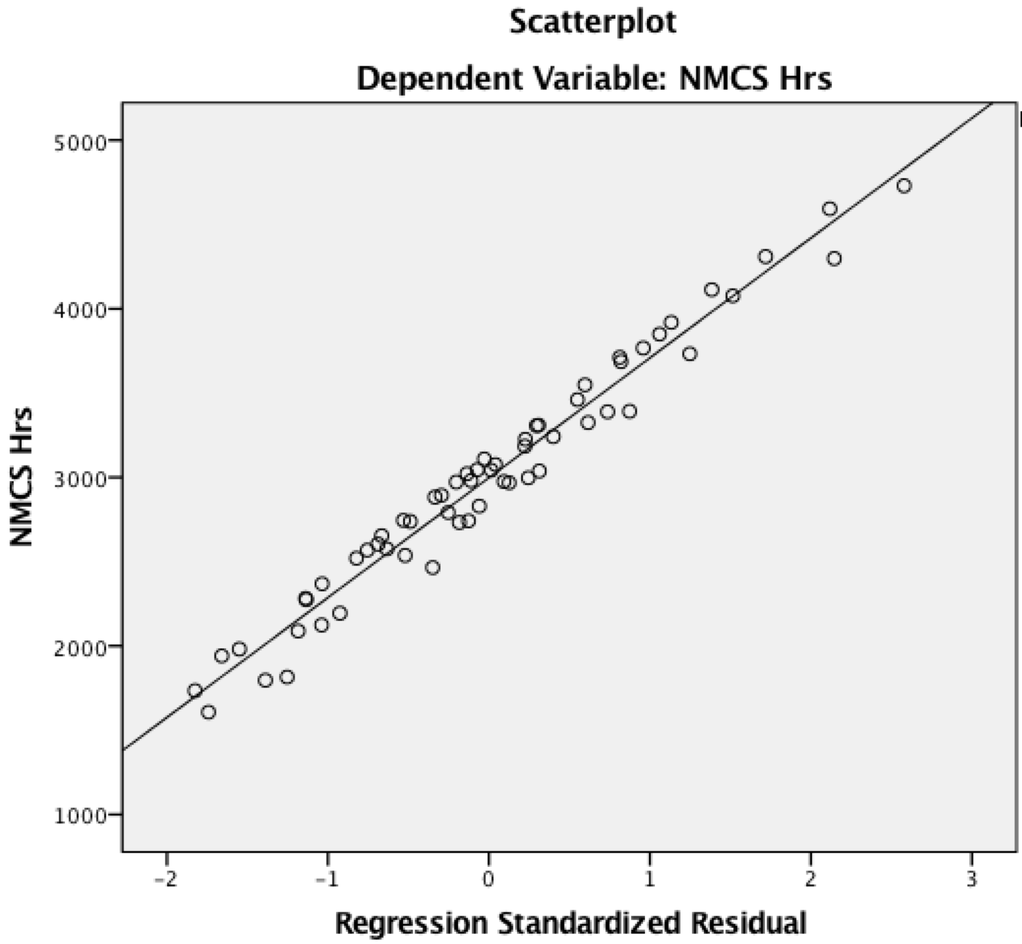
	SDR YesNo	Kolmogorov-Smirnov[a] Statistic	df	Sig.	Shapiro-Wilk Statistic	df	Sig.
MICAPHours	No	.392	36768	.000			
	Yes	.357	323	.000	.433	323	.000

a. Lilliefors Significance Correction

Test of Homogeneity of Variance

		Levene Statistic	df1	df2	Sig.
MICAPHours	Based on Mean	9.375	1	37089	.002
	Based on Median	4.337	1	37089	.037
	Based on Median and with adjusted df	4.337	1	37082.541	.037
	Based on trimmed mean	6.260	1	37089	.012

Appendix I. Kadena AB Tests of Normality and Heteroscedasticity

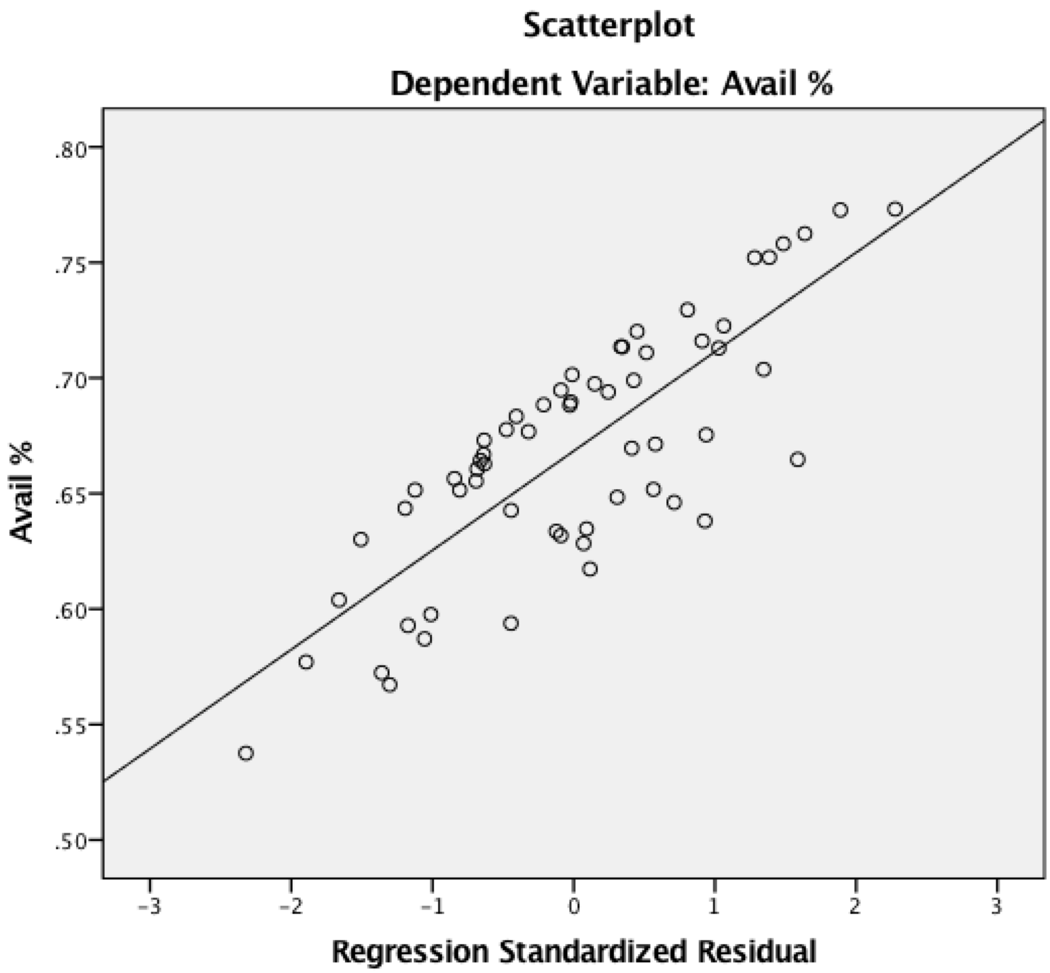


Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Unstandardized Residual	.076	60	.200*	.983	60	.557
Standardized Residual	.076	60	.200*	.983	60	.557

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

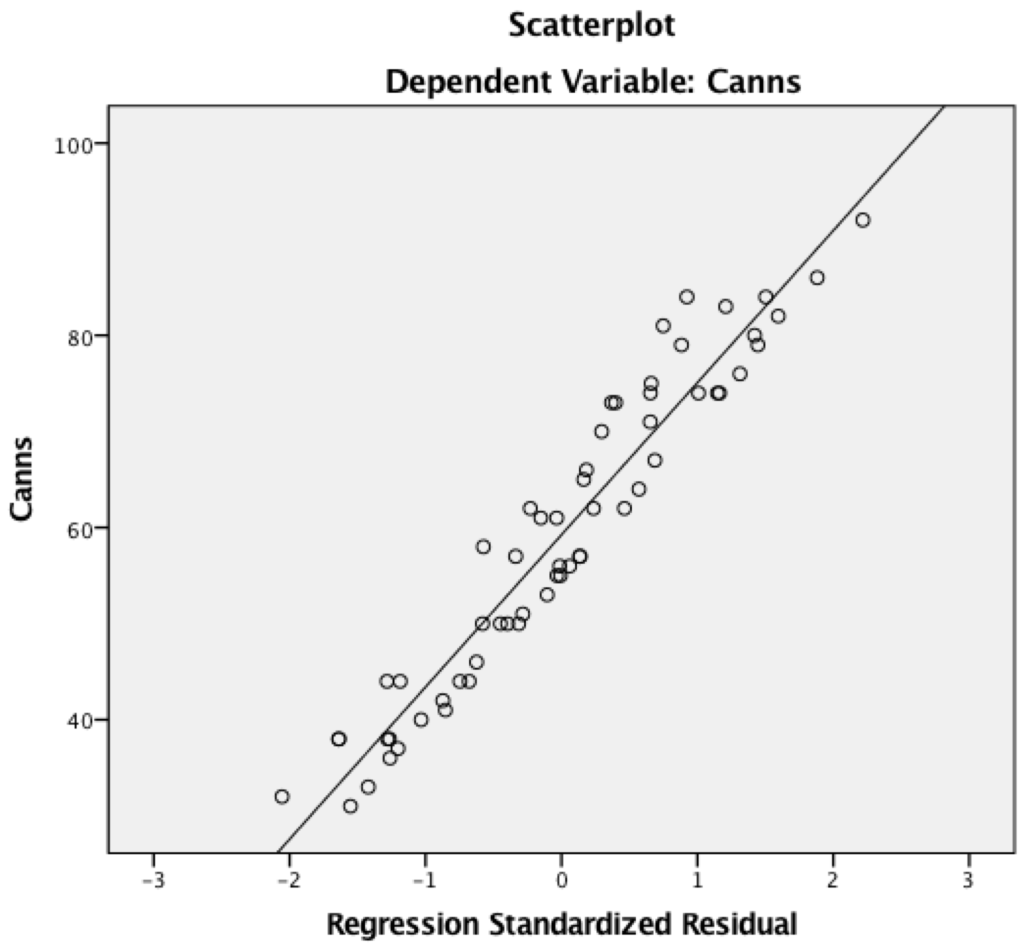


Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Standardized Residual	.055	60	.200*	.995	60	.998
Unstandardized Residual	.055	60	.200*	.995	60	.998

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction



Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Unstandardized Residual	.068	60	.200*	.987	60	.766
Standardized Residual	.068	60	.200*	.987	60	.766

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

Tests of Normality

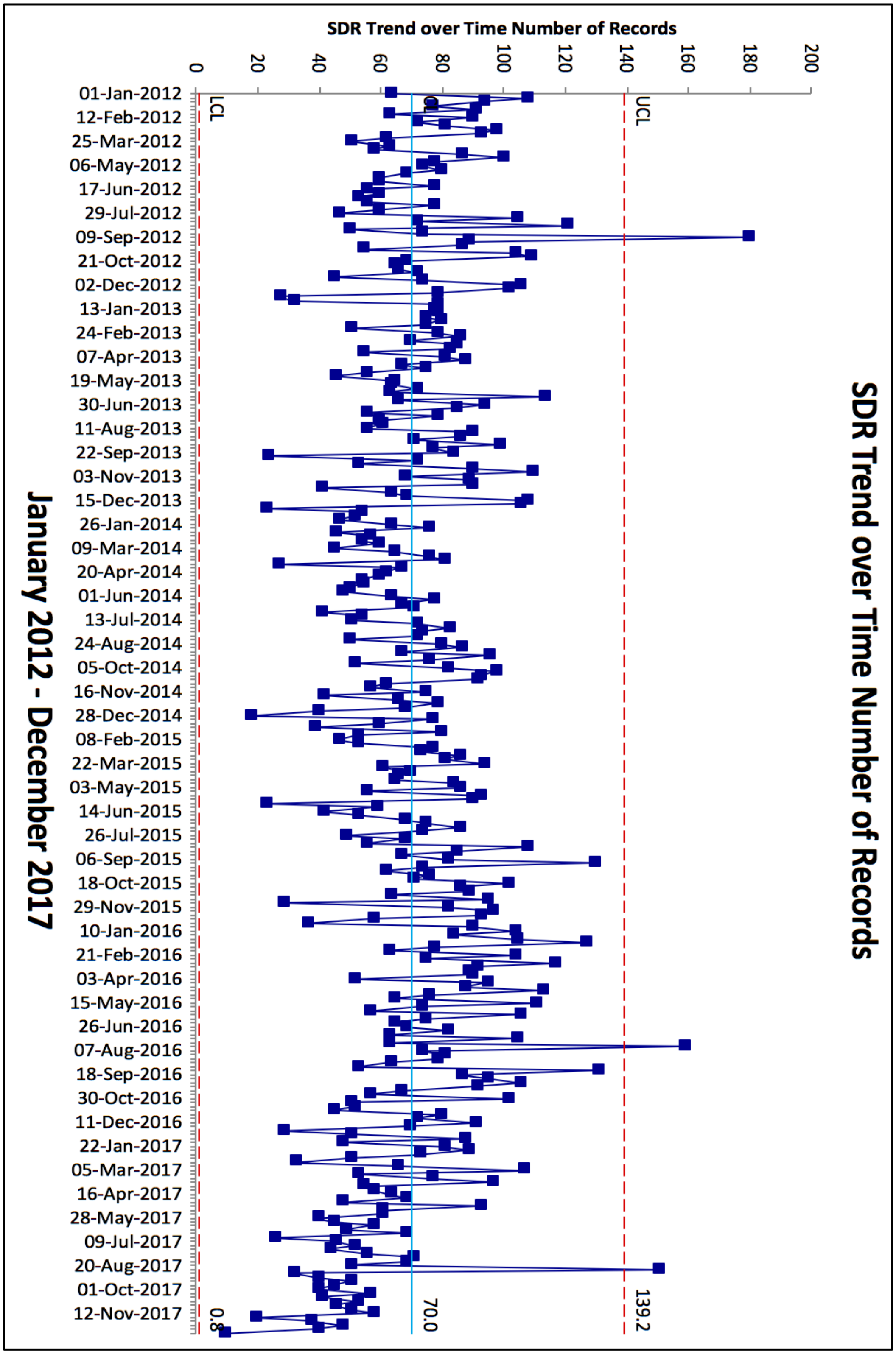
	SDR YesNo	Kolmogorov-Smirnov[a] Statistic	df	Sig.	Shapiro-Wilk Statistic	df	Sig.
MICAPHours		.348	22480	.000			
	Yes	.338	649	.000	.329	649	.000

a. Lilliefors Significance Correction

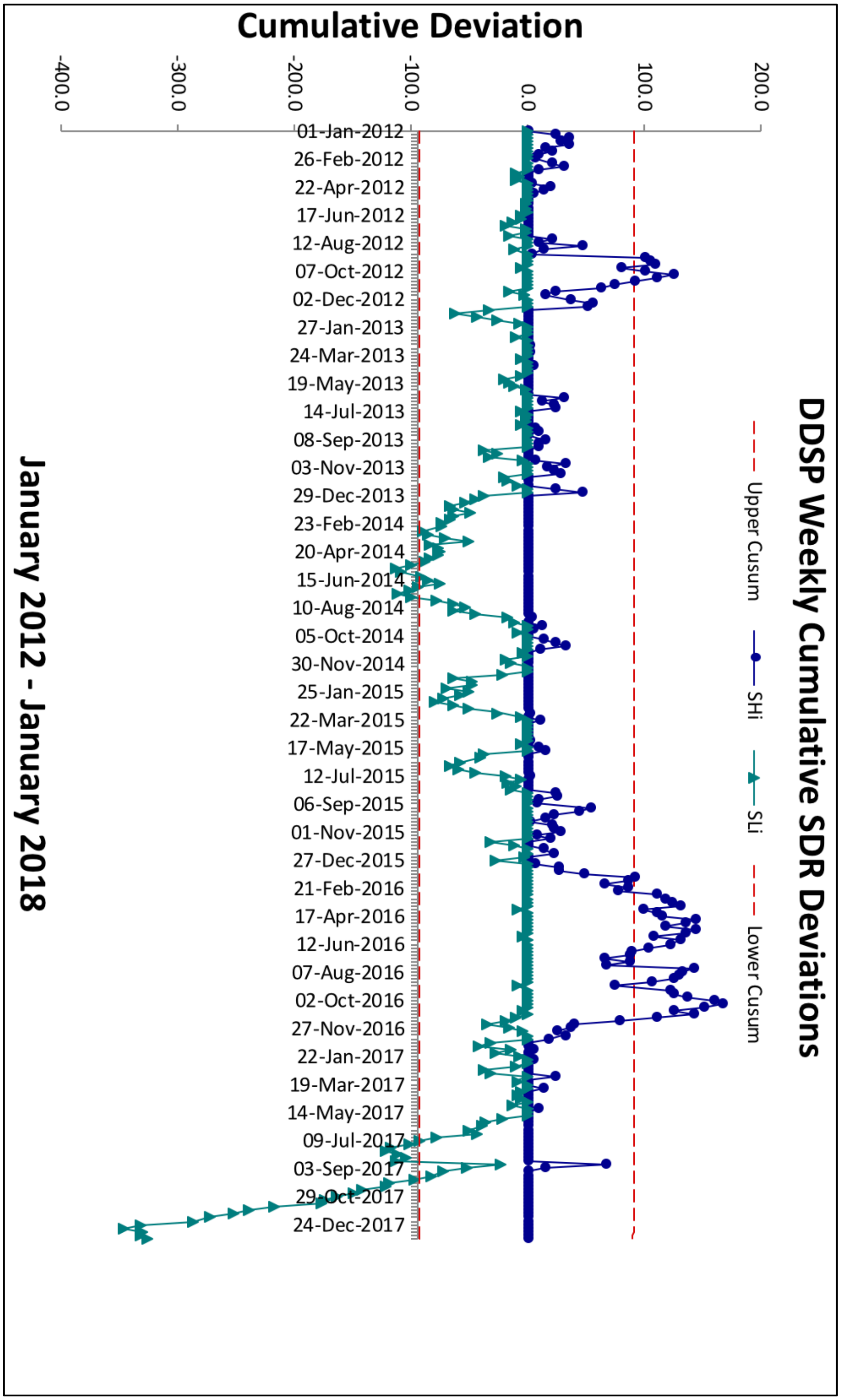
Test of Homogeneity of Variance

		Levene Statistic	df1	df2	Sig.
MICAPHours	Based on Mean	5.573	1	23127	.018
	Based on Median	2.561	1	23127	.110
	Based on Median and with adjusted df	2.561	1	23057.454	.110
	Based on trimmed mean	3.591	1	23127	.058

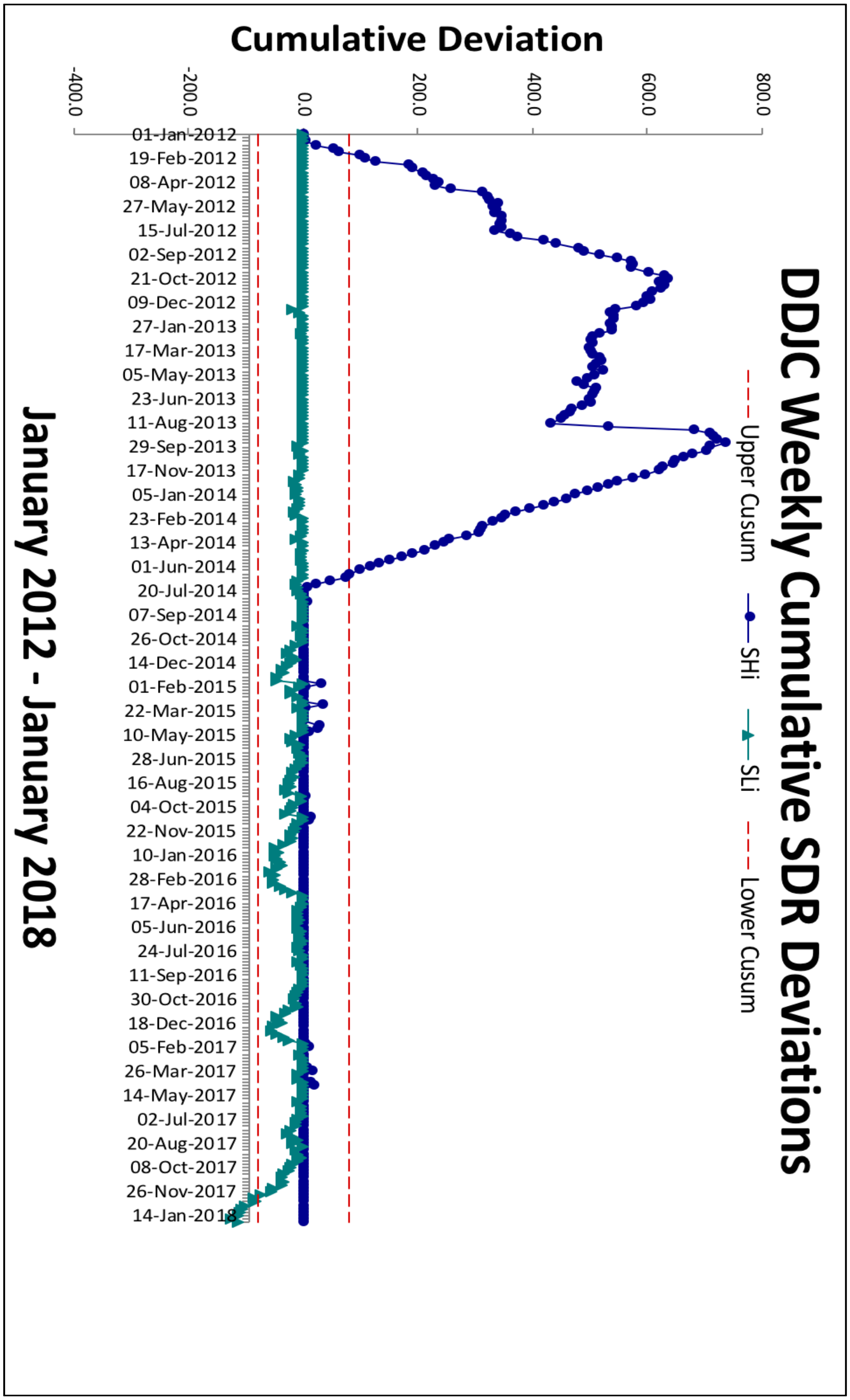
Appendix J. Shewhart Control Chart of Weekly DDSP SDRs from 2012-2017



Appendix K. CUSUM Control Chart of Weekly DDSP SDRs



Appendix L. CUSUM Control Chart of Weekly DDJC SDRs



References

Akyuz, G. A., & Erkan, T. E. (2010). Supply chain performance measurement: A literature review. *International Journal of Production Research, 48*(17), 5137-5155.

Anderson, E. W., Fornell, C., & Lehmann, D. R. (1994). Customer satisfaction, market share, and profitability: Findings from Sweden. *Journal of Marketing, 58*(3), 53-66.

Babbie, E. R. (2010). *The practice of social research* (10th ed.). Belmont, CA: Wadsworth Cengage.

Balzer, W. K., Doherty, M. E., & O'Connor, R., Jr. (1989). Effects of cognitive feedback on performance. *Psychological Bulletin*, *106*(3), 410–433.

Bienstock, C. C., Mentzer, J. T., & Bird, M. M. (1997). Measuring physical distribution service quality. *Journal of the Academy of Marketing Science, 25*(1), 31-44.

Bitici, U., Garengo, P., Dorfler, V., & Nudurupati, S. (2012). Performance measurement: Challenges for tomorrow. *International Journal of Management Review, 14*(3), 305-327.

Borg, W. R., & Gall, M. D. (1989). *Educational research: An introduction* (5th ed.). White Plains, NY: Longman.

Bourne, M. C., Neely, A. D., Mills, J. F., & Platts, K. W. (2003). Implementing performance measurement systems: A literature review. *International Journal of Business Performance Management, 5*(1), 1-24.

Bowersox, D. J., Closs, D. J., & Cooper, M. B. (2002). *Supply chain logistics management.* New York, NY: McGraw-Hill.

Box, G. E. P., & Jenkins, G. M. (1970). *Time series analysis: Forecasting and control.* San Francisco, CA: Holden-Day.

Bray, J. (1990). *Administrative and holding costs resulting from processing reports of discrepancy*. Alexandria, VA: Defense Logistics Agency.

Brennan, R. L. (2001). An essay on the history and future of reliability from the perspective of replications. *Journal of Educational Measurement, 38*(4), 295-317.

Carver, C. S., & Scheier, M. F. (1981). *Attention and self-regulation: A control theory approach to human behavior*. New York, NY: Springer-Verlag.

Chesbrough, H. W., & Garman, A. R. (2009). How open innovation can help you cope in lean times. *Harvard Business Review, 87*, 68-76.

Chiou, J. S., & Droge, C. (2006). Service quality, trust, specific asset investment, and expertise: Direct and indirect effects in a satisfaction-loyalty framework. *Journal of the Academy of Marketing Science, 34*(4), 613-627.

Christopher, M. (2005). *Logistics and supply chain management* (3rd ed.). Dorchester, Great Britain: Pearson.

Cooper, D. J., Ezzmel, M., & Qu, S. (2016). Popularizing a management accounting idea: The case of the balanced scorecard. *Social Science Resource Network.* Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2765258.

Copacino, W. C. (1997). *Supply chain management: The basics and beyond.* Boca Raton, FL: St. Lucie Press.

Coyle, J. J., Bardi, E. J., & Langley, C. J. (2002). *Management of business logistics: A supply chain perspective.* Cincinnati, OH: South-Western College Pub.

Coyle, J. J., Langley, C. J., Gibson, B. J., Novack, R. A., & Bardi, E. J. (2009). *Supply chain management: A logistics perspective* (8th ed.). Mason, OH: Cengage.

Croxton, K. L. (2014). The order fulfillment process. In Lambert, D. M. (Ed.), *Supply chain management: Processes, partnerships, performance* (4th ed.) (pp. 107-122). Ponte Vedra Beach, FL: Supply Chain Management Institute.

Davis-Sramek, B., Droge, C., Mentzer, J. T., & Myers, M. B. (2009). Creating commitment and loyalty behavior among retailers: What are the roles of service quality and satisfaction? *Journal of the Academy of Marketing Science, 37*, 440-454.

Davis-Sramek, B., Mentzer, J. T., & Stank, T. P. (2008). Creating consumer durable loyalty through order fulfillment service operations. *Journal of Operations Management, 26*, 781-797.

Defense Logistics Agency. (2017). DLA at a Glance. Retrieved from http://www.dla.mil/AtaGlance.aspx.

Defense Logistics Manual 4000.25. (2016). *Defense Logistics Management Standards* (Vol. 2) *Supply Standards and Procedures.* Washington, DC: Department of Defense.

Department of Defense Manual 4140.01 Volume 10. (2017). *DoD Supply Chain Materiel Management Procedures: Supply Chain Inventory Reporting and Metrics.* Washington, DC: Department of Defense.

Department of Defense. (2016). *Supply Chain Metrics Guide.* Washington, DC: Department of Defense.

Edgell, S. E., & Noon, S. M. (1984). Effect of violation of normality on the t test of the correlation coefficient. *Psychological Bulletin, 95*(3), 576-583.

Fornell, C. (1992). A national customer satisfaction barometer: The Swedish experience. *Journal of Marketing, 55*(1), 1-21.

Frese, M., & Zapf, D. (1994). Action as the core of work psychology: A German approach. In H. C. Triandis, M. D. Dunnette, & L. M. Hough (Eds.), Handbook of industrial and organizational psychology (2nd ed., Vol. 4, pp. 271-340). Palo Alto, CA: Consulting Psychologists Press.

Goh, S. C. (2012). Making performance measurement systems more effective in public sector organizations. *Measuring Business Excellence, 16*(1), 31-42.

Griffis, S. E., Rao, S., Goldsby, T. J., Voorhees, C. M., & Iyengar, D. (2012). Linking order fulfillment performance to referrals in online retailing: An empirical analysis. *Journal of Business Logistics, 33*(4), 279-294.

Gunasekaran, A., & Kobu, B. (2007). Performance measures and metrics in logistics and supply chain management: A review of recent literature (1995-2004) for research and applications. *International Journal of Production Research, 45*(12), 2819-2840.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81-112.

Hauser, J. R., & Katz, G. M. (1998). Metrics: You are what you measure! *The International Center for Research on the Management of Technology.* Cambridge, MA: Massachusetts Institute of Technology.

Johnson, C. D., (2000). *USAF aircraft maintenance management: Is there a better way?* Air Command and Staff College, Maxwell Air Force Base, Alabama.

Kaplan, R. S., & Norton, D. P. (1992). The balanced scorecard: Measures that drive performance. *Harvard Business Review, 70* (January-February), 71-79.

Kernan, M. C., & Lord, R. G. (1991). An application of control theory to understanding the relationship between performance and satisfaction. *Human Performance, 4,* 173-185.

Klapper, L. S., Hamblin, N., Hutchison, L., Novak, L., & Vivar, J. (1999). Supply chain management: A recommended performance measurement scorecard. McLean, VA: Logistics Management Institute.

Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback-intervention theory. *Psychological Bulletin, 119*(2), 254-284.

Knupfer, N. N., & McLellan, H. (1996). Descriptive research methodologies. In D. H. Jonassen (Ed.), *Handbook of Research for Educational Communications and Technology* (pp. 1196-1212). New York: Macmillan.

Kulhavy, R. W., White, M. T., Topp, B. W., Chan, A. L., & Adams, J. (1985). Feedback complexity and corrective efficiency. *Contemporary Educational Psychology*, *10*, 285–291.

Kumar, U., Galar, D., Parida, A., Stenstrom, C., & Berges, L. (2013). Maintenance performance metrics: A state of the art review. *Journal of Quality in Maintenance Engineering, 19*(3), 233-277.

Lin, F. R., & Shaw, M. J. (1998). Reengineering the order fulfillment process in supply chain networks. *International Journal of Flexible Manufacturing Systems, 10*(3), 197-229.

Lysakowski, R. S., & Walberg, H. J. (1982). Instructional effects of cues, participation, and corrective feedback: A quantitative synthesis. American Educational Research Journal, *19*(4), 559–578.

McCauley, J. L., Bassler, K. E., and Gunaratne, G. H. (2008). On the analysis of time series with nonstationary increments. In B. Rosser (Ed.), *Handbook of Complexity Research* (pp. 259-284). Cheltenham, UK: Edward Elgar.

McKinney, J. S. (1995). *Cost of processing reports of discrepancy: Administrative costs and holding costs* (Department of Defense Publication DLA-95-P50094). Fort Belvoir, VA.

Mentzer, J. T., Flint, D. J., & Hult, G. T. M. (2001). Logistics service quality as a segment-customized process. *Journal of Marketing*, *65*(2), 82-104.

Monica, F. S., Kennerley, M., Micheli, P., Martinez, V., Mason, S., Marr, B., Gray, D., & Neely, A. (2007). Towards a definition of a business performance measurement system. *International Journal of Operations and Production Management*, *27*(8), 784-801.

Muckstadt, J. A. (1973). A model for multi-item, multi-echelon, multi-indenture inventory system. *Management Science*, *20*, 472-481.

National Institute of Standards and Technology. (2008). Shewhart control chart. Retrieved from http://www.itl.nist.gov/div898/handbook/mpc/section2/mpc221

Neely, A. D. (2002). *Business performance measurement: Theory and practice.* Cambridge, UK: Cambridge University Press.

Neely, A. D., Mills, J., Platts, K., Richards, H., Gregory, M., & Bourne, M. (2000). Performance measurement system design: Developing and testing a process based approach. *International Journal of Operations and Production Management*, *20*(10), 1119-1145.

Petcoff, R. P. (2010, May 5). Air Force program recognized for excellence in government. *U.S. Air Force.* Retrieved from http://www.af.mil/News/Article-Display/Article/116767/air-force-program-recognized-for-excellence-in-government/.

Rao, S., Griffis, S., & Goldsby, T. (2011). Failure to deliver? Linking online order fulfilment glitches with future purchase behavior. *Journal of Operations Management, 29*(8), 692-703.

Reichheld, F. F. & Sasser, W. E. (1990). Zero defections: Quality comes to services. *Harvard Business Review, 68* (September-October), 105-111.

Stahl, A. L. (2014). *A survey and analysis of aircraft maintenance metrics: A Balanced Scorecard approach.* MS Thesis. Graduate School of Engineering and Management, Air Force Institute of Technology (AU), Wright-Patterson Air Force Base, Ohio.

Stank, T.P., Goldsby, T.J.,Vickery, S.K., & Savitskie, K. (2003). Logistics service performance: Estimating its influence on market share. *Journal of Business Logistics 24*(1), 27–56.

Stapenhurst, T. (2005). *Mastering Statistical Process Control.* Abingdon, UK: Routledge.

Tenenbaum, G., & Goldring, E. (1989). A meta-analysis of the effect of enhanced instruction: Cues, participation, reinforcement and feedback and correctives on motor skill learning. *Journal of Research and Development in Education, 22*(3), 53–64.

Thirumalai, S., & Sinha, K. K. (2005). Customer satisfaction with order fulfillment in retail supply chains: Implications of product type in electronic B2C transactions. *Journal of Operations Management, 23*(3), 291-303.

Vector Research. (1997). *Department of Defense Guide for Managing Information Technology (IT) as an Investment and Measuring Performance*. Arlington, TX: Vector Research.

Voxware (2012). Impact of late deliveries can be disastrous, study shows. Retrieved from http://www.supplychainbrain.com/content/industry-verticals/retail/single-article-page/article/impact-of-late-or-inaccurate-deliveries-can-be-disastrous-study-shows/.

Walberg, H. J. (1982). What makes schooling effective? *Contemporary Education Review*, *1*(1), 1–34.

White, H., & Sabarwal, S. (2014). Quasi-experimental design and methods. *Methodological Briefs: Impact Evaluation.* Florence, IT: *Unicef Office of Research.*

Wouters, M., & Wilderom, C. (2008). Developing performance measurement systems as enabling formalization: A longitudinal field study of a logistics department. *Accounting, Organizations, and Society, 33*(4), 488-516.

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)*	**2. REPORT TYPE**	**3. DATES COVERED** *(From – To)*
22-03-2018	Master's Thesis	Oct 2016 – Mar 2018

TITLE AND SUBTITLE	
MANAGING SUPPLY DISCREPANCIES: THE EFFECT OF PERFORMANCE MEASUREMENT AND FEEDBACK ON ORDER FULFILLMENT QUALITY	**5a. CONTRACT NUMBER**
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	**5d. PROJECT NUMBER**
Weber, Michael J., Captain, USAF	**5e. TASK NUMBER**
	5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAMES(S) AND ADDRESS(S)	**8. PERFORMING ORGANIZATION REPORT NUMBER**
Air Force Institute of Technology Graduate School of Engineering and Management (AFIT/EN) 2950 Hobson Way, Building 640 WPAFB OH 45433-8865	AFIT-ENS-MS-18-M-168

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

DLA Distribution Susquehanna	635th Supply Chain Operations Wing
2001 Mission Drive	South Drive
New Cumberland, PA 17070-5000	Scott AFB, IL 62225
717-770-7421	618-256-4241
ATTN: Col Brad Eungard	ATTN: Lt. Col Daniel Johnstone

10. SPONSOR/MONITOR'S ACRONYM(S)

AFRL/RHIQ (example)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
DISTRUBTION STATEMENT A. APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

13. SUPPLEMENTARY NOTES

14. ABSTRACT
Extensive research on the impact of shipping and packaging errors in the private sector finds numerous negative outcomes, including reduced customer satisfaction, reduced customer loyalty, and lower profitability. However, little research has been done examining the impact of order fulfillment errors on military operations. The purpose of this research is to quantify the impact of supply discrepancy reports (SDRs) on military aircraft readiness metrics, including cannibalizations, not mission capable supply (NMCS) hours, aircraft availability and MICAP hours. Results show SDRs significantly impact aircraft readiness metrics in seven of the fifteen analyses conducted. Additionally, a quasi-experimental study is implemented at DLA Distribution Susquehanna, Pennsylvania (DDSP) aimed at reducing supply discrepancies using performance measurement and feedback over a seventeen-week period. Cumulative sum (CUSUM) control charts showed a decline in the number of reported SDRs for fifteen consecutive weeks, amounting to the lowest average in over six years. The results of this research suggest that aircraft readiness metrics across the Air Force could show measurable improvement if similar SDR reduction strategies are implemented throughout more DoD suppliers.

15. SUBJECT TERMS
Aircraft, Logistics, Performance Measurement, Readiness, Supply Discrepancies

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
					Dr. Daniel Steeneck, AFIT/ENS
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			**19b. TELEPHONE NUMBER** *(Include area code)*
U	U	U	UU	118	(937) 785-3636, ext 4702 daniel.steeneck@afit.edu

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