Air Force Institute of Technology

AFIT Scholar

Theses and Dissertations | Student Graduate Works

6-2008

Forecasting Flying Hour Costs of the B-1, B-2, and B-52 Bomber Aircraft

Stefanie L. Van Dyk

FORECASTING FLYING HOUR COSTS OF THE B-1, B-2 AND B-52 BOMBER AIRCRAFT

THESIS

Stefanie L. Van Dyk, First Lieutenant, USAF

AFIT/GCA/ENV/08-M02

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY

AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

AFIT/GCA/ENV/08-M02

FORECASTING FLYING HOUR COSTS OF THE B-1, B-2, AND B-52 BOMBER AIRCRAFT

THESIS

Presented to the Faculty

Department of Systems Engineering and Management

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 Cost Analysis

Stefanie L. Van Dyk, BS

First Lieutenant, USAF

March 2008

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED

AFIT/GCA/ENV/08-M02

FORECASTING FLYING HOUR COSTS OF THE B-1, B-2, AND B-52 BOMBER AIRCRAFT

Stefanie L. Van Dyk, BS
First Lieutenant, USAF

Approved:

_____________//signed//______________ 30 May 2008
Eric J. Unger (Chairman) Date

_____________//signed//______________ 30 May 2008
Edward D. White (Member) Date

AFIT/GCA/ENV/08-M02

Abstract

This thesis both evaluates, and presents improvements to, the current method of forecasting flying costs of Air Force aircraft. It uses depot level repairable (DLR) and consumable (CONS) data for the Air Force's bomber platforms: B-1B, B-2, and B-52H. The current forecasting method assumes a proportional relationship between costs and flying hours such that 1) when no hours are flown costs are zero, and 2) a 1% increase in flying hours will increase costs by 1%. The findings of this research indicate that applying log-linear ordinary least squares regression techniques may be an improved fit of flying cost data over the current proportional model; the actual data indicate a non-zero intercept and a less than proportional relationship between costs and flying hours. This research also found that models including factors other than flying hours as independent variables, such as sorties, lagged costs, and fiscal trends, may be more useful than models based solely on flying hours. Finally, this research found that estimating quarterly costs at the base-level may yield more accurate estimates than estimating at the monthly level, or mission design series level.

Acknowledgements

First, I would like to thank my family for all the support they have given me over the years. Their constant encouragement fueled my enthusiasm to not only develop the best product I could, but to learn as much as possible along the way.

I would like to express my sincere appreciation to my committee, Lt Col Eric Unger and Dr. Tony White. I feel very privileged to have had the opportunity to work with such intelligent and patient academics. Also, thank you to Lt Col Jeffrey Smith and Dr. Michael Hicks for their motivation in the classroom and as mentors. I would also like to thank Mark Gossett, Larry Klapper, and William "Crash" Lively for all of their timely help in providing and explaining the AFTOC and REMIS data necessary to complete this research. I am also grateful to TSgt Pichai Polprasert from the AFCCC for providing me climatology data. Finally, thanks to Billy Kirby for his time explaining CAM.

Stefanie L. Van Dyk

Table of Contents

List of Figures

Page

List of Tables

FORECASTING FLYING HOUR COSTS OF THE B-1, B-2, AND B-52 BOMBER AIRCRAFT

I. Introduction

Background

"The Air Force at one time stood head and shoulders above all the Services in their capabilities, organization, and processes for developing cost estimates for major weapon systems" (Kammerer, 2003:1). Developing accurate cost estimates is important for every program within the Department of Defense (DOD); poor estimates can cause funds to be misallocated, which can lead to under-funded or over-funded programs.

Programs which suffer from a lack of funding must request additional funds from Congress who, in turn, pulls funds from other programs. Further, over-funding a program causes money to be tied up obligated within a program that does not need it, leaving other areas under-funded. Generally, funds are taken from research and development programs, leading to a reduction in technology progression and weapon development (Kammerer, 2002:1). Operations and Maintenance (O&M) costs, which represent about 40% of the overall DOD budget, have seen significant growth in recent years resulting in insufficient budgeting and leading to requests for additional funding.

O&M costs are costs required to operate and maintain the nation's military forces. The O&M budget includes a variety of costs, ranging from flying costs, to health care and environmental programs (Kiley, 2001:1). From FY2000 to 2005 these costs grew from $133.4 billion (FY07$) to $209.5 billion (FY07$) (GAO, 2007:9-11). This significant growth in cost was funded primarily through supplemental funding, as depicted in Figure 1. The GAO (2007) states that the significant increase in O&M costs are due not only to increased operations in support of the global war on terror, but also to aging military equipment (GAO, 2007:13-14).

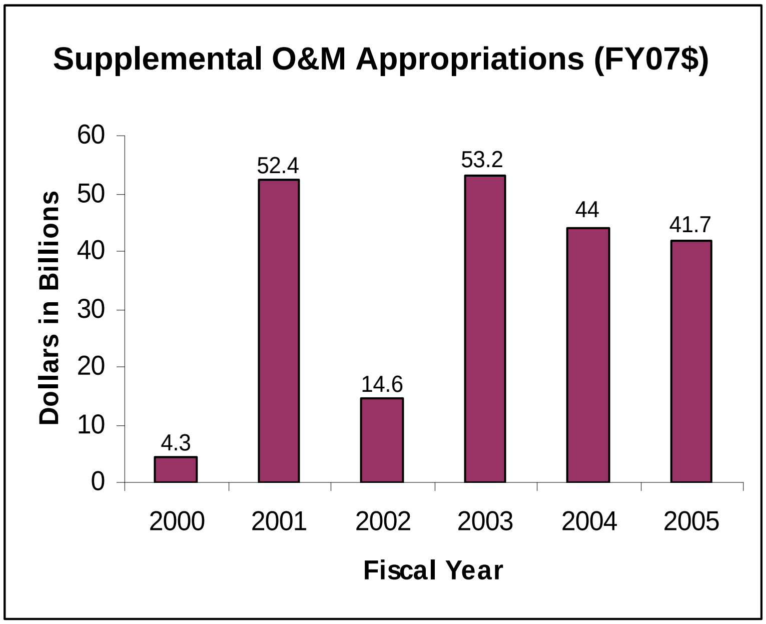

Figure 1: Supplemental O&M Appropriations (GAO, 2007)

The Air Force's flying cost program, which represents about 20% of the O&M budget (Kiley, 2001), has seen significant change over the last several decades. As shown in Figure 2, while O&M costs have grown nearly 600% (CY99$) over the last several decades, total flying hours have dropped nearly 75% (Defense, 2006). Likewise, flying programs in particular have seen nearly 10% cost growth in recent years, specifically on reparable and consumable parts (Kammerer, 2002).

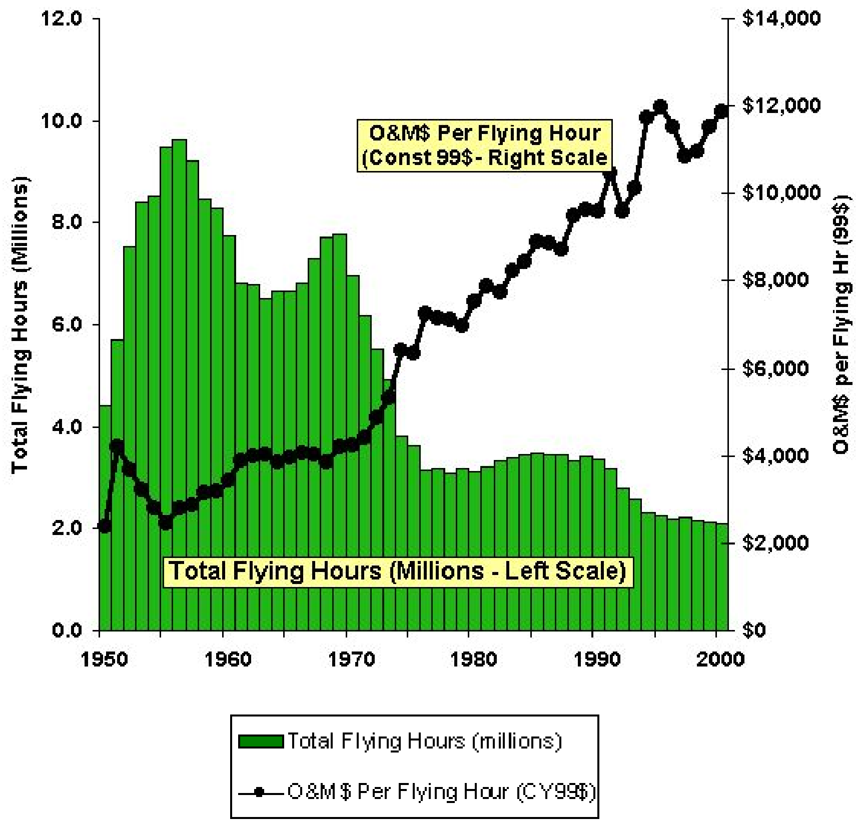

Figure 2: The Rising Cost of O&M per Flying Hour (Defense, 2006)

The Congressional Budget Office (Congressional, 2007) suggests that a reason behind these trends is that new weapon systems are becoming increasingly complex, leading to larger O&M costs than the systems they replace. Additionally, the existing weapon systems are reaching unprecedented ages, leading to increased parts consumption (Kammerer, 2002). Many have suggested that costs will continue to grow as aircraft continue to age (Kiley, 2001; Hart, 2003; Hebert, 2003; Pyles, 2003). In fact, Kiley (2001) suggested that continuing to fly these aging aircraft may cause O&M costs to rise much faster than the rates that have been realized to date. Older equipment will “wear out faster” than new equipment due to problems associated with corrosion and fatigue (Hebert, 2003:4).

Flying Hour Program Costs

Until FY2008, developing flying program cost estimates was the responsibility of each major command (MAJCOM). However, in an effort to reduce cost overages and streamline financial processes, General T. Michael Moseley began implementation of a plan which will centralize all budgeting for the flying hour program in a single program office (Scaggs, 2006). This program office, called the Centralized Asset Management (CAM) program office, was introduced in FY2006. The Air Force is currently in the process of consolidating budgeting and spending processes within the CAM office. Centralization of budgeting for the flying hour programs is expected to be complete by the end of FY2008 (Rumple, 2007). While Air Force senior leaders have acknowledged problems with the current flying cost forecasting process and attempted to mitigate these problems with the CAM program office, the new office has adopted the old flying cost forecasting methodology.

The current process of forecasting flying costs involves calculating cost per flying hour (CPFH) factors based on historical data. Each MAJCOM estimates the CPFH factors for each Mission Design Series (MDS), or aircraft type, based on wing-level cost and flying hour inputs. Costs from a previous time period are divided by hours flown to determine the CPFH factor. Forecasters then modify these factors for economic and other approved adjustments. To forecast flying costs, forecasters multiply the adjusted CPFH factor by the projected number of flying hours. Figure 3 diagrams this process.

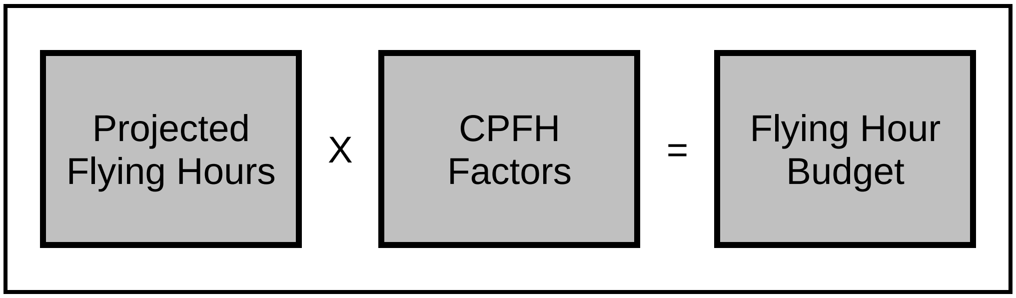

Figure 3: The CPFH Budgeting Process

Forecasters calculate CPFH factors for three different types of costs: depot level repairables (DLRs), consumables (CONS), and fuel. DLRs are defined as those which are repaired rather than disposed of, such as avionics systems and engines. In contrast, CONS are parts which are disposed of after their useful life, such as batteries and screws. The third type of flying cost, fuel, is defined simply as the fuel used during flight (GAO, 1999).

The CPFH process assumes the relationship between flying hours and operating costs is proportional. We define a proportional relationship as one with two characteristics: 1) the relationship between costs and flying hours is linear, such that a one percent increase in flying hours leads to a one percent increase in flying costs, and 2) the intercept is equal to zero, indicating that if no flying hours were accumulated for a given time period, the costs would be zero. In other words, the current proportional method assumes that operating costs are entirely variable. However, there are costs associated with flying programs that are fixed, such as routine maintenance checks. Figure 4 shows the disagreement between forecasting costs using the proportional CPFH method and actual costs, which indicate fixed costs associated with the flying program.

The CPFH factor methodology is appropriate for fuel costs, as an increase in the amount of hours flown generally leads to a proportional increase in the amount of fuel burned. Likewise, if no hours are flown, no fuel is used. However, as Kammerer mentioned, forecasts for DLR and CONS parts have not been as accurate (Kammerer, 2002). As Figure 4 shows, the proportional method, which considers all flying costs to be variable, will under-predict these costs when a low number of hours are flown in a given time period.

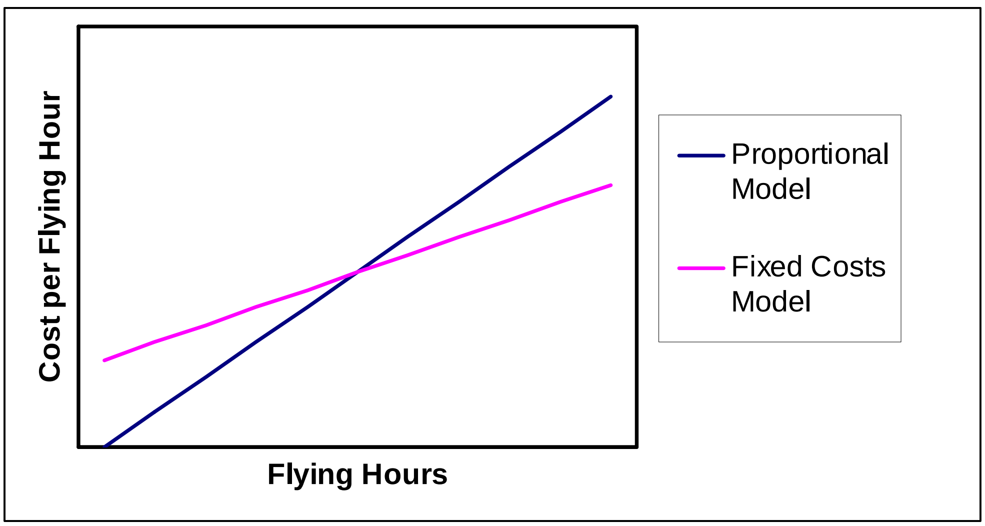

Figure 4: The Relationship Between Costs and Flying Hours

Likewise, the proportional method will over-predict costs when flying hours are high in a given time period.

As the Air Force centralizes its budgeting processes and continues to develop increasingly complex aircraft, which are believed to be driving up flying costs (Congressional, 2007), this potential flaw in the forecasting process must be tested and changes need to be made, if necessary. Estimators must seek more detailed knowledge of each individual aircraft, as well as all potential cost predictors, in order to develop accurate cost estimating models. This research effort will explore the relationships between cost and flying hours for three Air Force aircraft, the B-1B, B-2, ad B-52H. Specifically, the focus of this research is to forecast DLR and CONS costs for the bomber aircraft.

Purpose

The purpose of this research is to explore the usefulness of the current CPFH methodology in estimating DLR and CONS flying costs of the bomber aircraft. Additionally, we will use simple ordinary least squares (OLS) regression of costs on flying hours to see if the proportional model adequately captures the relationship between costs and flying hours given real data of bomber aircraft. Finally, we will identify and quantify other potential cost predictor variables. The variables we evaluate are based on those highlighted in previous research and are discussed in the next chapter.

The following two research questions will guide this research:

1. Does the current CPFH methodology, which assumes a proportional relationship, capture the true relationship between flying hours and costs?
2. Are factors other than flying hours useful in estimating flying costs?

Summary

The rest of this paper is structured as follows: Chapter II presents a literature review which summarizes multiple studies that have explored both operating costs and the CPFH approach. Multiple variables have been evaluated and several different estimating techniques have been employed. Chapter III will discuss the data and methodology used in this study and how it will be used to evaluate both the current CPFH methodology and the predictability of variables other than flying hours. Chapter IV presents the results from the analyses and Chapter V will summarize the findings and answer each of the research questions.

II. Literature Review

Overview

This chapter provides a discussion of completed research concerning operations and maintenance (O&M) costs. We will summarize several studies and explain how the findings of those studies shaped the context of our research. In this chapter, we also discuss the decision to evaluate the effects of cannibalization rates – a variable which has not been evaluated in previous research – on consumables and reparable costs.

Previous Research

Several researchers have conducted studies attempting to better understand flying hour costs. These researchers have all aimed to improve the current budgeting process through a better understanding of the behavior of costs. Each presents different approaches to estimating the costs of flying through analysis of different variables and evaluation of the operating costs for different groups of aircraft. As a whole, this body of research has examined over 20 different variables. Eight of the flying hour cost studies, as well as the variables examined, are chronologically listed in Table 1, beginning with Hildebrandt (1990) and ending with the current study of bomber aircraft. The dependent variables range from overall Operations and Support (O&S) costs, to CPFH factors for individual aircraft. The independent variables represent a variety of different factors, from aircraft characteristics, to operations tempo, to environmental characteristics. This chart shows that aircraft age has been evaluated in several studies. Additionally, one can see that location characteristics, such as temperature and dew point, and operational factors, such as utilization rate and mission capable rate, are new to the list of potential cost predictors.

Table 1: Explanatory Variables Used in CPFH Research

	Hildebrandt (1990)	Wallace (2000)	Kiley (2001)	Pyles (2003)	Laubacher (2004)	Hawkes (2005)	Armstrong (2006)	Bryant (2007)	Van Dyk (2008)
Dependent Variables									
Operations/Support Costs	X		X						
Number of Parts Replacements		X							
Maintenance Work Hours				X					
Net Flying Costs					X				X
CPFH Factor						X	X	X	
Independent Variables									
Flying Hours	X	X	X				X	X	X
Lagged Costs					X	X			X
Aircraft Age	X	X	X	X		X	X		X
Average Total Operating Hours								X	X
Flyaway Cost	X		X	X					
IOC Year	X								
Aircraft Type (categorical)	X			X					
MAJCOM				X		X			
Percent Engine Type						X			
Percent Block						X			
Sorties		X							X
Average Sortie Duration						X	X	X	X
Utilization Rate						X		X	X
Mission Capable Rate								X	X
Cannibalization Rate									X
Deployments						X	X	X	
Ground Days		X							
Total Aircraft Inventory	X								
Program Change							X	X	
Base Location						X			
Petroleum Proxy							X	X	X
Temperature							X	X	X
Dew Point								X	
Month/Seasonality							X	X	X

The following pages will both describe the research summarized in the table above and explain how the findings will be applied to the present models for Air Force bombers.

RAND Finds Flying Hours and Age Affect Flying Hour Costs.

Gregory Hildebrandt and Man-bing Sze of RAND conducted a study concerning Air Force O&S costs nearly 18 years ago (Hildebrandt, 1990). This research, entitled "An Estimation of US Air Force Aircraft Operating and Support Cost Relations," seeks to link available O&S cost data to aircraft in order to provide a method of approximating O&S costs for acquisition programs (Hildebrandt, 1990:v). Hildebrandt and Sze's found their data in the Visibility and Management of Operating and Support Costs (VAMOSC) database. Their data consisted of 400 observations ranging from 1981 to 1986. The VAMOSC data are supplemented with information on aircraft flyaway cost, average mission design (MD) age, and the MDS year (defined as the year the MDS entered service) (Hildebrandt, 1990:16).

The authors use log-linear regression models to examine the relationship between O&S costs and several explanatory variables. Transforming linear relationships into a logarithmic relationship has a number of advantages, such as stabilizing the variance of the data and simplifying the interpretation of the regression coefficients. Logarithmic transformations will be discussed in greater detail in Chapter III. The authors found flying hours and flyaway costs to be statistically significant with relatively large coefficients. Specifically, a one percent increase in flying hours increases O&S costs by more than 0.6%, while a one percent increase in flyaway cost increases costs by more than 0.4%, all else being equal. The flyaway cost variable performs very well as a proxy for aircraft type and MDS year. While this research clearly supports including the flyaway cost as a variable in regression models evaluating multiple airframes, it is not applicable to our research of bomber aircraft since only one platform will be evaluated in each model, and each platform only consists of one variant (i.e. the B-1B and B-52H). The total number of aircraft within each MDS fleet, or total aircraft inventory (TAI), is also statistically significant and negatively correlated in each model. However, a one percent increase in the number of aircraft only results in about a 0.05% decrease in O&S costs. Additionally, TAI is highly correlated with flying hours. We will not include TAI as a variable in this study of bomber aircraft; our research used the number of aircraft in the derivation of other variables. These variables will be discussed in detail in the following chapter.

Hildebrandt and Sze (1990) discuss the "Bathtub Effect," which is the belief that costs initially experience a decline due to learning curves and the elimination of initial reliability and maintainability problems. Costs then experience a steady-state, followed by an increase in costs due to increased repair and maintenance requirements associated with age (Hildebrandt, 1990:20-21). Figure 5 depicts this trend in operating costs.

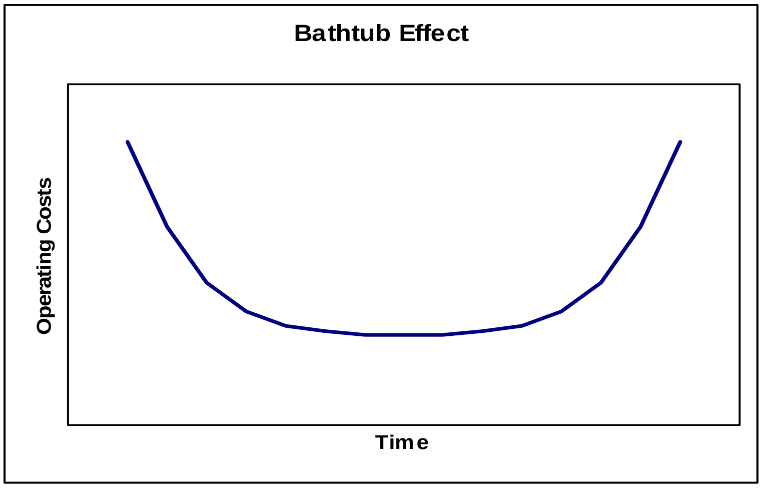

Figure 5: The Bathtub Effect

The B-1B and B-52H are mature aircraft at the point of our research and, therefore may be experiencing age related increases in costs, as depicted in Figure 5. This study of bomber aircraft will evaluate aircraft age to determine if age is significant in forecasting CONS and DLR costs of bombers. The following study supports Hildebrandt and Sze's findings that age drives cost. It also evaluates factors related to sorties as predictors of cost.

Wallace Evaluates Usage as a Predictor of Costs.

In the study "A Physics-Based Alternative to Cost-Per-Flying-Hour Models of Aircraft Consumption Costs," Wallace, Hauser, and Lee (2000) sought a model which would more accurately forecast flying hour costs during times of varying usage than the proportional CPFH model. The current CPFH model, they argued, is adequate during times when aircraft usage is relatively constant but insufficient during a usage "surge," which tends to correspond with military conflicts. These researchers forecasted costs by using explanatory variables that measure these surges, such as the number of take-offs and landings, the total flying hours, and the total number of hours an aircraft spends on the ground (Wallace, 2000:2-1). The authors chose to use data for the C-5B, the C-17, the KC-10, and the F-15C because they represent different types of aircraft. Additionally, these aircraft flew during major conflicts, which allowed the authors to compare the effectiveness of their model with the proportional model for both steady state usage and a usage surge.

To test the effects of their independent variables on the number of removals of parts (a proxy for costs) the authors split their data into four calibration sets. The model that performed the best used about 1/3 of the data, representing the months prior to a surge and the first few months of a surge. The researchers used the remaining 2/3 of the data as validation data (Wallace, 2000:4-1). Comparing the models developed in this research with the proportional model revealed that the physics-based model was much more efficient in forecasting costs during periods of a usage surge. In the example of the C-5B model, the physics-based method underestimated costs by 15%, while the proportional model overestimated costs by as much as 236% (Wallace, 2000:4-5, 4-10).

The findings of this study indicate that both flying hours, and variables associated with sorties (landings and flying hours) are effective in forecasting costs. Intuitively, sorties are a significant predictor of costs for a number of reasons. First, a sortie is associated with starting up and shutting down an aircraft's engine and other flight systems, which put stress on the aircraft. Second, a sortie is also associated with one or more take-offs and landings, which cause significant wear on an aircraft's parts. We will include sorties as a control variable in this study of Air Force bomber aircraft. While our study will omit ground hours as an explanatory variable, it will include temperature, which describes the environment in which the aircraft are stored.

The physics-based study (Wallace, 2000), like the Hildebrandt (1990) report, discussed findings of a "Bathtub Effect," implying a relationship between the age of an aircraft and flying hour costs (Wallace, 2000:2-5). The researchers explained that as parts age, they experience fatigue and corrosion (Wallace, 2000: 3-1). This supports the decision to include aircraft age as an explanatory variable.

Finally, this physics-based study mentions that cannibalizations increase during times of conflict (Wallace, 2000: 4-12). Our research will evaluate the effect of parts removal in bomber aircraft during the global war on terror, as well as the years leading up to it. This will be done to quantify the effect that cannibalization rates have on costs. We discuss cannibalization rates in more detail later in the chapter. The next two studies further highlight problems associated with aircraft age. While Wallace (2000) focused his research on exploring the relationship between costs and flying hours, the Congressional Budget Office (CBO) focused on determining the effect aircraft age has on costs (Kiley, 2001).

CBO Supports Previous Findings.

In the paper "The Effects of Aging on the Costs of Operating and Maintaining Military Equipment," the CBO used regression analysis to measure the relationship between aircraft O&M costs and the following explanatory variables: average fleet age, flying hours, and average aircraft purchase price. The CBO models also include dummy variables for each year (Kiley, 2001:34).

The CBO report investigated the relationship between O&M costs and the independent variables with three models. The first model focused on Air Force aircraft and revealed aircraft age to be positively correlated but statistically insignificant in forecasting O&M costs. However, the second and third models, the first focusing on Navy aircraft and the second aggregating both Air Force and Navy aircraft data, found aircraft age to be a statistically significant predictor of aircraft O&M costs. Specifically, "an additional year of average age is associated with an increase in O&M costs of 1 percent to 3 percent per year" (Kiley, 2001:36). These findings support the findings of Hildebrandt (1990) and Wallace (2000). Additionally, the CBO found average purchase price and flying hours to be positively correlated and statistically significant in forecasting aircraft O&M costs.

The yearly dummy variables were insignificant (Kiley, 2001:36). This could be due to the fact that these variables may be capturing the same effect as the age variable. Unless new aircraft are added to the fleet, both of these variables will increase in equal intervals. Our research of bomber aircraft will account for time, when necessary, by including a lagged cost term as an independent variable. To further capture the effects of time, such as fatigue and corrosion, aircraft age will also be included as an explanatory variable. The next study focuses on how aging equipment affects the number of hours required for maintenance and modifications. Hildebrandt (1990), Wallace (2000), and Kiley (2001) all found aircraft age to be significant in forecasting CPFH. In his study "Aging Aircraft," Pyles (2003) explored this relationship in more detail.

Focusing on Aging Aircraft.

Pyles (2003) examined aircraft fleets' workloads and material consumption patterns for growth associated with age. Using ordinary least squares multiple regression, Pyles (2003) built regression equations for each of 13 maintenance and modification categories (such as flightline maintenance, DLR repair, etc.) (Pyles, 2003:50). Pyles (2003) then used a method of backwards stepwise regression to eliminate statistically insignificant variables in order to arrive at a predictive model. Finally, in order to test for late-life acceleration in maintenance and modification hours, Pyles (2003) incorporated a second-order variable which captures the effect of aircraft age beyond the age of 20. If this variable was significant, he concluded that age was a driver of costs (Pyles, 2003:51).

Pyles (2003) incorporated variables such as flyaway cost, MAJCOM, and aircraft mission (bomber, tanker, cargo, among others). Using this methodology, Pyles (2003) discovered that flyaway cost is highly significant in forecasting maintenance and modification hours (Pyles, 2003:90). After controlling for this factor, Pyles (2003) finds that 11 of the 13 maintenance and modification categories show signs of late-life growth in maintenance and modification hours.

The findings of Hildebrandt (1990), Wallace (2000), Kiley (2001), and Pyles (2003) all support the inclusion of aircraft age as an explanatory variable in CPFH models. However, each of these researchers evaluated models based on data representing multiple platforms. Meanwhile, MAJCOMs create the flying hour budgets based on estimates for each individual MDS-level. The CBO report highlights the fact that studies focusing on individual airframes are more credible and are better able to distinguish between the effects of age from other factors (Kiley, 2001:26). If flying costs are truly correlated with aircraft age, accurately measuring and controlling for these effects will be increasingly important as bomber aircraft continue to age. The next four researchers evaluate individual aircraft and attempt to determine the variables affecting flying costs and to quantify their affects.

Laubacher Explores Cost Trends.

The following study attempted to build a forecasting model by evaluating trends in recent cost data. In his study, "Analysis and Forecasting of Air Force Operating and Support Cost for Rotary Aircraft," Laubacher (2004) focuses on evaluating annual cost data for three helicopters across MAJCOMs. This analysis begins by evaluating the difference between actual costs and the predicted costs in recent years. Laubacher (2004) proceeds to evaluate the actual cost data and attempts to fit a predictive model using three forecasting methods: three-year moving average, single exponential smoothing (SES), and Holt's linear method (Laubacher, 2004:62-64).

Laubacher (2004) develops eight models in his study, one for each of the eight MAJCOMs supporting one of the following three helicopters: the MH-53J, the UH-1N, and the HH-60G. Six of his eight models found Holt's method to be the most predictive. The remaining two models found the SES method to be the best. Additionally, six of his eight models (five using Holt's method and one using the SES method) outperformed the budgets set by the Air Force in the years evaluated.

While Laubacher's (2004) research involved a small data set, his findings suggest that accounting for cost trends in the forecasting process may improve forecasting accuracy. These findings appear to be inconsistent with the results from the pooled CBO models mentioned earlier, which found dummy variables representing years to be insignificant (Kiley, 2001). However, as mentioned, the CBO report also included age as an explanatory variable, which may have acted as a proxy for the effects of time. As previously mentioned, we include both average age and lagged costs as independent variables in this study.

Exploring CPFH for Individual Aircraft.

Three recent AFIT thesis efforts have addressed CPFH for specific aircraft. Hawkes (2005) attempted to find a relationship between the depot-level reparable (DLR) CPFH rate and programmatic and operational explanatory variables for the F-16C/D. He evaluated annual base-level data from FY1998 through 2004. Through OLS regression, Hawkes (2005) concluded that aircraft age, utilization rate (the average number of sorties per aircraft), and location are significantly significant predictors of the DLR CPFH factor for the F-16. Additionally, Hawkes (2005) found average sortie duration (ASD) and controlling for deployments to be insignificant in forecasting CPFH.

The following year, Armstrong (2006) researched the causes of CONS and DLR CPFH factors for the F-15 aircraft. The purpose of this research was to find a marginal CPFH rate based on the number of hours flown in a given time period. Armstrong (2006) used a panel model to evaluate base-level monthly data for the F-15 from FY2001 through 2005. He found ASD and temperature to be significant factors affecting CPFH. Additionally, Armstrong (2006) found a seasonality trend in his data, suggesting that costs increase toward the end of the fiscal year (Armstrong, 2006).

Bryant (2007) also conducted a study based on panel models where he developed forecasting tools for CONS and DLR CPFH factors of the KC-135. Bryant (2007) developed separate models for the ANG, active duty, and reserve units to determine if there were any differences between the three branches. He found mission capable rate (MCR), utilization rate, and total airframe hours (which was evaluated in place of average age) to be significant in forecasting CPFH. He also found a seasonality trend, similar to Armstrong's (2006) findings, in his CONS models; he also found costs to increase toward the end of the fiscal year. Additionally, Bryant (2007) evaluated dew point as a potential predictor of CPFH; however, his research revealed mixed results in the significance of dew point.

Both Hawkes (2005) and Bryant (2007) found utilization rate to be a significant predictor of CPFH, therefore, our research will include utilization rate as a predictor of bomber costs. Research has revealed mixed results for the inclusion of ASD as a predictive variable; however, we will include it in our research of bomber aircraft in an attempt to capture the effects of deployments. The three AFIT theses highlighted above attempted to account for deployments in various ways and none were found to be significant. Armstrong (2006) found temperature to be significant, while Bryant (2007) was unable to draw conclusions on the effects of dew point. This research will attempt to evaluate temperature as a predictor of bomber flying costs. Finally, seasonality will also be evaluated in this research as it was found significant by both Armstrong (2006) and Bryant (2005). The final section in this chapter focuses on cannibalizations.

Cannibalizations Adversely Affect Maintenance.

The Government Accountability Office (GAO) has expressed concern with the military's reliance on cannibalization to keep aircraft in the air. DOD defines cannibalization as "removing serviceable parts from one piece of equipment and installing them into another" (GAO, 2001:1). Units cannibalize aircraft in order to maintain readiness and operational needs. Cannibalizing an aircraft is generally faster than ordering the desired part from the supply system. However, there are several adverse affects to this process, such as increased maintenance costs through increased mechanics' workloads as well as mechanical problems (GAO, 2001:2). This process also leads to grounding aircraft that are missing parts due to cannibalizations (GAO, 2001:17). Additionally, cannibalizations increase the risk of collateral damage of other parts, since removing a specific part usually requires moving nearby parts (GAO, 2001:19).

Since cannibalizing parts involves replacing broken parts with used ones, rather than new ones, and can result in collateral damage (GAO, 2001:19), the level of cannibalizations can affect the performance of aircraft components, potentially resulting in increased flying hour costs. The GAO report (2001) included a list of the cannibalization rates (defined as the number of cannibalizations per 100 sorties) of 91 Air Force, Navy, and Marine Corps aircraft. In FY2000, the B-1B was found to have the highest cannibalization rate at 85.4, while the B-52H was fifth on the list with a cannibalization rate of 30.2; the average cannibalization rate is 9.4 (GAO, 2001:34). While the B-2 was not included on the list, cannibalizations are a significant event for B-1B and the B-52H aircraft. Cannibalization rates appear to have an impact of maintenance levels which may lead to increased costs; therefore, we will include cannibalization rates in the flying cost models for the B-1B and the B-52H.

Summary

Recent studies suggest that basing flying cost estimates on more than simply the number of hours flown will result in more accurate cost forecasts. These additional variables include aircraft age and total operating hours, variables quantifying operations tempo, lagged costs, and environmental characteristics at the bases supporting the aircraft. While each of these studies reveals significant insight into flying costs, none evaluated bomber aircraft specifically. Evaluating the CONS and DLR costs for each of these aircraft may reveal variables which are significant in developing more predictive flying cost models.

Research has also been conducted concerning cannibalizations. Cannibalizations can lead to increased mechanical failures and the grounding of aircraft which have lost parts due to cannibalizations. This research will evaluate the affects of cannibalizations on CPFH for the first time.

Chapter III will discuss the data and methodology used in this study to develop predictive models for CONS and DLR costs for each of the B-1B, B-2, and B-52H. It will also describe how we created the database used in this research.

III. Data and Methodology

Overview

The purpose of this chapter is to describe the data and methodology used to answer the two research questions defined in Chapter I. The methodology discussion will start with a description of the sources which provided the data for this study, as well as how we aggregated these data into a common spreadsheet. In the next step, we will perform regressions of costs on flying hours and compare the results to the current methodology. This will allow us to determine how effective the current cost per flying hour (CPFH) process is in estimating costs. Next we will build new forecasting models through the examination of several other potential predictor variables, in addition to flying hours. Lastly, this chapter will also discuss the diagnostic tests used throughout the model building process.

Data Sources

Four sources provided the data used in this study: the Air Force Total Ownership Cost (AFTOC) database, the Air Force Reliability and Maintainability Information System (REMIS), the Multi-Echelon Resource and Logistics Information Network (MERLIN), and the Air Force Combat Climatology Center (AFCCC). We aggregated the information from these four sources into a single database used for developing the models. The following sections summarize these four data sources. Additionally, Appendix A includes samples from each database. After each of the four data sources are explained, this chapter will discuss the database used in this research, and will define each variable.

AFTOC

The AFTOC database provides a single source of information for Air Force O&M costs, as well as personnel, infrastructure, and research and development costs. AFTOC receives information on transactions affecting wing-level O&M budgets from a series of data feeds that provide daily transaction information. The Air Force Cost Analysis Agency (AFCAA) provided the AFTOC data used in this study. The data included information on monthly base-level costs for both of the dependent variables in question – depot level repairables (DLRs) and consumables (CONS) costs associated with each of the three bomber platforms. The data provided was for nine years: from FY1998 through 2006. AFTOC provided the data in CY2007 dollars.

REMIS

The Air Force REMIS database stores information on maintenance and logistics data for each of the Air Force's weapons systems. AFCAA also provided the REMIS data, which included information on the operational factors that will be evaluated as predictor variables of the costs described above. REMIS provided tail-level monthly data on age, sorties, hours flown, and utilization rate for each of the three bomber platforms. These individual aircraft data were aggregated to the monthly base-level for use in this analysis.

MERLIN

MERLIN provided data for two of the operational factors considered in this research. REMIS supplies the MERLIN data. However, while REMIS provides tail-level information, MERLIN provides base-level data. This research obtained data for mission capable rates and cannibalization rates from the MERLIN database.

AFCCC

The AFCCC database includes historical daily weather observations from over 10,000 locations worldwide. The Air Force uses this data as a planning tool for executing military operations. This database provided the monthly temperature data for each base evaluated in this research. The mean temperature (in degrees Fahrenheit) for each month will be included to capture the effect that temperature has on flying hour costs.

The next section details the variables this research collected from the above sources. The discussion of the dataset will define both the dependent variables and the potential predictor variables.

Dataset

This research built three databases, one for each of the three aircraft analyzed. Each database contains monthly, base-level information on the two dependent variables, and 12 control variables, including flying hours. Each of the variables contained in the three databases are summarized below.

Dependent Variables

We will develop models to forecast two types of cost for three different types of aircraft. These costs include the costs of reparable parts (DLR) for each of the three bombers, as well as the costs of consumable (CONS) parts for each of the three types of bombers.

In our datasets of net costs for the B-1B, we discovered several negative values of net costs for both the DLR and CONS costs. This is due to the system the Air Force uses to manage aircraft parts. If maintenance units are able to repair broken parts that have been removed from aircraft, the units are able to return those parts for credits. Additionally, we found that certain units spent significantly less money than others. This is due to the repairing systems available at different bases; certain bases are equipped to repair certain parts. Therefore, if one of these parts breaks down at Base A, but Base B has the repairing capabilities, Base A may send that part to Base B for repair. This results in Base B receiving the credit for the repaired part, while Base A has to spend money on a new part with no opportunity for a credit.

To correct for this issue of negative costs associated with the credit system, we added a constant term to each observation in the monthly base-level B-1B DLR and CONS datasets. Specifically, we added \$12,070,261 to each observation's net DLR cost and \$5,383 to each observation's net CONS cost. This increased all of the net costs by the same amount, making each of the observations positive, while preserving the relationship between net costs and each of the independent variables. This methodology allows us to take the log of the data without losing any data.

A second method we used to correct for negative costs, in addition to adding a constant term to the net costs, was to group the data at different levels of aggregation. The first level of aggregation involves grouping the base-level data at mission design series (MDS) level. This means that costs incurred at each individual base are summed together for each time period resulting in a single net cost for the entire MDS level. Most of the negative net costs were omitted when the data was aggregated at the MDS level.

The second level of aggregation involved aggregating the monthly data at the quarterly level. We did this for both the base and MDS-level data. This resulted in four datasets for each of the six types of cost: base-level monthly, MDS-level monthly, base-level quarterly, and MDS-level quarterly. Since there is just one B-2 operational base, there is no difference between the base-level and MDS-level data for this airframe. Therefore, this research will result in a total of 20 models. Table 2 presents of list of each of the 20 dependent variables we evaluated. In addition to eliminating negative net costs, aggregating the data provides analysts with the flexibility to use different types of models.

Table 2: List of Dependent Variables

Dependent Variables	
B-1B DLR monthly by base	B-1B CONS monthly by base
B-1B DLR monthly by MDS	B-1B CONS monthly by MDS
B-1B DLR quarterly by base	B-1B CONS quarterly by base
B-1B DLR quarterly by MDS	B-1B CONS quarterly by MDS
B-2 DLR monthly by base/MDS*	B-2 CONS monthly by base/MDS*
B-2 DLR quarterly by base/MDS*	B-2 CONS quarterly by base/MDS*
B-52H DLR monthly by base	B-52H CONS monthly by base
B-52H DLR monthly by MDS	B-52H CONS monthly by MDS
B-52H DLR quarterly by base	B-52H CONS quarterly by base
B-52H DLR quarterly by MDS	B-52H CONS quarterly by MDS
* Since there is only one B-2 base, base-level is equivalent to MDS-level	

DLR. AFTOC provided the monthly net DLR costs for each of the three platforms in CY2008 dollars. These costs are defined as the total charges for DLR parts less the total credits reimbursed to a unit for unused or repaired parts, per the system described previously. AFTOC provided these monthly costs for each base supporting the aircraft evaluated.

CONS. Like DLR costs, AFTOC provided the monthly net CONS costs for each of the three platforms this research evaluated. The net CONS costs are defined similar to net DLR costs; net CONS costs are total charges less total credits. AFTOC provided the information for this variable at the base-level in CY2008 dollars.

Independent Variables

This research will evaluate several control variables as it develops the forecasting models described above. The motivation for choosing these variables was detailed in the previous chapter. The following sections will define each of the independent variables.

Flying Hours. As outlined in the previous chapter, the Air Force builds its flying program budgets based on projected flying hours. Additionally, several studies have identified flying hours as a significant predictor of flying costs, as discussed in Chapter II. Therefore, we will include flying hours as an independent variable. REMIS provided monthly, tail-level flying hour data. In order to use these data in our research, we had to aggregate the data at the base-level.

Average Aircraft Age. The average aircraft age for each base will also be included as an independent variable, as it was found to be significant in forecasting operations related costs. This variable will also act as a proxy for fiscal year fixed effects when aircraft fleets maintain a stable composition of aircraft; that is, no new planes enter the fleet and no existing planes are removed. We calculate this variable as the average age of all the aircraft at a specific location in a specific month. The REMIS database provided the information on aircraft age by tail-number. Therefore, we had to aggregate the information to the base-level in order to calculate average age.

Average Airframe Operating Hours. Average airframe operating hours is defined as the average total hours flown for each aircraft within a fleet. This variable has only been evaluated once and it was found to be significant. Therefore, we will include it in this study in order to further examine its effectiveness. REMIS provided the information on this aircraft characteristic and we will calculate it similar to average aircraft age. Total flying hours for a specific bomber platform will be calculated for a specific base. This number is then divided by the number of aircraft at the location to arrive at average airframe operating hours.

Sorties. A sortie is defined as a flying mission associated with one or more take-offs and landings. Research suggests that flying costs and sorties are positively correlated; that is, as sorties increase costs will increase. Therefore, sorties will also be examined as a predictor of DLR and CONS costs for each aircraft. REMIS also provided data for this variable at the tail-level, se we had to aggregate the tail-level sorties to arrive at total base-level sorties.

Average Sortie Duration. Average sortie duration (ASD) measures the average sortie length for a particular aircraft. We include this variable in our study in order to measure the effect deployments has on costs; a sortie in support of a contingency will be longer than training missions. Using REMIS data, we first calculate ASD for each individual aircraft. Wing-level ASD, or average ASD, is calculated by summing the tail-level ASD data and dividing by the number of aircraft.

Mission Capable Rate. Mission capable rate (MCR) is defined as the percentage of wing possessed aircraft capable of flying at least one specified mission in a given time period. A high MCR for a base means that fewer aircraft are grounded due to maintenance or modification problems. This variable is included in order to determine how readiness affects costs. MERLIN provided this information at the monthly base-level for each of the three aircraft.

Utilization Rate. Utilization rate represents the amount each aircraft is flown in a given time period at a given location. It is defined as the number of sorties flown, on average, per aircraft. That is, the higher the utilization rate for a base, the more sorties flown, per aircraft, at that base. This variable was found to be significant in recent studies. We calculate this variable from the provided REMIS data as the number of sorties flown at a given installation divided by the number of aircraft at that installation.

Average Crude Oil. Recent studies included jet fuel prices as an independent variable in order to capture how flying hour costs are affected by changes in costs in the petroleum industry, and thus the cost of moving parts from one place to another. Studies did not find this variable to be significant. However, this research will analyze average monthly crude oil prices as a proxy for the petroleum industry. The Energy Information Administration provided the data for this variable in cents per gallon.

Month. Previous research found that flying costs may exhibit a seasonality trend associated with the fiscal year; specifically, Armstrong (2006) and Bryant (2007) found costs to increase toward the end of the fiscal year. Therefore, this research will include dummy variables representing the months as independent variables. A dummy variable represents whether or not a variable is present for a specific observation. For example, a dummy variable representing the month of October would have a value of one if the observation took place in the month of October, or zero if the observation took place in any other month.

Temperature. Previous studies have found temperature, and other location characteristics, to be significant predictors of cost. Therefore, this research will examine the effect temperature has on flying costs. AFCCC provided the information on average monthly temperature in degrees Fahrenheit for each base evaluated.

Cannibalization Rate. Cannibalization rate is calculated at the base level. It is defined as the number of times a cannibalization occurs per 100 sorties for a given base. A cannibalization is the removal of a working part from one aircraft in order to install it on another aircraft so that the receiving aircraft can fly a mission. This research will not include this variable in the development of either of the two B-2 forecasting models because the B-2 has not experienced cannibalization. MERLIN provided the monthly, base-level cannibalization rates for each platform.

Summary of Data Collection

This research uses monthly data from FY1998 through FY2006 in estimating the CPFH models built in this analysis. Two types of costs are modeled for each of three bomber platforms, DLR costs and CONS costs. The database provides as many as 108 monthly data points for each of the operational bases supporting bomber aircraft. A sample of the database of monthly B-52H costs is shown in Table 3. Similar databases were built for the monthly costs of the B-1B and B-2, as well as the quarterly costs for the B-1B, B-2, and B-52H. In addition to the variables shown in the table below, dummy variables were created to represent each month and quarter.

Table 3: Sample from Database Developed in This Research

Base	FY	FM	CONS (CY08)	DLR (CY08)	FH	Age	TOH	Sorties	ASD	MCR	U Rate	Oil $	Temp	C Rate
Barksdale	1998	1	1344645.06	9598034.05	1341.90	430.63	14210.88	202.00	6.92	0.77	9.05	21.33	49.12	0.77
Barksdale	1998	2	411421.78	2096488.62	1496.10	431.40	14251.33	196.00	8.58	0.78	10.97	20.19	49.37	0.39
Barksdale	1998	3	748906.02	11005049.03	1151.70	432.45	14214.26	185.00	6.72	0.79	8.46	18.33	52.97	0.26
Barksdale	1998	4	785176.60	7418413.73	1188.70	433.53	14177.39	210.00	6.69	0.77	8.73	16.72	61.48	0.19
Barksdale	1998	5	718688.30	6674792.86	1473.40	434.96	14285.39	195.00	9.34	0.83	11.30	16.06	73.10	0.17
Barksdale	1998	6	1407521.02	7193840.88	1437.80	435.57	14421.20	217.00	7.51	0.77	9.76	15.12	80.63	0.13

Table 3 lists the number of observations collected for each base. As shown in the table, only one base has fewer than 108 observations. Due to reallocation of aircraft, Mountain Home AFB ceased supporting the B-1B by 2002.

Table 4: Observations per Base (FY1998 – FY2006)

Aircraft	Base	Number of Observations
B-1	Dyess	108
	Ellsworth	108
	Mountain Home	44
B-2	Whiteman	108
B-52	Barksdale	108
	Minot	108

* FY1998 – FY2001

After building the datasets, we evaluated simple regressions of cost on flying hours to determine how useful a proportional model is in forecasting costs.

Evaluating the Proportional Relationship

Once we have collected the data and constructed a useful database, we examine the strength of a proportional relationship between costs and flying hours. As described in the previous chapters, the Air Force currently employs a flying cost forecasting technique which involves multiplying the anticipated number of flying hours by a cost per flying hour (CPFH) factor for each type of flying cost. This method is simple and straightforward; however, it assumes that, for a given time period, 1) costs will be zero if no flying occurs, and 2) that costs and flying hours are linearly related, that is, increasing flying hours by one percent will increase costs by one percent.

To test this theory, we will examine regressions of quarterly MDS-level costs for each type of aircraft, DLR and CONS, on flying hours and assess the results. The findings from this step of the analysis will initiate the examination of a new, potentially more predictive, forecasting model.

Building the New Models

We applied ordinary least squares (OLS) regression to the datasets described previously in order to develop the cost forecasting models. An OLS regression model is one which maximizes the amount of explained variation in the dependent variable by minimizing the squared distances between the model's predicted values, based on the explanatory variables, and the actual values given in the dataset.

The general form of OLS regression is shown in the equation below:

$$y_i = \alpha + \beta_1 x_{1i} + \beta_2 x_{2i} + \ldots + \beta_p x_{pi} + \varepsilon_i$$

In this equation, y_i represents the response, or dependent, variable for a specific observation, *i*. The control, or independent, variables are represented by x_{1i}, x_{2i}, …, x_{pi}, for each observation, *i*, while the coefficients for each of the *p* parameters are given by β_1, β_2, …, β_p. The intercept, or constant term, is given by α, and ε_i represents the error term, or residual, in each observation.

Additionally, this research will examine the relationship between both DLR and CONS costs and each of the independent variables using a log-linear model. Taking the log of the dependent variable and certain independent variables means we are now capturing the elasticity of cost on the independent variables (Wooldridge, 2006). Therefore, we are able to describe the variations in the dependent variable as a percentage change given a percentage change in the independent variable; the units are unimportant. Another advantage of using a log-linear model is its ability to reduce heteroskedasticity problems associated with the residuals of the model. Heteroskedasticity, as well as other post-estimation problems, will be discussed in the next section.

While performing a log transformation on certain variables presents several advantages, it also has some major limitations. First, log transforms cannot be used if a variable is zero or negative. As discussed previously, we discovered several B-1B observations with negative net costs due to the credit system employed by the Air Force. These negative costs led us to both aggregate the data at the four levels of aggregation, defined earlier, and to add a constant term to the net costs of the B-1B datasets.

A second limitation of log transformations is that the coefficients on the independent variables represent instantaneous elasticities. Therefore, when we interpret the coefficient as an elasticity, describing how a change in an independent variable relates to a change in the dependent variable, the relationship is an approximation. Further, the interpretation of the coefficients as approximations is only valid for small changes in the independent variable; the larger the change in the dependent variable, the less accurate the approximation. To illustrate this, suppose $x_0 = 40$ and $x_1 = 41$. The percentage change from x_0 to x_1, given by $(x_1 - x_0)/x_0$, is 2.5%. However, $\ln(41) - \ln(40) = 0.0247$, or 2.47%. If $x_1 = 60$, then the change is much larger; the change in x is now 50%. However, $\ln(60) - \ln(40) = 0.4055$, or 40.55%. The approximation is much less accurate when the change is large. Therefore, the equation only yields an approximation for small changes in x. However, for the purposes of this analysis, the log-linear models developed in this paper represent constant elasticity models (Wooldridge, 2006).

Model Diagnostics

To determine if a linear model is appropriate for describing the relationship between costs and flying hours, we perform a series of analyses. We must test for both highly influential data points, and the residual assumptions of a classical linear model: normality, independence, and homoskedasticity.

Influential Data Points

An influential data point is an observation which has a large affect on any part of the regression analysis, such as the estimated coefficients or p-values of the independent variables present in the model. This implies that including the point in the modeling process may distort the accuracy of the regression, as the estimated coefficients and the statistical significance of the independent variables are in question. This analysis will test for influential data points using Cook's Distance. Any data point with a Cook's Distance value greater than 0.5 indicates that the point is a potentially overly influential data point. This point is evaluated and removed if it significantly affects the parameters of the model. If any points are omitted due to large Cook's Distance values, it will be noted in the results tables presented in Chapter IV.

In addition to examining the Cook's Distance, we perform three diagnostic tests to test three assumptions of OLS regression, independence, normality, and constant variance of the residuals.

Independence

The first assumption is that the error terms are independent. This means that the error term has the same variance given any values of the explanatory variables, or that the error term from one period is unrelated to the error term in the following period. To test this assumption of independence, we use the Durbin-Watson statistic to test for the presence of autocorrelation in the residuals; that is, to test for first order serial correlation. With this test, a p-value less than 0.05 means we reject the null hypothesis that the residuals are independent. To control for this, we will graph the autocorrelations of the residuals. This graph will allow us to visually determine if there is a first order serial correlation in the data and will identify that a lag term to include in the model as an independent variable.

Normality and Constant Variance

The remaining two diagnostic tests are the Shapiro Wilk test for normality of the studentized residuals, and the Breusch Pagan test for constant variance of the residuals. In the Shapiro Wilk test, similar to the Durbin Watson test, a p-value greater than 0.05 fails to reject the null hypothesis that the residuals are normally distributed. Likewise, in the Breusch Pagan test, a p-value greater than 0.05 means we fail to reject the null hypothesis that the residuals exhibit constant variance.

In some cases, we found a model to fail one or both of these tests. However, OLS regression is robust to deviations from normality and constant variance. This means that even if a model fails one or both of these tests, the estimated coefficients are unbiased. Therefore, if a visual inspection of the histogram of residuals appears similar to a bell curve, with a peak in the center and is symmetric on both sides, we meet the assumption of normality. Additionally, if a model fails the Breusch Pagan test but a visual inspection reveals that the residuals have a common variance and are spread evenly about the mean, we meet the assumption of constant variance. If a model fails to satisfy one or both of these tests, but a visual inspection reveals that the deviation is a robust deviation from normality or constant variance, it will be noted in the results tables in Chapter IV.

The Shapiro Wilk and Breusch Pagan tests are also very sensitive to outliers. Therefore, some models may fail these tests when all points are included, but pass the tests when an outlier is omitted. In this case, we say that the deviation was verified to be a robust deviation from normality or constant variance. These deviations will also be notes in the results tables in Chapter IV.

Model Utility

To assess the usefulness and predictability of the model, we evaluated both the p-values of the independent variables and the R^2 and adjusted R^2 of the model. In this step of the methodology, we ran models using both a forward regression methodology and a backward regression methodology. In forward regression, we begin with a model consisting of just one independent variable and continue adding variables that are statistically significant at the 95% confidence level, one at a time, until each variable has been evaluated. Backward regression begins with a model consisting of all of the independent variable evaluated. Independent variables which are insignificant at the 95% confidence level are omitted, one at a time, until we achieve a model consisting of only statistically significant variables. Therefore, several models reported in Chapter IV involve a single independent variable, while the model with the most independent variables has just four statistically significant variables.

To measure the explanatory power of the models, we evaluate the R^2 and adjusted R^2. The R^2 tells us what percentage of the variability in the dependent variable is explained by the model. The adjusted R^2 accounts for both the sample size and the number of independent variables included in the model. If the adjusted R2 is significantly smaller than the R^2, then some of the independent variables are not adding predictive value to the model (McClave, 2005). Of each of the final models reported in Chapter IV, the greatest distance between the R^2 and the adjusted R^2 is 0.037 indicating that all of the predictor variables included in each model are useful in explaining the net costs.

Multicollinearity

The independent variables of each model were tested for multicollinearity, which is linear redundancy in the model. This means that some of the independent variables are correlated with each other and have the same effect on the dependent variable, which biases the beta coefficients (Wooldridge, 2006). To determine if any of the models suffered from multicollinearity, we evaluated the variance inflation factor (VIF) of each independent variable. The VIF is a scaled factor which measures how much the variance of an independent variable's beta coefficient is increased due to collinearity with another independent variable. A VIF larger than 5 indicates that the model suffers from multicollinearity. Of each of the final models reported in Chapter IV, the largest VIF for each of the included independent variables is 4.57.

The Final Models

The final models we report in the Chapter IV are built using the methodology outlined in this chapter and satisfy each of the diagnostics just described. The four models for each of the six types of cost are presented in a table. These tables will identify:

- the independent variables included in each model;
- the estimated coefficients for each independent variable, along with the standard errors and p-values (a p-value less than or equal to 0.05 indicates that the variable is statistically significant);
- the R^2 and adjusted R^2 for the model, which reveals how much variation in the response variance is explained by the model; and
- the results of the Shapiro Wilk, Breusch Pagan, and Durbin Watson tests.

Summary

This chapter outlines the methodology used in building the predictive models of DLR and CONS costs for each of the three Air Force bomber platforms, the B-1B, the B-2, and the B-52H. The first step is to collect and normalize the data used in building the forecasting models. Next, we test the current CPFH methodology used by the Air Force by regressing costs on flying hours. Then, we will use OLS regression to determine if accounting for other potential predictor variables may be useful in forecasting flying costs.

Chapter IV summarizes the results of the models we built using the methodology just described. In Chapter V we will highlight the conclusions that were drawn based on the results obtained in Chapter IV.

IV. Results and Analysis

Overview

This chapter presents the results of each of the predictive models discussed in Chapter III. First, we present the results of using the current proportional model, based on flying hours, to forecast costs. Next, we present the final models for each type of cost, developed using the methodology outlined in Chapter III. We developed and examined four models for each of the following six costs: B-1B depot level reparable (DLR) costs, B-1B consumables (CONS) costs, B-2 DLR costs, B-2 CONS costs, B-52H DLR costs, and B-52H CONS costs.

Proportional Models

As discussed in Chapter II, the Air Force currently uses a cost per flying hour (CPFH) factor to determine flying program budgets. As described in Figure 4, this methodology is based on two assumptions. First, it assumes zero costs when zero hours are flown. Second, it assumes a linear, proportional relationship between costs and flying hours. That is, when flying hours increases by one percent, costs will increase by one percent.

The first step of this analysis is to test the hypothesis that flying hours and costs are related proportionally. To do this, we performed simple regressions of cost on flying hours to see if 1) the intercept was zero, and 2) the coefficient on flying hours was equal to the CPFH factor used in the proportional methodology. The equations below define the proportional relationships the Air Force uses to forecast DLR and CONS costs.

$$DLR_t = \text{CPFH factor}_{DLR} * \text{Projected Flying Hours}_t$$

$$\text{CONS}_t = \text{CPFH factor}_{\text{CONS}} * \text{Projected Flying Hours}_t$$

To test this relationship, we perform simple regressions of net DLR and net CONS costs on flying hours.

Net DLR Costs

Table 4 shows the results of regressing quarterly mission design series (MDS) level DLR costs on flying hours. The table shows both the estimated intercept and the estimated coefficient for flying hours. Examining the B-1B quarterly model, for example, tells us that when zero hours are flown, we can anticipate DLR costs of $66,936,060 (CY08$). Additionally, each flying hour will increase costs by an additional $3,754 (CY08$). The rightmost column of the table shows the FY2008 CPFH factor, provided by the Air Force Total Ownership Cost (AFTOC) database, used to project active duty net DLR costs.

The table reveals significant disagreement between the data and the current forecasting method used to project flying costs. First, Table 5 shows that our regressions reveal a multi-million dollar intercept for both the B-1B and the B-52H; this is consistent with substantial fixed costs associated with flying programs that are now accounted for in the proportional model. This contradicts the first assumption underlying the proportional model used by the Air Force; flying zero hours does not imply zero DLR costs. Additionally, the estimated coefficient for flying hours for each of the three aircraft is less than the current CPFH factor used by the Air Force.

The B-2 model also yields a multi-million dollar intercept; however, the intercept is not statistically different from zero, as indicated by the high p-value. Additionally, the B-2 model revealed a coefficient less than the CPFH factor; however, this coefficient is not statistically different from the CPFH factor at the 95% confidence interval. The results of the B-2 model indicate that the current proportional CPFH methodology may be useful in forecasting B-2 DLR costs.

We ran an F test to test the null hypothesis that the regression and proportional models are statistically equivalent. Our findings indicate that we fail to reject the null hypothesis that the models are statistically equivalent. Therefore, the intercept is not statistically different from zero and the proportional method is appropriate in modeling the B-2 DLR data.

Table 5: Regressing DLR Costs on Flying Hours

Dependent Variable: MDS-Level Deport Level Repairables Costs (FY08$)					
Model		Intercept	Flying Hours	Adj R^2	FY2008 AD CPFH Factor
B-1B Quarterly	Coefficient	66,936,060	3,754	0.101	$18,664
	Standard Error	*7,691,020*	*1,688*		
	p-Value	*<0.001*	*0.033*		
B-2 Quarterly	Coefficient	2,372,281	11,850	0.391	$12,990
	Standard Error	*3,693,852*	*2,448*		
	p-Value	*0.525*	*<0.001*		
B-52H Quarterly	Coefficient	17,524,921	3,935	0.433	$5,633
	Standard Error	*4,100,060*	*747*		
	p-Value	*<0.001*	*<0.001*		

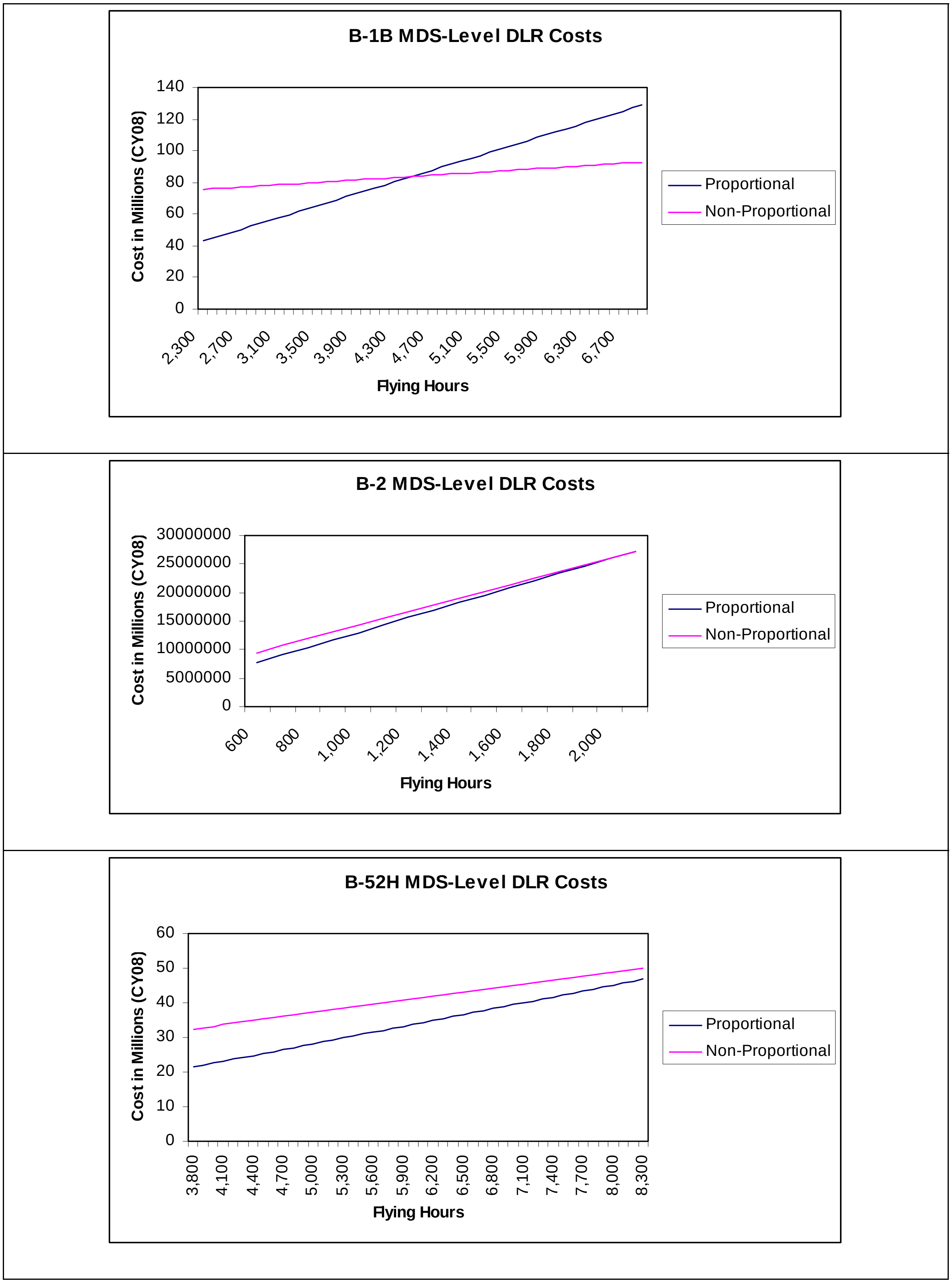

Figure 6: Proportional Forecasts versus Non-Proportional Forecasts for DLR Costs

The charts in Figure 6 show the relationship between the proportional models used by the Air Force and the non-proportional models created by regressing DLR costs on flying hours. The B-1B model closely resembles Figure 4 from Chapter I. Given the different slopes between the regression line and the proportional relationship, the intersection between the two lines represents the point where the proportional model shifts from underestimating flying costs to overestimating flying costs. The B-2 data reveals that the proportional model and our model yield similar forecasts. That is, given historical levels of flying hours, ranging from 596 to 1,976 per quarter, the intercept calculated in Table 5 is insignificant at the 95% confidence level. Evaluating the chart of B-52H DLR costs reveals that the proportional forecasting method consistently underestimates costs, given the range of historical flying hours.

Net CONS Costs

Table 6 shows the results of regressing MDS-level CONS costs on flying hours. Similarly to the results of the models representing net DLR costs for the B-1B and B-52H discussed previously, we find that the data indicate an intercept, while the coefficient on flying hours is considerably less than the CY2008 CPFH factor used in the proportional model. Surprisingly, CONS costs for the B-2 reveal a negative estimated coefficient for flying hours. This indicates that increasing flying hours will reduce overall CONS costs for the B-2. However, the p-value for this coefficient is 0.749, indicating that it is not statistically different than zero. Therefore, we can conclude that flying hours is not a predictor of B-2 CONS costs.

We make a visual comparison between the proportional CPFH models and the regressions developed in this research in Figure 7. All three charts reveal that the proportional method underestimates CONS costs at low levels of flying hours and overestimates costs at high levels of flying hours.

Table 6: Regressing CONS Costs on Flying Hours

Dependent Variable: MDS-Level Consumables Costs (FY08$)						
Model		Intercept	Flying Hours	Adj R^2	FY 2008 AD CPFH Factor
B-1B Quarterly	Coefficient	7,418,918	623	0.053	$1,933
	Standard Error	*1,652,172*	*363*		
	p-Value	*<0.001*	*0.095*		
B-2 Quarterly	Coefficient	1,803,612	-86	-0.026	$969
	Standard Error	*403,864*	*268*		
	p-Value	*<0.001*	*0.749*		
B-52H Quarterly	Coefficient	1,710,543	575	0.235	$932
	Standard Error	*920,520*	*168*		
	p-Value	*0.072*	*0.002*		

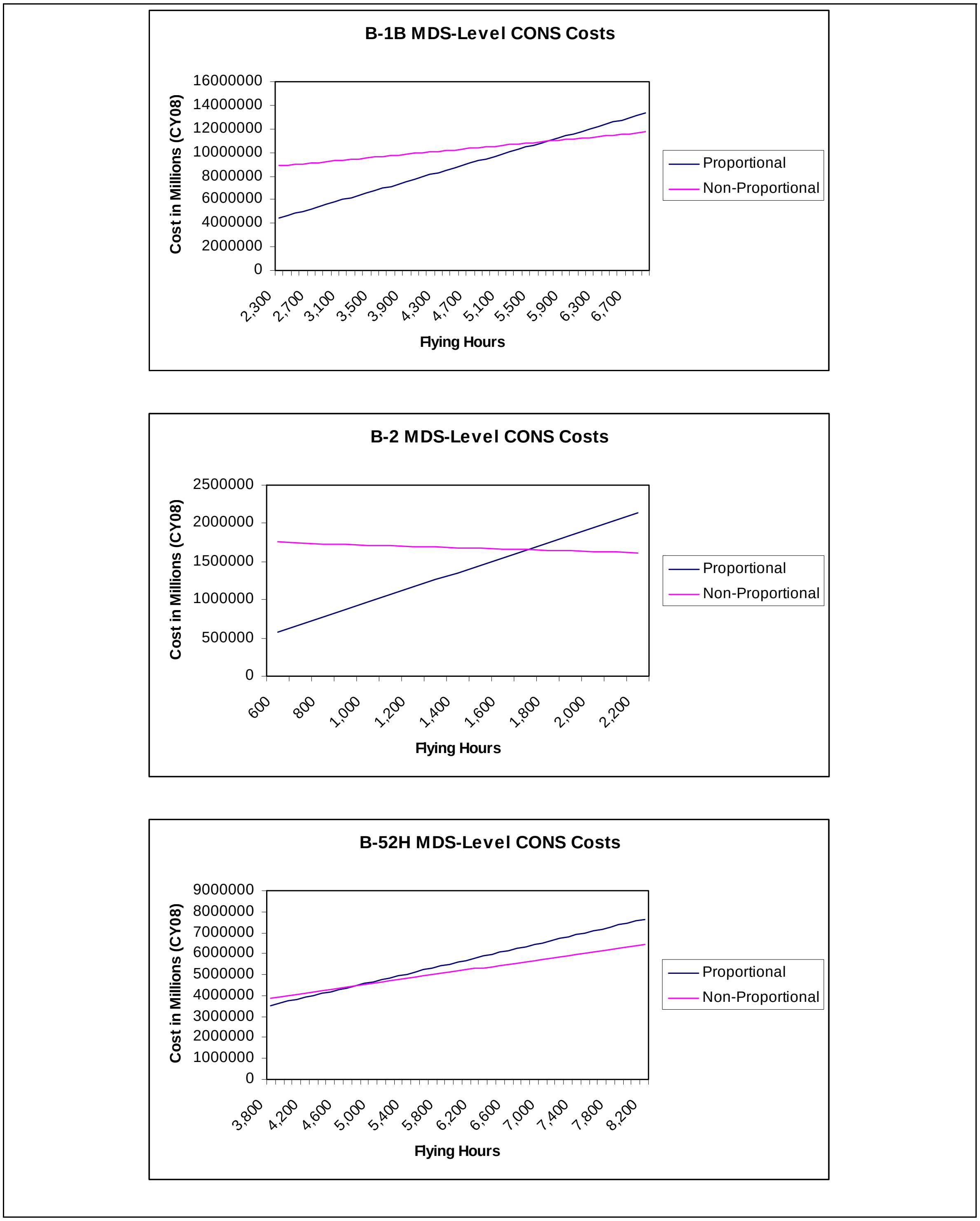

Figure 7: Proportional Forecasts versus Non-Proportional Forecasts for CONS Costs

The findings discussed previously indicate that the current proportional method of forecasting flying costs may not be appropriate for these three aircraft. The next step of this analysis, as described in Chapter III, is to apply OLS regression techniques to several variables flagged as potential cost drivers in previous research. Our goal is to determine if controlling for factors other than flying hours is useful in forecasting flying costs. First, we examine B-1B DLR costs.

B-1B DLR Model

This section will summarize the models of DLR costs for the B-1B. We developed five models associated with B-1B DLR costs. Table 7 shows the results for each regression. The first two models developed for the B-1B represent monthly base-level data. The remaining three models represent monthly MDS-level costs, quarterly base-level costs, and quarterly MDS-level costs.

Our analysis revealed that the factors that drive costs at Dyess AFB are different that the factors that drive costs at Ellsworth AFB and Mountain Home AFB. As we developed the model for Dyess AFB, we found that the model's residuals failed the Durbin-Watson test of independence at a 95% confidence level (an alpha level of 0.05). Therefore, the log of the previous month's DLR cost was included as an independent variable to correct for the lack of independence. The final model for Dyess AFB relates the log of monthly DLR costs to the log of the previous month's DLR costs. Once again, we evaluated the log of net costs as the dependent variable to correct for non-constant variance.

$$ln(Monthly\ DLR)_{(t)} = 9.411 + 0.449 * ln(Monthly\ DLR)_{(t-1)}$$

The estimated coefficient on the log of the previous month's DLR costs for Dyess AFB is 0.449, indicating that a one percent increase in last month's net DLR costs will increase the current month's costs by 0.449%. Therefore, to forecast net DLR costs for Dyess AFB, we must know the previous month's net DLR costs. Surprisingly, flying hours was not found to be a statistically significant predictor of DLR costs for Dyess AFB. The R^2 for the model of monthly DLR costs at Dyess AFB is 0.201, indicating that our model explains 20.1% of the variation in monthly DLR costs at Dyess AFB. Table 7 also provides the p-values of each of the three diagnostic tests performed on the model to test three assumptions of OLS regression: independence, normality, and constant variance of the model residuals. As defined in Chapter III, a p-value of 0.05 or larger satisfies each of the three tests. For the Dyess AFB model, the p-value for the Durbin Watson test is 0.593. Therefore, our model satisfies the assumption of independence.

Before explaining the results of the Shapiro Wilk test for normality and the Breusch Pagan test for constant variance, we must first explain a few terms. First, as discussed in Chapter III, OLS regression is robust to deviations from normality and constant variance (McClave, 2005:572). This means that even if a model fails one or both of the tests, the estimated coefficients are unbiased.

In this research, if we find that a model fails to satisfy the Shapiro Wilk test, we inspect a histogram of the studentized residuals. If we find that removing one or two points results in satisfying the Shapiro Wilk test, we say that the deviation was verified to be a robust deviation from normality. If we find that the model fails the Shapiro Wilk test, but the studentized residuals appear to be approximately normal, we say that the deviation was visually determined to be a robust deviation from normality.

Table 7: B-1B DLR Cost Models

Dependent Variable: Log of B-1B Depot Level Reparables Costs														
Model		Intercept	Log of Flying Hours	Log of Sorties	Log of ASD	Utilization Rate	Age	Cohort	Lag 1 of ln(DLR)	R^2 Adj R^2	Number of Obs	Durbin Watson	Shapiro Wilk	Breusch Pagan
Base Monthly (Dyess)	Coefficient	9.411	--	--	--	--	--	--	0.449	0.201 0.193	107	0.593	0.009**	0.255
	Standard Error	*1.491*	–	–	–	–	–	–	*0.087*					
	p-Value	*<0.001*	–	–	–	–	–	–	*<0.001*					
Base Monthly (Ellsworth, Mt Home)	Coefficient	15.992	0.189	--	-0.325	--	--	--	--	0.288 0.279	148	0.599 0.283	0.314	0.053
	Standard Error	*0.122*	*0.025*	–	*0.066*	–	–	–	–					
	p-Value	*<0.001*	*<0.001*	–	*<0.001*	–	–	–	–					
MDS Monthly	Coefficient	14.715	0.328	--	--	--	--	--	--	0.087 0.078	108	0.191	<0.001*	0.001*
	Standard Error	*0.745*	*0.103*	–	–	–	–	–	–					
	p-Value	*<0.001*	*0.002*	–	–	–	–	–	–					
Base Quarterly	Coefficient	4.458	--	0.849	--	--	-0.006	-1.027	0.510	0.811 0.801	82	0.834 0.740 0.109	<0.001*	<0.001**
	Standard Error	*0.924*	–	*0.176*	–	–	*0.002*	*0.165*	*0.084*					
	p-Value	*<0.001*	–	*<0.001*	–	–	*0.004*	*<0.001*	*<0.001*					
MDS Quarterly	Coefficient	14.714	0.308	--	--	0.046	--	--	--	0.403 0.366	36	0.622	0.006*	0.027*
	Standard Error	*1.756*	*0.083*	–	–	*0.013*	–	–	–					
	p-Value	*<0.001*	*0.001*	–	–	*0.001*	–	–	–					

* Verified to be a robust deviation from normality (Shapiro Wilk) or constant variance (Breusch Pagan)

** Visually determined to be a robust deviation from normality (Shapiro Wilk) or constant variance (Breusch Pagan)

We conduct a similar process if a model fails the Breusch Pagan test for constant variance. We begin by inspecting the residuals versus fitted plot. If we find that removing one or two points results in satisfying the Breusch Pagan test, we say that the deviation was verified to be a robust deviation from constant variance. If we find that a model fails the Breusch Pagan test, but the residuals appear to exhibit constant variance, then we say that the deviation was visually determined to be a robust deviation from constant variance.

The model of monthly DLR costs for Dyess AFB fails the Shapiro Wilk test with a p-value of 0.009. The histogram of studentized residuals, along with a normal quantile plot, is shown in Figure 8. These graphs reveal that a single point may be causing the deviation from normality.

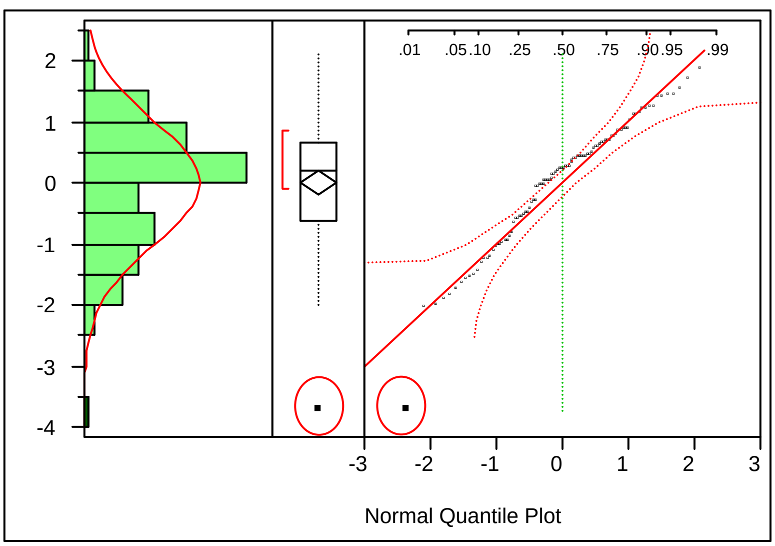

Figure 8: Histogram and Normal Quantile Plot of the Studentized Residuals

The data point of interest represents an observation where DLR costs are much lower than the rest of the dataset. Therefore, the model significantly overestimates DLR costs for this time period, resulting in a very low studentized residual.

Figure 9 shows the histogram and normal quantile plot when this data point is excluded. The p-value of the Shapiro Wilk test is 0.035 when this point is excluded. However, the studentized residuals appear to be approximately normal. Therefore, we visually determined that this deviation is a robust deviation from normality.

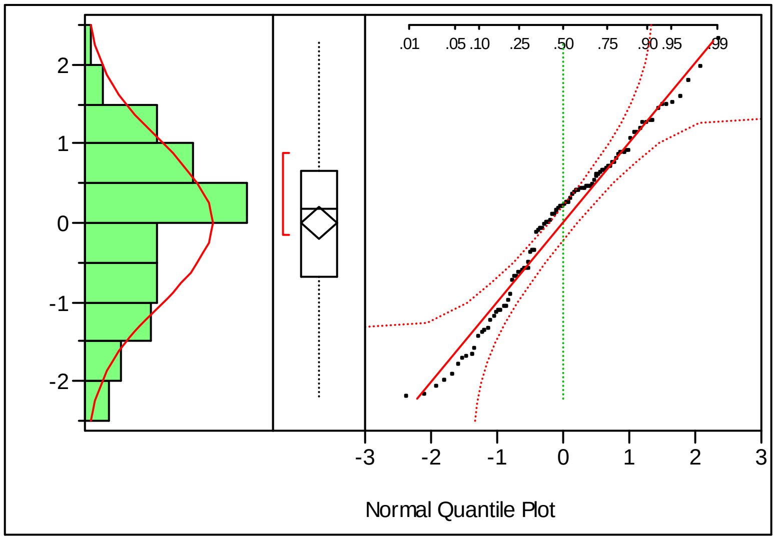

Figure 9: Histogram and Normal Quantile Plot of Studentized Residuals for Model 1 (Excluding Outlier)

The third diagnostic test shown in Table 7 is the Breusch Pagan test. The model of monthly DLR costs for Dyess AFB satisfies the Breusch Pagan test with a p-value of 0.225.

The second model in Table 7 relates the log of monthly DLR costs for Ellsworth AFB and Mountain Home AFB to the log of flying hours and the log of average sortie duration (ASD).

$$ln(Monthly\ DLR) = 15.992 + 0.189 * ln(Flying\ Hours) – 0.325 * ln(ASD)$$

The estimated coefficient on the log of flying hours is 0.189, indicating that a one percent increase in flying hours leads to a 0.189% increase in net CONS costs. This finding indicates that an increase in flying hours has about one-fifth the effect on cost as is assumed in the proportional model. The log of ASD is negatively correlated with monthly DLR costs and has an estimated coefficient of -0.325. This indicates that increasing the average sortie duration for a fleet of B-1Bs at Ellsworth AFB or Mountain Home AFB by one percent will lead to a reduction in DLR costs of 0.325%. The R^2 for this model is 0.288, indicating that 28.8% of the variance in DLR costs at Ellsworth AFB and Mountain Home AFB is explained by our model. Additionally, this model yields satisfactory results for all three diagnostic tests.

The third model in Table 7 relates the log of MDS-level monthly costs to the log of flying hours.

$$ln(Monthly\ DLR) = 14.715 + 0.328 * ln(Flying\ Hours)$$

This model indicates that a one percent increase in flying hours will increase costs by 0.328%. The R^2 for this model is 0.087, indicating that only 8.7% of the variation in the dependent variable is explained by this model. The p-value of the Shapiro Wilk test for this model is 0.001 indicating that we reject the null hypothesis of normality. However, we verified the deviation to be robust deviation from normality. Additionally, the model failed to satisfy the Breusch Pagan test. The residual versus fitted plot is shown in Figure 10.

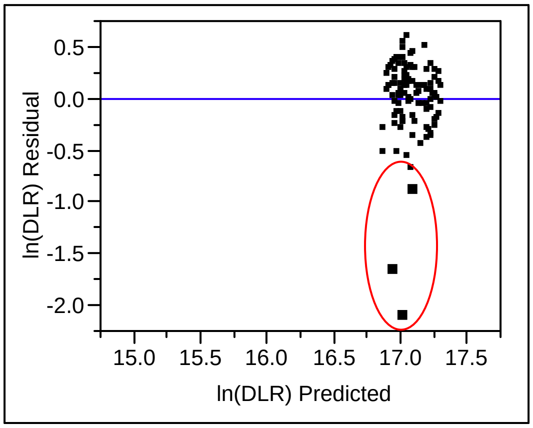

Figure 10: Residuals Versus Fitted Plot for Model 3

We determined that the three circled points were causing the model to fail the test for constant variance. Therefore, omitting these three points resulted in the residuals versus fitted plot shown in Figure 11.

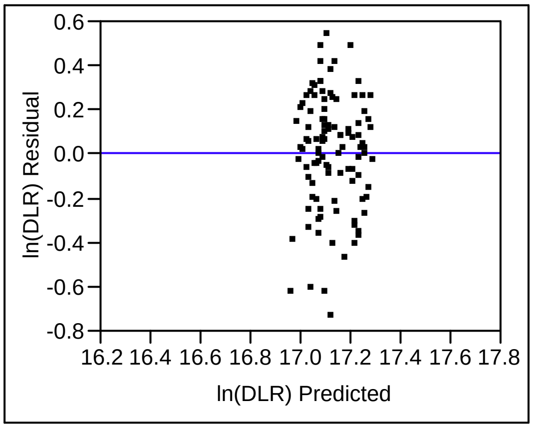

Figure 11: Residuals Versus Fitted Plot for Model 3 (Excluding Anomalies)

The p-value for the Breusch Pagan test is now 0.204. We fail to reject the null hypothesis that residuals exhibit constant variance. Therefore, we have verified that the deviation from constant variance is a robust deviation.

The fourth model in Table 7 relates the log of base-level quarterly DLR costs to the log of sorties, age, the log of the previous quarter's DLR costs, and a cohort.

$$ln(Quarterly\ DLR)_{(t)} = 4.458 + 0.849 * ln(Sorties) - 0.006 * Age$$
$$- 1.027 * Cohort + 0.510 * ln(Quarterly\ DLR)_{(t-1)}$$

The cohort represents four observations – one from Mountain Home AFB (a base which no longer flies B-1B aircraft, as discussed in Chapter III) and three from Ellsworth AFB – which exhibit very low net DLR costs. Specifically, the average net DLR cost in this dataset is approximately $27 million (CY08$), while all four of these points represent DLR costs less than $14 million (CY08$). Without the cohort dummy variable, our model overestimates net DLR costs for these observations.

This model indicates that an increase in last quarter's DLR cost will lead to an increase in the current quarter's DLR cost. However, the observations in the cohort have costs that are very low relative to the previous quarter's DLR costs; each point has a current DLR cost that is more than 60% less than the costs in the previous quarter.

Additionally, our model indicates that an increase in sorties leads to an increase in net DLR costs. The observations in the cohort represent observations with higher than average sorties. Specifically, the Ellsworth AFB observations represent quarters where 818 sorties or more were flown compared to an average of 371 sorties.

As discussed previously, net DLR costs represent DLR charges less DLR credits. In examining the charges and credits for this dataset, we found that the cohort represents observations where the units experience large credits relative to charges. Specifically, the credits for these observations are greater than or equal to 74% of the total charges, while the credits for the rest of the dataset are an average of 58% of the total charges. Further, the cohort represents observations where the DLR charges in the previous quarter are relatively high at Dyess AFB. Specifically, Dyess AFB experiences an average of $60M (CY08) in charges, while the average for the cohort is $76M (CY08). Ellsworth AFB provides repair capabilities that are not available at Dyess AFB. Therefore, when certain parts break at Dyess AFB they are shipped to Ellsworth AFB for repair and Ellsworth AFB gets the credits for the repairs it does on parts from Dyess AFB. The increase in quarterly charges at Dyess AFB combined with the increase in credits at Ellsworth AFB in the following quarter indicate that Ellsworth AFB may have repaired, and received credits for, parts removed from aircraft at Dyess AFB.

When using this model to forecast costs, Air Force estimators can anticipate that a quarter will classify as an observation in this cohort if 1) the observation is for Ellsworth AFB (Mountain Home AFB no longer supports B-1B aircraft), and 2) they anticipate a number of sorties that is significantly over the average and a high level of credits in relation to charges for a given quarter.

The log of sorties has an estimated coefficient of 0.849, indicating that a one percent increase in the number of sorties flown in a quarter will cause DLR costs to increase by 0.849%, all else being equal. Net DLR costs and age are negatively correlated, as shown by the estimated coefficient for age; -0.006. This indicates that increasing the average age of a B-1B fleet by one month leads to a decrease in costs of 0.6%. The estimated coefficient on the log of the previous quarter's DLR costs is 0.510, indicating that a one percent increase in DLR costs last quarter will lead to a 0.51% increase in costs this quarter. The estimated coefficient for the cohort is -1.027, indicating that these six points have DLR costs that are significantly less than the rest of the dataset, all else being equal. The R^2 for this model is 0.811. Therefore, this model explains 81.1% of the variation in DLR costs. Additionally, this model failed both the Shapiro Wilk and Breusch Pagan tests. The deviation from normality was verified to be a robust deviation. The deviation from constant variance was visually determined to be a robust deviation.

The fifth model of B-1B DLR costs, relates the log of MDS-level quarterly costs to the log of flying hours and utilization rate.

$$ln(Quarterly\ DLR) = 14.714 + 0.308 * ln(Flying\ Hours) + 0.046 * Utilization\ Rate$$

The estimated coefficient for the log of flying hours is 0.308, indicating that a one percent increase in the number of flying hours in a quarter with increase net DLR costs by 0.308%. This model indicates that flying hours has a much smaller effect on cost than is assumed in the proportional model. The estimated coefficient for utilization rate is 0.046. This indicates that increasing the utilization rate by one sortie per aircraft increases costs by 4.6%, all else being equal. With an R^2 of 0.403, 40.3% of the variation in DLR costs is explained by our model. Additionally, this model failed both the Shapiro Wilk test and the Breusch Pagan test. However, both failures were verified to be robust deviations.

Interestingly, of each of the four levels of aggregation evaluated, these findings indicated that the model representing base-level quarterly costs is the most explanatory. Additionally, this model does not include flying hours as an independent variable; the model suggests that variables such as sorties and lagged costs may be more correlated with net DLR costs for the B-1B than flying hours.

B-1B CONS Model

This section describes the results of the four models developed to model net CONS costs for the B-1B. All of the B-1B CONS models are listed in Table 8. The first two models represent base-level monthly costs and MDS-level monthly costs. The third model, representing base-level quarterly costs, was split into two models, one representing Dyess AFB and Mountain Home AFB, and one for Ellsworth AFB. We had to build two models because we found that costs at Ellsworth AFB are driven by different factors than those that drive costs at Dyass AFB and Mountain Home AFB. The final model represents MDS-level quarterly CONS costs for the B-1B.

The first model relates monthly base-level B-1B CONS costs to the log of sorties and the log of the previous month's CONS costs.

$$ln(Monthly\ CONS)_{(t)} = 0.922 + 0.363 * ln(Sorties) + 0.806 * ln(Monthly\ CONS)_{(t-1)}$$

The estimated coefficient on the log of sorties is 0.363, indicating that a one percent increase in the number of sorties flown in a given month will lead to a 0.363% increase in net CONS costs. The log of the previous month's CONS costs has a coefficient of 0.806, indicating that a one percent increase in the previous month's costs will result in a 0.806% increase in the current month's CONS costs. This model has an R^2 of 0.832, indicating that the model explains 83.2% of the variation in the dependent variable. While this model failed to satisfy both the Shapiro Wilk and Breusch Pagan tests, we visually determined the deviations to be robust deviations from normality and constant variance.

Table 8: B-1B CONS Cost Models

Dependent Variable: Log of B-1B Consumables Costs												
Model		Intercept	Log of Sorties	November	Year Effect	Lag 1 of ln(CONS)	Lag 2 of ln(CONS)	R^2 Adj R^2	Number of Obs	Durbin Watson	Shapiro Wilk	Breusch Pagan
Base Monthly	Coefficient	0.922	0.363	--	--	0.806	--	0.832 0.830	254	0.969	<0.001**	0.001**
	Standard Error	*0.368*	*0.085*	--	--	*0.033*	--			0.999		
	p-Value	*0.011*	*<0.001*	--	--	*<0.001*	--			0.621		
MDS Monthly	Coefficient	4.373	--	-0.511	--	0.711	--	0.619 0.612	107	0.814	0.050	<0.001**
	Standard Error	*0.879*	--	*0.096*	--	*0.059*	--					
	p-Value	*<0.001*	--	*<0.001*	--	*<0.001*	--					
Base Quarterly (Dyess, Mountain Home)	Coefficient	1.33	0.998	--	--	0.52	--	0.787 0.778	50	0.529 0.270	<0.001*	<0.001**
	Standard Error	*1.021*	*0.267*	--	--	*0.116*	--					
	p-Value	*0.199*	*0.001*	--	--	*<0.001*	--					
Base Quarterly (Ellsworth)	Coefficient	13.322	--	--	-2.979	--	0.108	0.961 0.958	33	0.065	0.330	0.477
	Standard Error	*0.702*	--	--	*0.257*	--	*0.048*					
	p-Value	*<0.001*	--	--	*<0.001*	--	*0.031*					
MDS Quarterly	Coefficient	9.605	--	--	--	0.405	--	0.260 0.237	35	0.715	<0.001*	0.075
	Standard Error	*1.911*	--	--	--	*0.119*	--					
	p-Value	*<0.001*	--	--	--	*0.002*	--					

* Verified to be a robust deviation from normality

** Visually determined to be a robust deviation from normality (Shapiro Wilk) or constant variance (Breusch Pagan)

The second model relates the log of MDS-level monthly CONS costs to a dummy variable representing the costs occurring in November and the log of the previous month's CONS costs.

$$ln(Monthly\ CONS)_{(t)} = 4.373 - 0.511 * November + 0.711 * ln(Monthly\ CONS)_{(t-1)}$$

The model indicates that costs occurring in the month of November are significantly less than costs in each of the other months. We hypothesize that this is due to the variations in cost associated with the fiscal year. Costs tend to be greater toward the end of the year as units attempt to deplete their budgets, and costs are less at the beginning of the year as units begin planning their spending patterns. Additionally, the model suggests that a one percent increase in the previous month's CONS costs leads to a 0.711% increase in the current month's costs. This model explains 61.9% of the variation in MDS-level monthly CONS costs for the B-1B. The model fails the Breusch Pagan test; however, the deviation was visually determined to be a robust deviation from constant variance.

The third model in Table 8 models the quarterly net CONS costs for B-1Bs at Dyess AFB and Mountain Home AFB. As discussed, we found that the factors that drive B-1B CONS costs at Dyess AFB and Mountain Home AFB are different than the factors that drive costs at Ellsworth AFB. This model finds the log of sorties and the log of the previous month's CONS costs to be statistically significant in forecasting the log of costs.

$$ln(Quarterly\ CONS)_{(t)} = 1.330 + 0.998 * ln(Sorties) + 0.520 * ln(Quarterly\ CONS)_{(t-1)}$$

The model suggests that a one percent increase in sorties in a given quarter results in a 99.8% increase in net CONS costs, a significantly larger impact than in the base-level monthly model. Additionally, a one percent increase in the previous quarter's CONS costs results in a 0.520% increase in the current quarter's CONS costs. This model explains 78.7% of the variation in quarterly net CONS costs for Dyess AFB and Mountain Home AFB. While the model failed both the Shapiro Wilk and Breusch Pagan tests, both deviations were determined to be robust deviations.

The model of quarterly CONS costs for Ellsworth relates the log of costs to a year effect and the log of the second lag of CONS costs, that is, the CONS costs from two quarters prior.

$$ln(Quarterly\ CONS)_{(t)} = 13.322 - 2.979 * Year\ Effect + 0.108 * ln(Quarterly\ CONS)_{(t-2)}$$

The year effect in this model represents the data points from the first quarter of FY1998 through the second quarter of FY1999. Costs during this time period were significantly lower than the costs in the rest of the dataset. We controlled for the differences in costs for this time period with a dummy variable. The model also suggests that a one percent increase in net CONS costs from two quarters prior leads to a 0.108% increase in the current quarter's CONS costs. This model explains 96.1% of the variation in quarterly CONS costs for Ellsworth AFB. This model satisfies all the diagnostic tests.

This model of quarterly net CONS costs for Ellsworth also suffered from an overly influential data point. The overlay plot of Cook's Distance values is shown in Figure 12. The plot reveals two data points that may be overly influential. The circled point represents a quarter where CONS costs were very high compared to two quarters prior. Therefore, the model significantly underestimates the CONS costs for this time period. Figure 13 shows the overlay plot of the Cook's Distance values after the anomaly is removed. Removing this data point causes the coefficient on the year effect to decrease from -2.493 to -2.979. Additionally, the coefficient on the lag variable decreased from 0.162 to 0.108. Removing this single point caused each of the other points to yield Cook's Distance values which indicate that they are not overly influential. The model reported in Table 8 does not include the overly influential data point.

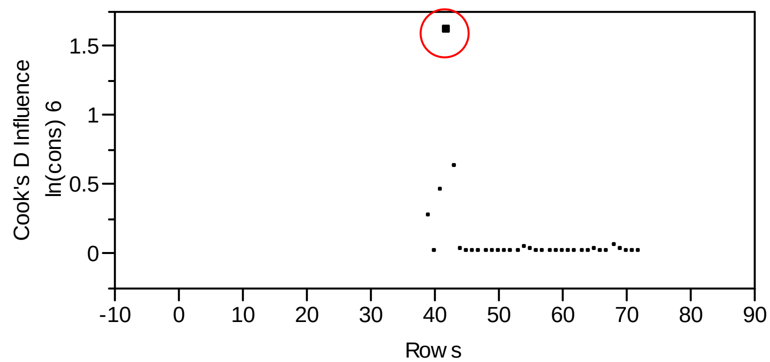

Figure 12: Overlay Plot of Cook's Distance Values for Model 4

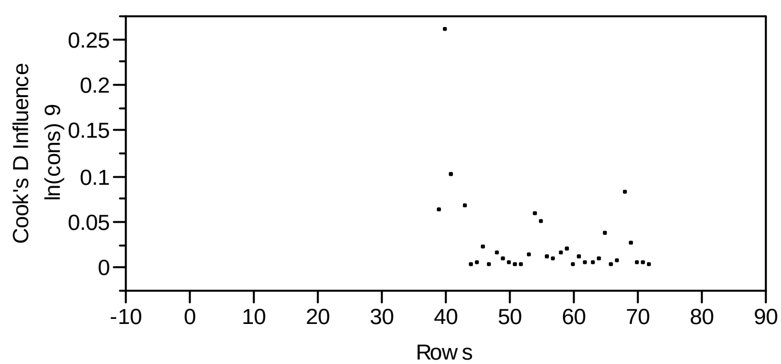

Figure 13: Overlay Plot for Cook's Distance for Model 4 (Excluding the Anomaly)

The final model of B-1B net CONS costs relates the log of MDS-level quarterly costs to the log of the previous quarter's costs.

$$ln(Quarterly\ CONS)_{(t)} = 9.605 + 0.405 * ln(Quarterly\ CONS)_{(t-2)}$$

This model indicates that a one percent increase in net CONS costs in the previous quarter will result in a 0.405% increase in CONS costs in the current quarter. This model explains 26.0% of the variance in MDS-level quarterly costs. While this model failed the Shapiro Wilk test, it was verified to be a robust deviation from normality.

Of each of the models representing B-1B net CONS costs, we found that the three models representing base-level costs are more explanatory than the MDS-level models. Specifically, the base quarterly model for Ellsworth AFB yields the largest R^2 value None of the base-level models include flying hours as an independent variable, rather, they indicate that sorties and lagged costs may be more correlated with CONS costs for the B-1B.

B-2 DLR Model

This section summarizes the two models we developed to model B-2 DLR costs. As mentioned in the previous chapter, since there is only one base which supports B-2 aircraft, base-level and MDS-level costs are equivalent. The two models developed in the section, MDS-level monthly DLR costs and MDS-level quarterly DLR costs, are summarized in Table 9.

The first model relates the log of monthly B-2 DLR costs to the log of flying hours, average age, and the dummy variable representing the month of November.

$$ln(Monthly\ DLR) =$$

$$= 10.858 + 0.689 * ln(Flying\ Hours) + 0.006 * Age – 1.505 * November$$

The model suggests that a one percent increase in flying hours will increase B-2 DLR costs by 0.689%. Additionally, as a fleet ages by one month, costs will increase by 0.6%. Likewise, as a fleet ages by one year, costs will increase by 7.2%. Finally, we found that costs incurred in the month of November are significantly less than costs incurred

Table 9: B-2 DLR Cost Models

Model		Intercept	Log of Flying Hours	Age	November	R^2 / Adj R^2	Number of Obs	Durbin Watson	Shapiro Wilk	Breusch Pagan
		Dependent Variable: Log of B-2 Depot Level Reparables Costs								
Base/MDS Monthly*	Coefficient	10.858	0.689	0.006	-1.505	0.521 / 0.507	105	0.446	0.204	0.696
	Standard Error	*1.005*	*0.182*	*0.002*	*0.21*					
	p-Value	*<0.001*	*<0.001*	*0.009*	*<0.001*					
Base/MDS Quarterly	Coefficient	10.998	0.725	0.005	--	0.592 / 0.567	36	0.140	0.629	0.325
	Standard Error	*1.657*	*0.250*	*0.002*	–					
	p-Value	*<0.001*	*0.007*	*0.044*	–					

* One overly influential data point was removed due to a high Cook's Distance (0.5) and a high studentized residual (> 4)

throughout the rest of the year. This supports the fiscal year hypothesis mentioned in the previous section. This model has an R^2 of 0.521, indicating that the model explains 52.1% of the variation in monthly DLR costs for the B-2. The results of the diagnostic tests for normality and constant variance were satisfactory.

One data point was removed from this model. It had both a high Cook's Distance value, shown in Figure 14, and a very high studentized residual, shown in Figure 15. The data point represents an observation taking place in the month of November. DLR costs for this particular month are very high, especially relative to the other November data points. Therefore, the model significantly underestimates the DLR costs for this observation.

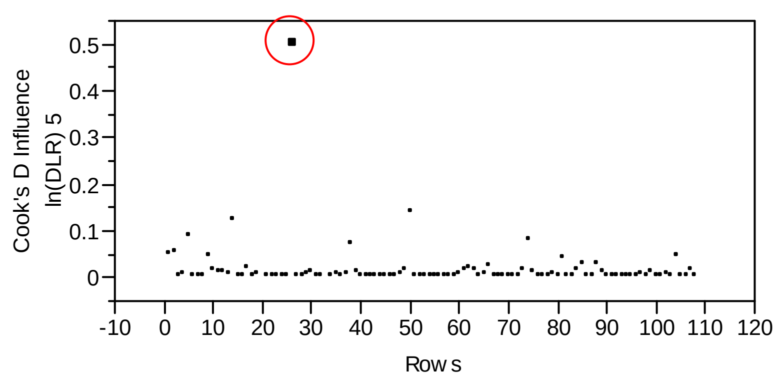

Figure 14: Overlay Plot of Cooks Distance for Model 1

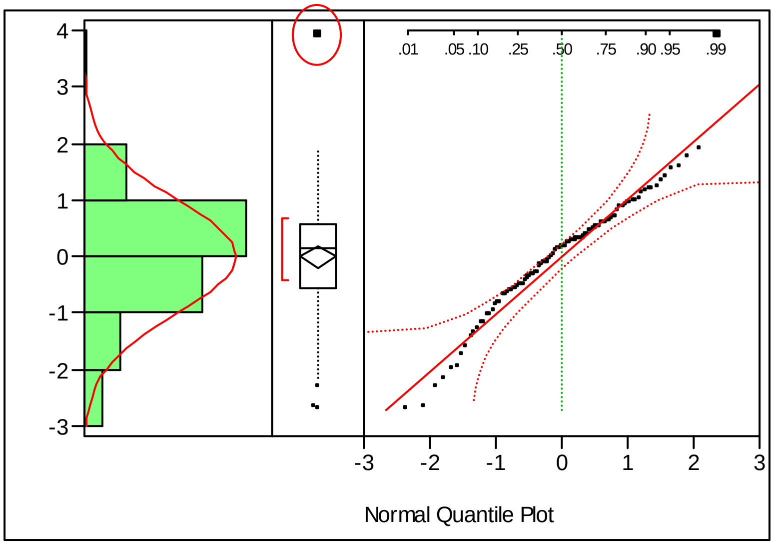

Figure 15: Histogram and Normal Quantile Plot of Studentized Residuals for Model 1

The second model relates the log of MDS-level quarterly DLR costs to the log of flying hours and average age.

*ln(Quarterly DLR) = 10.998 + 0.725 * ln(Flying Hours) + 0.005 * Age*

This model indicates that a one percent increase in flying hours in a given quarter results in a 0.689% increase in net DLR costs. Additionally, increasing the average age of the B-2 fleet by one month, results in a 0.6% increase in DLR costs. This is equivalent to a 1.8% increase per quarter. This model explains 59.2% of the variation in quarterly DLR costs for the B-2. All diagnostics yielded satisfactory results.

Of the two models described in this section, we found that the base-level quarterly model yielded a slightly higher R^2 than the monthly model. This model suggests that age, in addition to flying hours, is correlated with B-2 net DLR costs.

B-2 CONS Model

This section discusses the predictive models for B-2 CONS costs. For both the monthly and quarterly CONS models, none of the potential predictor variables considered in this research, to include flying hours, is statistically significant in forecasting B-2 CONS costs.

We include the results of the models of the log of CONS costs regressed on the log of flying hours. These models reveal that flying hours is not a statistically significant predictor of B-2 CONS costs. We hypothesize that these unusual results may be due to the small sample size of B-2 aircraft. While the B-1B and B-52H each provide a sample size of over 60, there are only 20 B-2 aircraft. Additionally, the B-2 is the youngest of the three bombers; therefore, perhaps consumable expenditure patterns have not yet stabilized for this aircraft.

Table 10: B-2 CONS Cost Models

Dependent Variable: Log of B-2 Consumables Costs					
Model		Intercept	Log of Flying Hours	R^2 / Adj R^2	Number of Obs
Base/MDS Monthly	Coefficient	12.831	0.046	0.001	107
	Standard Error	*0.798*	*0.130*	-0.008	
	p-Value	*<0.001*	*0.722*		
Base/MDS Quarterly	Coefficient	14.094	0.025	0.001	36
	Standard Error	*1.281*	*0.177*	-0.029	
	p-Value	*<0.001*	*0.997*		

B-52H DLR Model

This section discusses the four models we developed to forecast net DLR costs for the B-52H. The four models are similar to those developed for the B-1B; the first two models evaluate base and MDS-level data at the monthly level, and the remaining two models evaluate base and MDS-level data at the quarterly level. Table 11 summarizes the four B-52H DLR models.

The first model relates the log of monthly base-level DLR costs for the B-52H to the log of flying hours, average base-level ASD, and the month of November.

$$ln(Monthly\ DLR) =$$

$$= 12.101 + 0.760 * ln(Flying\ Hours) - 0.268 * ASD - 0.976 * November$$

The model indicates that a one percent increase in the number of flying hours in a month will increase DLR costs by 0.760%. Additionally, increasing the ASD for a fleet of B-52Hs by one hour will reduce DLR costs by 26.8%. The model also indicates a seasonality trend associated with the fiscal year. Costs incurred in the month of November are significantly less than costs incurred during other months of the year. The R^2 for this model is 0.702. Therefore, this model explains 70.2% of the variance in monthly, base-level DLR costs for the B-52H. Additionally, this model failed both the Shapiro Wilk and Breusch Pagan tests. However, the deviation from normality was verified to be robust. The deviation from constant variance was visually determined to be robust.

Table 11: B-52H DLR Cost Models

Model		Intercept	Log of Flying Hours	Log of Sorties	ASD	November	Lag 1 of ln(DLR)	Lag 12 of ln(DLR)	R^2 Adj R^2	Number of Obs	Durbin Watson	Shapiro Wilk	Breusch Pagan
Base Monthly	Coefficient	12.101	0.76	--	-0.268	-0.976	--	--	0.702 0.698	214	0.276 0.071	0.010*	<0.001**
	Standard Error	*0.264*	*0.04*	–	*0.027*	*0.095*	–	–					
	p-Value	*<0.001*	*<0.001*	–	*<0.001*	*<0.001*	–	–					
MDS Monthly	Coefficient	10.465	0.28	--	--	-0.867	--	0.235	0.603 0.59	96	0.560	<0.001*	0.228
	Standard Error	*1.756*	*0.111*	–	–	*0.146*	–	*0.102*					
	p-Value	*<0.001*	*0.013*	–	–	*<0.001*	–	*0.023*					
Base Quarterly***	Coefficient	7.478	--	0.645	-0.115	--	0.367	--	0.901 0.896	67	0.171 0.555	0.009**	0.063
	Standard Error	*0.984*	–	*0.085*	*0.026*	–	*0.082*	–					
	p-Value	*<0.001*	–	*<0.001*	*<0.001*	–	*<0.001*	–					
MDS Quarterly	Coefficient	3.263	--	0.712	--	--	0.543	--	0.584 0.558	35	0.385	0.131	0.071
	Standard Error	*2.164*	–	*0.181*	–	–	*0.115*	–					
	p-Value	*<0.001*	–	*<0.001*	–	–	*<0.001*	–					

Dependent Variable: Log of B-52H Depot level Reparables Costs

* Verified to be a robust deviation from normality (Shapiro Wilk) or constant variance (Breusch Pagan)

** Visually determined to be a robust deviation from normality (Shapiro Wilk) or constant variance (Breusch Pagan)

*** Two overly influential data points were removed due to high Cook's Distance values and high-magnitude studentized residuals

The second model listed in Table 11 relates the log of monthly B-52H DLR costs aggregated at the MDS-level to the log of flying hours, the month of November, and the log of DLR costs incurred in the same month one year prior.

$$ln(Monthly\ DLR)_{(t)} =$$

$$= 10.465 + 0.280 * ln(Flying\ Hours) - 0.867 * November + 0.235 * ln(DLR)_{(t-12)}$$

A one percent increase in flying hours will increase monthly MDS-level DLR costs by 0.280%, while costs incurred in the month of November are significantly less than DLR costs during the rest of the year. Additionally, a one percent increase in costs incurred during the same month, one year prior, leads to a 0.235% increase in DLR costs for the current month. This model explains 60.3% of the variance in monthly MDS-level DLR costs for the B-52H. This model failed the Shapiro Wilk test; however, it was verified to be a robust deviation from normality.

The third model aggregates B-52H DLR costs by quarter at the base-level. This model relates the log of DLR costs to the log of sorties, base-level ASD, and the log of DLR costs from the previous quarter.

$$ln(Quarterly\ DLR)_{(t)} = 7.478 + 0.645 * ln(Sorties) - 0.115 * ASD + 0.367 * ln(DLR)_{(t-1)}$$

A one percent increase in the number of sorties flown in a quarter will increase DLR costs by 0.645%. Increasing the base-level ASD by one hour will reduce costs by 11.5%. Additionally, a one percent increase in net DLR costs from the previous quarter will lead to an increase in this month's DLR costs by 0.367%. This model has an R^2 of 0.901, indicating that it explains 90.1% of the variation in DLR costs. Though this model failed the Shapiro Wilk test, the deviation was visually determined to be a robust deviation from normality.

Two overly influential points were omitted from the final model representing quarterly base-level costs, listed in Table 11. Figure 16 highlights these two points in an overlay plot of the Cook's Distance values. Additionally, Figure 17 shows the studentized residuals of these two points.

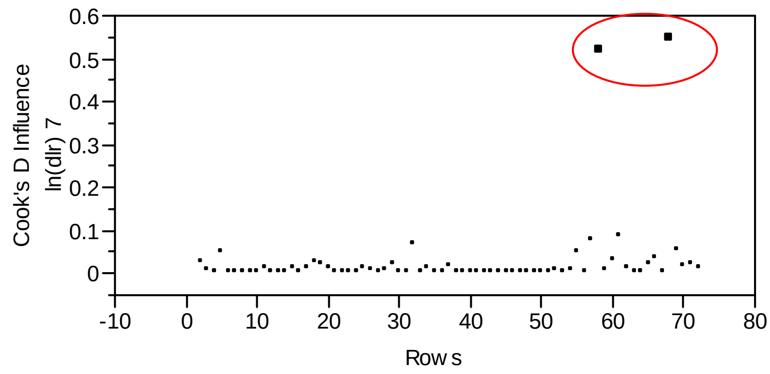

Figure 16: Overlay Plot of Cook's Distance Values for Model 3

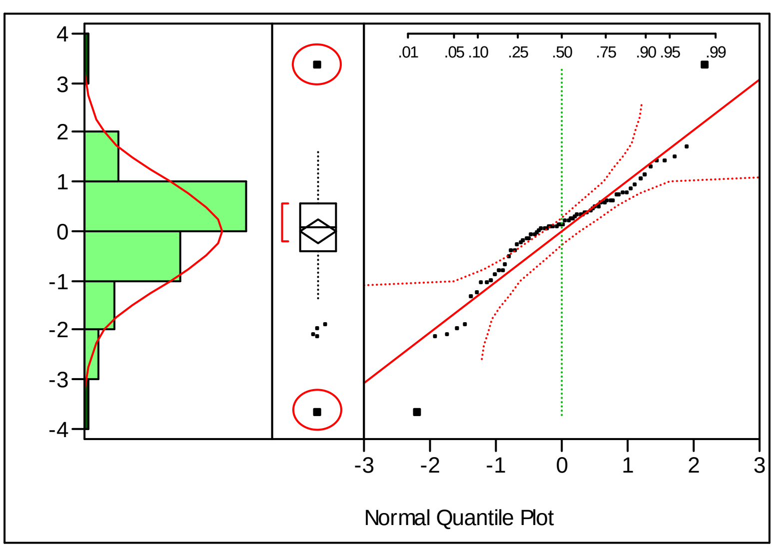

Figure 17: Studentized Residuals for Model 3

One of these data points represents a quarter where DLR costs were unusually high. Therefore, the model significantly underestimated costs for this period, causing the point to have a very high studentized residual. The other omitted data point represents a quarter where costs were unusually low. In this case, the model significantly overestimated DLR costs for this period. Excluding these two data points resulted in a decrease in the estimated coefficient for the log of sorties, from 0.766 to 0.645. Additionally, the coefficient on ASD changed from -0.147 to -0.115. Finally, the coefficient on log of the previous quarter's DLR costs increased from 0.261 to 0.367. Figure 18 and Figure 19 depict the Cook's Distance values and the studentized residuals after these points are excluded.

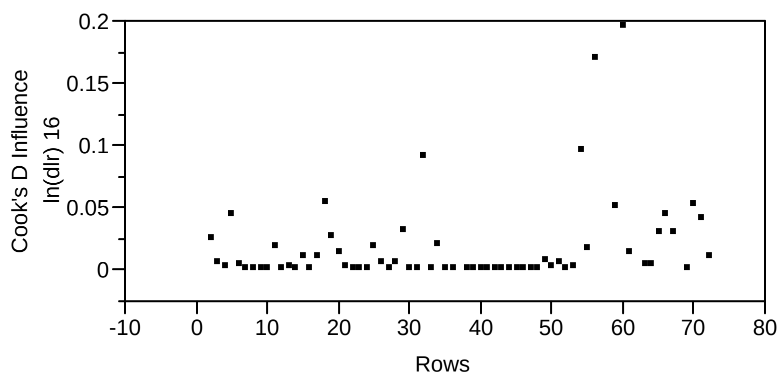

Figure 18: Overlay Plot of Cook's Distance for Model 3 (Excluding Anomalies)

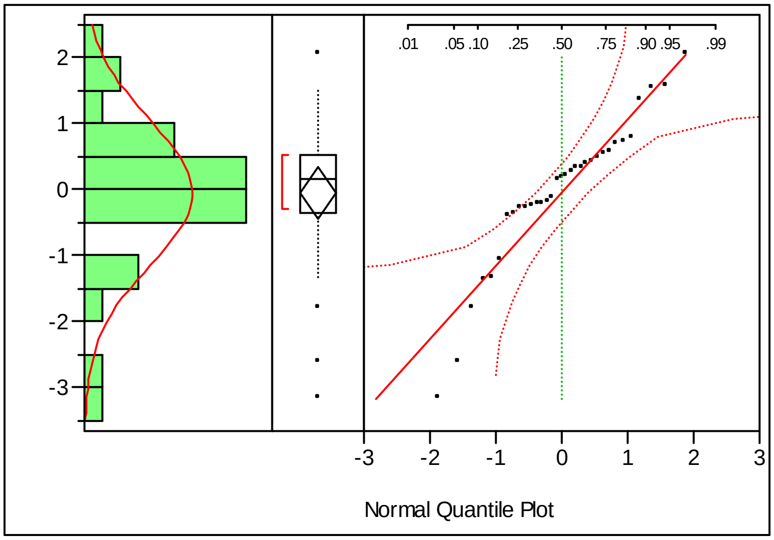

Figure 19: Studentized Residuals for Model 3 (Excluding Anomalies)

The final model in Table 11 relates the log of quarterly MDS-level DLR costs to the log of sorties and the log of the previous quarter's DLR costs.

$$ln(Quarterly\ DLR)_{(t)} = 7.478 + 0.712 * ln(Sorties) + 0.543 * ln(DLR)_{(t-1)}$$

This model indicates that a one percent increase in the number of sorties flown by the B-52H fleet will lead to a 0.712% increase in DLR costs. Additionally, if the previous quarter's costs increase by one percent, we can anticipate a 0.543% increase in DLR costs for the current quarter. This model explains 58.4% of the variance in quarterly MDS-level DLR costs for the B-52H. All diagnostic tests were satisfactory.

The models developed to forecast B-52H DLR costs reveal that the base-level quarterly level of aggregation yields the most predictive model. Additionally, this model suggests that sorties, ASD, and lagged costs may be more useful in forecasting DLR costs than flying hours.

B-52H CONS Model

This section summarizes the final set of models developed in this study, those representing quarterly MDS-level CONS costs for the B-52H. The first two models in this section represent monthly base-level CONS costs; one model represents Barksdale AFB and the other represents Minot AFB. The final three models represent monthly costs aggregated at the MDS-level, quarterly costs aggregated at the base-level, and quarterly costs aggregated at the MDS-level.

The first model relates the log of monthly CONS costs for Barksdale AFB to the log of average base-level total operating hours (TOH), average base-level ASD, and the month of November.

$$ln(Monthly\ DLR) = 0.594 + 1.431 * ln(TOH) – 0.089 * ASD – 0.379 * November$$

The model indicates that a one percent increase in the average total operating hours for a fleet of aircraft will lead to a 1.431% increase in CONS costs. Additionally, increasing the average ASD length by one hour will lead to an 8.9% decrease in CONS costs for Barksdale AFB. Finally, costs in November are, once again, significantly less than costs incurred in other months. This model explains 25.1% of the variance in the dependent variable, monthly CONS costs for Barksdale AFB. All of the diagnostic tests were satisfactory for this model.

The next model relates the log of monthly CONS costs for Minot AFB to the base-level mission capable rate (MCR), the month of November, and a year effect.

$$ln(Monthly\ DLR) = 0.594 + 1.431 * ln(TOH) – 0.089 * ASD – 0.379 * November$$

The year effect in this model is a dummy variable representing CONS costs incurred from September 2002 through April 2005. During these months, CONS costs at Minot

Table 12: B-52H CONS Cost Models

Dependent Variable: Log of B-52H Consumables Costs																
Model		Intercept	Log of Flying Hours	Log of TOH	ASD	MCR	Age	Cannibalization Rate	November	Year Effect	Lag 1 of ln(CONS)	R^2 Adj R^2	Number of Obs	Durbin Watson	Shapiro Wilk	Breusch Pagan
Base Monthly (Barksdale)	Coefficient	0.594	--	1.431	-0.089	--	--	--	-0.379	--	--	0.251 0.229	107	0.149	0.100	0.822
	Standard Error	*5.266*	–	*0.541*	*0.033*	–	–	–	*0.109*	–	–					
	p-Value	*0.911*	–	*0.009*	*0.008*	–	–	–	*0.001*	–	–					
Base Monthly (Minot)	Coefficient	14.623	--	--	--	-2.121	--	--	-0.631	0.360	--	0.336 0.316	108	0.154	0.063	0.016*
	Standard Error	*0.844*	–	–	–	*0.968*	–	–	*0.131*	*0.082*	–					
	p-Value	*<0.001*	–	–	–	*0.031*	–	–	*<0.001*	*<0.001*	–					
MDS Monthly	Coefficient	9.544	--	--	--	--	0.002	0.008	-0.502	--	0.238	0.411 0.388	107	0.539	0.068	0.214
	Standard Error	*1.120*	–	–	–	–	*0.001*	*0.002*	*0.097*	–	*0.086*					
	p-Value	*<0.001*	–	–	–	–	*0.018*	*0.001*	*<0.001*	–	*0.006*					
Base Quarterly	Coefficient	8.760	0.736	--	--	--	--	--	--	--	--	0.662 0.657	72	0.261 0.559	0.403	0.551
	Standard Error	*0.488*	*0.063*	–	–	–	–	–	–	–	–					
	p-Value	*<0.001*	*<0.001*	–	–	–	–	–	–	–	–					
MDS Quarterly	Coefficient	9.700	0.660	--	--	--	--	--	--	--	--	0.284 0.263	36	0.111	0.117	0.220
	Standard Error	*1.542*	*0.180*	–	–	–	–	–	–	–	–					
	p-Value	*<0.001*	*0.001*	–	–	–	–	–	–	–	–					

* Visually determined to be a robust deviation from normality constant variance

AFB were significantly larger than CONS costs in the rest of the dataset. The points represented by the year effect are circled in the overlay plot in Figure 20.

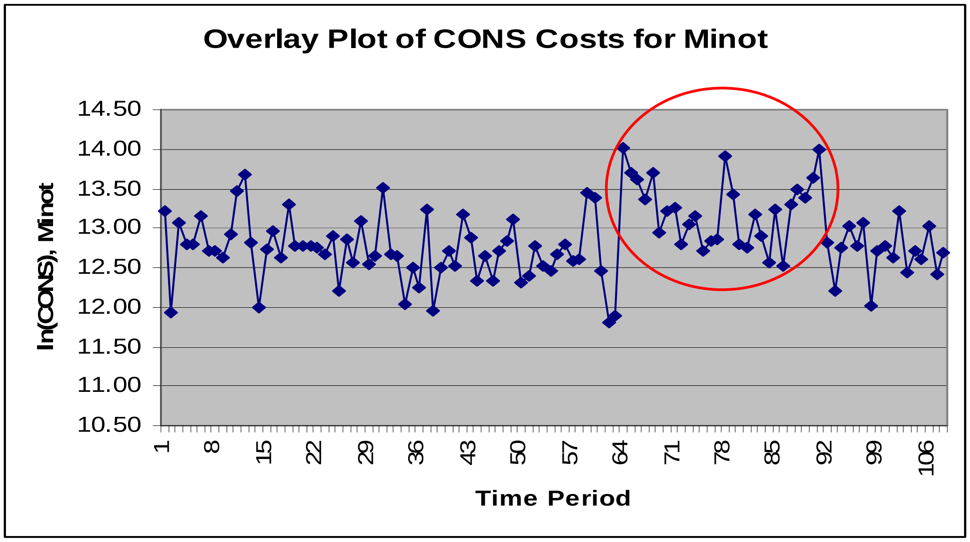

Figure 20: Overlay Plot of CONS Costs for Minot

This model indicates that a one percent increase in MCR, or the percentage of aircraft at Minot AFB available to fly a mission, will lead to a 2.120% decrease in CONS costs. As a lower MCR indicates a higher percentage of aircraft that are incapable of flying a mission, this finding makes logical sense. Additionally, costs incurred in the month of November are statistically less than CONS costs incurred during the other months of the year. Finally, CONS costs incurred from September 2002 through April 2005 are, on average, 36.0% larger than costs incurred before September 2002 or after April 2005. The Durbin Watson and Shapiro Wilk statistics were satisfactory for this model. The model failed the Breusch Pagan test; however, the deviation was visually determined to be a robust deviation from constant variance.

The third model listed in Table 12 relates the log of monthly MDS-level costs to the average age of the B-52H, the cannibalization rate, the month of November, and the log of the previous month's CONS costs.

$$ln(Monthly\ CONS)_{(t)} = 9.544 + 0.002 * Age +$$

$$0.008 * Cannibalization\ Rate - 0.502 * November + 0.238 * ln(DLR)_{(t-1)}$$

This model indicates that CONS costs will increase 0.2% as the average age of the B-52H increases by one month. Likewise, costs will increase 2.4% for every year the B-52H ages. This model was also the only model to find cannibalization rate to be a significant costs driver, though the impact is small. When cannibalization rate, or the number of cannibalizations per 100 sorties, increases by one percent costs will increase by 0.008%. Costs incurred in the month of November are, once again, statistically less than costs incurred during the other months of the year. Finally, a one percent increase in the previous month's CONS costs will lead to a 0.240% increase in the current month's costs. This model explains 41.1% of the variation in the dependent variable. Additionally, all diagnostic tests were satisfactory.

The fourth model listed in Table 12 relates the log of quarterly base-level costs to the log of flying hours. The model indicates that a one percent increase in flying hours will lead to a 0.736% increase in quarterly CONS costs. This model explains 66.2% of the variation in quarterly base-level CONS costs for the B-52H. All diagnostics are satisfactory.

The final model relates the log of quarterly MDS-level costs to the log of flying hours. This model suggests that a one percent increase in flying hours will increase costs by 0.660%. 28.4% of the variation in quarterly MDS-level CONS costs is explained by this model and all diagnostic tests are satisfactory.

Once again, we find that the most predictive level of aggregation is the one based on a base-level quarterly model. This model suggests that forecasting quarterly costs at the base-level using flying hours as an independent variable is the most useful for B-52H CONS costs.

Summary

This chapter both tested the current proportional cost forecasting model and explored the predictability of other variables associated with flying costs. To test the current proportional CPFH methodology, we performed regressions of quarterly MDS-level costs on flying hours. The results of those regressions were presented in this chapter and revealed that using a proportional model based on CPFH factors may be an inappropriate method of forecasting DLR and CONS costs for the three Air Force bombers. We then evaluated the effects of 12 potential predictor variables, chosen based on previous research, on DLR and CONS costs for each of the bombers. The final models were presented in this chapter. Chapter V discusses our conclusions based on these results.

V. Conclusions

Overview

This chapter uses the results from Chapter IV to answer the research questions defined in Chapter I. After evaluating these questions, this research will conclude by discussing the strengths and limitations of our study.

Findings

We defined two research questions in the first chapter. After conducting a literature review and defining the methodology, we gathered data on select variables and developed predictive cost models, discussed in Chapter IV. This section will answer both of the research questions based on the results from the previous chapter.

Q1: Does the current CPFH methodology, which assumes a proportional relationship, capture the true relationship between flying hours and costs?

With the exception of B-2 depot level repairable (DLR) costs, the data do not support the hypothesis that DLR or consumables (CONS) costs are proportionally related to flying hours. We define the proportional cost per flying hour (CPFH) relationship as one which 1) has a zero intercept, indicating that when zero hours are flown, costs will be zero, and 2) yields a linear relationship between flying hours and cost such that a one percent increase in flying hours leads to a one percent increase in costs. As shown in Chapter IV, simple regressions of cost on flying hours revealed that each type of cost has a large, non-zero intercept. Additionally, the slope of each regression was smaller in magnitude than the current FY2008 CPFH factor used by the Air Force. Figure 21 shows the variation between the current proportional forecasting methodology and the actual relationship between costs and flying hours based on actual data for the B-1B and B-52H, as shown in Chapter IV.

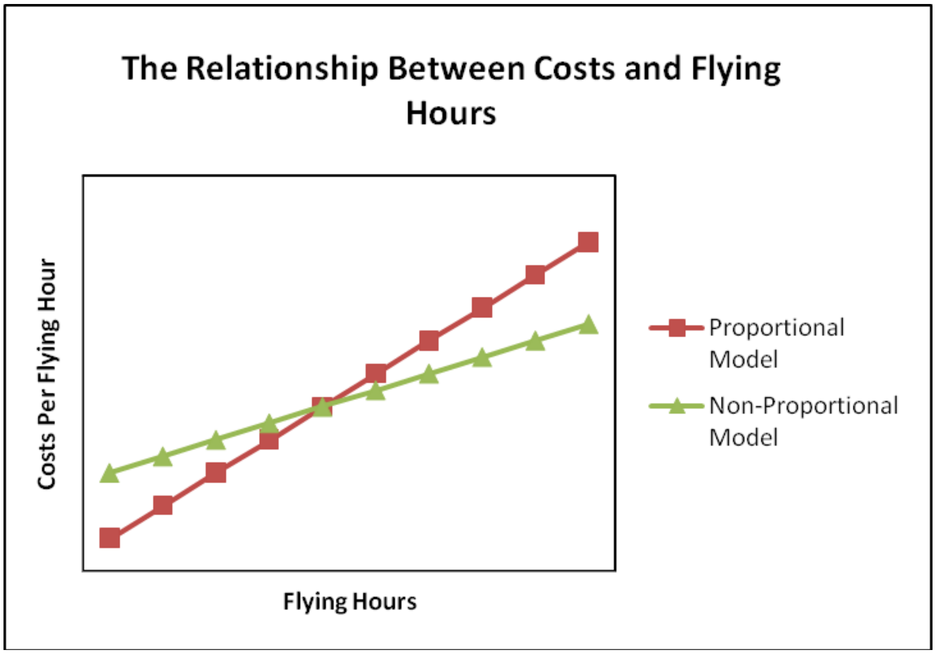

Figure 21: The Relationship between Costs and Flying Hours

The models we built in Chapter IV were based on quarterly data. While cost estimators can apply these models to monthly forecasts, it must be done with caution. Monthly level data are much more variable than the aggregated quarterly level data. Therefore, actual costs may be much different than estimated costs in a specific month, though the costs should balance out over time.

While we did not validate the usefulness of regression analysis over the current proportional model, our results suggest that Air Force analysts consider the use of simple regressions to forecast flying costs for the B-1B and B-52H, as these models indicate a nonzero intercept due to routine maintenance costs. Additionally, as shown in Figure 21, the current proportional model underestimates the actual costs at low levels of flying hours and overestimates costs at high levels of flying hours. While B-2 DLR costs are forecasted relatively well by a proportional model, given the historical range of flying hours, flying hours does not appear to be useful in forecasting B-2 CONS costs.

Q2: Are factors other than flying hours useful in estimating flying costs?

Our results indicate that building models based on flying hours alone is insufficient. In addition to evaluating the usefulness of the current forecasting methodology, we also evaluated several other factors as potential predictors of cost. Accounting for lagged costs, fiscal trends, and sorties, among other variables, results in much more predictive cost forecasting models.

The frequency of occurrence of each of the independent variables is shown in Table 13. Each of the columns represents of group of models, while each of the independent variables is listed down the left-hand side in order of most frequent to least frequent.

Table 13: Frequency of Occurrence of Each Independent Variable (Expressed as a Fraction of the Total Models in Each Category)

Independent Variable	B-1B DLR	B-1B CONS	B-2 DLR	B-2 CONS	B-52H DLR	B-52H CONS	Total	Total DLR	Total CONS
Lagged Costs	2/5	5/5	0/2	0/0	3/4	1/5	**11/21**	5/11	**6/10**
Flying Hours	3/5	0/5	2/2	0/0	2/4	2/5	9/21	**7/11**	2/10
Fiscal Trend	0/5	1/5	1/2	0/0	2/4	3/5	7/21	3/11	4/10
Sorties	1/5	2/5	0/2	0/0	2/4	0/5	5/21	3/11	2/10
ASD	1/5	0/5	0/2	0/0	2/4	1/5	4/21	3/11	1/10
Age	1/5	0/5	2/2	0/0	0/4	1/5	4/21	3/11	1/10
Utilization Rate	1/5	0/5	0/2	0/0	0/4	0/5	1/21	1/11	0/10
MCR	0/5	0/5	0/2	0/0	0/4	1/5	1/21	0/11	1/10
Cann Rate	0/5	0/5	0/2	0/0	0/4	1/5	1/21	0/11	1/10
TOH	0/5	0/5	0/2	0/0	0/4	1/5	1/21	0/11	1/10
Crude Oil Prices	0/5	0/5	0/2	0/0	0/4	0/5	0/21	0/11	0/10
Temperature	0/5	0/5	0/2	0/0	0/4	0/5	0/21	0/11	0/10

Our findings indicate that it may be useful to model DLR and CONS costs differently. In general, we found that flying hours was the most predictive variable in forecasting net DLR costs, while accounting for lagged costs was the most useful in forecasting net CONS costs. As discussed, the Air Force currently uses estimating models based solely on flying hours to forecast both of these types of costs. While flying hours was significant in seven of the eleven DLR cost models, it was only significant in two of the ten CONS models. Our findings indicate that accounting for lagged costs is more useful in estimating CONS costs than flying hours.

In addition to these two independent variables, fiscal trend (represented as the month of November in this study), sorties, ASD, and age were also found to be relatively significant, as supported by previous research. Additionally, we found that crude oil prices and temperature were not very predictive variables.

Interestingly, our models revealed costs in the month of November to be significantly less than costs throughout the rest of the year. As discussed in Chapter II, both Armstrong (2006) and Bryant (2005) found costs to be significantly larger toward the end of the fiscal year. Our findings, in conjunction with Armstrong (2006) and Bryant (2007) indicate a significant fiscal trend in the patterns of flying costs.

As previously discussed, we evaluated cannibalization rate as a potential predictor variable of flying costs. Cannibalization rate was not found to be statistically significant in any of the DLR models, while it was statistically significant in just one of the CONS models. These findings do not support our hypothesis that flying costs will increase as the number of cannibalizations increases.

Several previous studies highlighted aircraft age as a significant predictor of operating costs. Surprisingly, only four of the 21 models included age as a statistically significant independent variable. Further, only two of those four models were for the older B-1B and B-52H aircraft. These findings do not support the argument that aircraft operating costs experience a "bathtub effect," as discussed in Chapter II.

In addition to including variables other than flying hours, our research indicates that building models based on quarterly data for each base is the most predictive level of aggregation, given the four levels we evaluated. For each of the six types of costs, we found this level of aggregation to be the most predictive.

Finally, we found that log-linear regression analysis resulted in an improved model fit of bomber flying cost data over the proportional CPFH model.

Strengths and Limitations

As defined in Chapter I, we sought to explore the potential of developing more predictive flying cost forecasting models for the Air Force's bomber aircraft, as compared to the current methodology of forecasting budgets employed by Air Force budgeters. While we feel we were successful in identifying significant predictor variables that are not considered in the current budgeting process, these models are merely a starting point.

Evaluating base-level monthly data provided a large dataset from which to build the forecasting models of this study. However, it also resulted in data subject to a significant amount of noise. Therefore, we decided to evaluate the cost data at MDS and quarterly levels of aggregation, as well. This resulted in four forecasting models for each type of cost, allowing us to compare the results across different levels of aggregation. This also provides budgeters with flexibility in using our results to forecast costs.

One main reason for the large amount of variance in the monthly data may be due to the variation in monthly expenditures and credits experienced by different bases. This research found that, in some instances, the credits were larger than the costs for a given month, resulting in a negative net cost. This problem was significant for both DLR and CONS costs associated with the B-1B. To control for these negative costs, we added a constant term to the dependent variable to make all costs positive, as discussed in Chapter III. This allowed us to perform a log transformation without excluding data points and losing critical information. Additionally, evaluating quarterly and MDS-level data eliminated the problems associated with negative costs.

Finally, while the models we reported in Chapter IV satisfy the assumptions of OLS regression, we found that different models representing the same type of cost did not include the same statistically significant independent variables. So while our models are statistically sound, this presents a level of difficulty in comparing models across levels of aggregation and cost types. A possible explanation for the inconsistency of independent variables across models is the small datasets from which the models were built. It may be the case that there was so much variability at low levels of aggregation that variables found to be significant at higher levels of aggregation were not significant at the lower levels. Similarly, it may have been that there was so little variation associated with certain independent variables at higher levels of aggregation that variables found to be significant at lower levels were not significant at higher levels.

Perhaps models based on larger datasets would have resulted in more consistency across different levels of aggregation.

Follow-On Suggestions

As mentioned, the dependent variables evaluated in this research represent net costs; that is, charges less credits. We suggest further evaluation of the expenditure and credit systems. Evaluating expenditures and credits individually may reveal that building separate forecasts for each is more appropriate than forecasting the net costs, as is done in current practice.

We also recommend focusing specifically on the B-2 flying costs. We determined that the predictor variables evaluated in this study were insignificant in forecasting B-2 CONS costs, perhaps due to the small sample size or the young age of the aircraft. More research should be conducted to determine what drives these costs.

Summary

The results of this analysis indicate that regression analysis may outperform the current proportional model in forecasting flying costs when flying hours are above or below average. For example, if an unusually low level of flying hours are flown in a given time period, the proportional model underestimates flying costs. Our models capture the fixed costs associated with aircraft maintenance that occur regardless of flying hours. Further, if hours are higher than average, the current proportional model overestimates costs; a regression model captures the lower costs reflected in the data evaluated. Our research also reveals that other operational factors such as sorties, seasonality trends and lagged costs, may be useful predictors in forecasting flying costs.

Appendix A: Description of Aircraft

B-1B Lancer

The B-1B entered active service in 1985. The key features of this airframe, developed in the 1960s, are its variable-geometry wings and its ability to carry the largest payload of guided and unguided weapons in the Air Force inventory (Fact Sheet: B-1B, 2007). Additionally, the B-1B is the fastest of the three Air Force bombers, with an ability to fly Mach 1.2 (900 miles per hour). Each of these features has contributed to significant structural problems experienced by the aircraft.

B-2 Spirit

The unique frame of the B-2 was first publicly displayed in 1988. Its most distinguishing characteristic is its flying wing design which contributes to its "stealth" capability. The structural design of the B-2 is a relatively new aviation technology, and the aircraft is the only bomber which uses this technology. While the rest of the world develops anti-aircraft weapons, the use of stealth aircraft will surely increase (Fact Sheet: B-2, 2007). Additionally, as the aircraft continues to age, the potential for unknown and previously unseen problems increases.

B-52H Stratofortress

The B-52H was first delivered in 1962 and is expected to fly for another 35 years. The massive aircraft is known for its 185-foot wingspan and its ability to carry the widest array of weapons in the Air Force inventory (Fact sheet: B-52H, 2007). These two features have led to significant structural problems. Additionally, the aging aircraft are suffering from fuel tank erosion (Hebert, 2003:7-8). As the B-52H continues to creep closer to its 80-year expected retirement age, the Air Force will surely face new and unanticipated costs.

Appendix B: Samples from Each Database

Table 14: Sample from AFTOC Database

Data_Type	Base	MD_CAIG	Demand_FY_Year	Demand_FY_Month	Net_Cost_CurrentYear
GSD	BARKSDALE AFB (LA)	B-52H	2002	12	$2,227,017.73
SD	BARKSDALE AFB (LA)	B-52H	2002	12	$3,124,016.30
GSD	DYESS AFB (TX)	B-1B	1998	01	$77,969.38
MSD	DYESS AFB (TX)	B-1B	1998	01	$15,029,106.86
GSD	DYESS AFB (TX)	B-1B	1998	02	$87,907.03
MSD	DYESS AFB (TX)	B-1B	1998	02	$12,114,202.00

Table 15: Sample from REMIS Database

MDS	FY	Fscl_Month	Command	Base	Tail_Number	FH	Landings	Sorties
B-52H	FY1998	1	AFRC	BARKSDALE AFB (LA)	61000032	18.80	13	4
B-52H	FY1998	1	ACC	BARKSDALE AFB (LA)	61000038	13.70	15	2
B-52H	FY1998	1	ACC	BARKSDALE AFB (LA)	61000039	23.30	14	5
B-52H	FY1998	2	ACC	BARKSDALE AFB (LA)	60000001	15.70	4	2
B-52H	FY1998	2	ACC	BARKSDALE AFB (LA)	60000002	30.60	17	4
B-52H	FY1998	2	ACC	BARKSDALE AFB (LA)	60000003	10.60	1	1

Table 16: Sample from MERLIN Database

WEAPON	B-1B							
MAJCOM	ACC				ACC			
UNIT	0007BHVWG				0028BHVWG			
METRIC	NMCM	NMCS	NMCB	MCR	NMCM	NMCS	NMCB	MCR
Oct-97	14.4	14.6	12.9	58.1	17.7	22	2.8	57.5
Nov-97	17.2	11	12	59.8	14	24.8	6.2	55
Dec-97	21.6	10.1	10.2	58.1	14.7	20.3	6.7	58.3
Jan-98	17	9.4	8.9	64.7	12.3	25.8	10.1	51.8
Feb-98	17.7	9.2	6.2	66.9	15.7	19.5	5.2	59.6
Mar-98	17.1	15.7	11.2	56	17.6	17.6	3	61.8

Table 17: Sample from AFCCC Database

Locations	Year	Month	Mean Temp C	Mean Temp F	Mean Dew C	Mean Dew F
WHITEMAN AFB	2006	10	11.9	53.4	6.2	43.2
WHITEMAN AFB	2006	11	7.4	45.3	2.5	36.5
WHITEMAN AFB	2006	12	2.7	36.9	-1.5	29.3
MCCONNELL AFB	1998	1	1.5	34.7	-1.0	30.1
MCCONNELL AFB	1998	2	5.3	41.6	1.0	33.7
MCCONNELL AFB	1998	3	4.7	40.4	0.6	33.0

Appendix C: Charges versus Credits

Figures 22 and 23 show the relationship between total charges and total credits for all six types of costs modeled. While some charts reveal similarities between charges and credits over time, it also appears that there may be differences between the two.

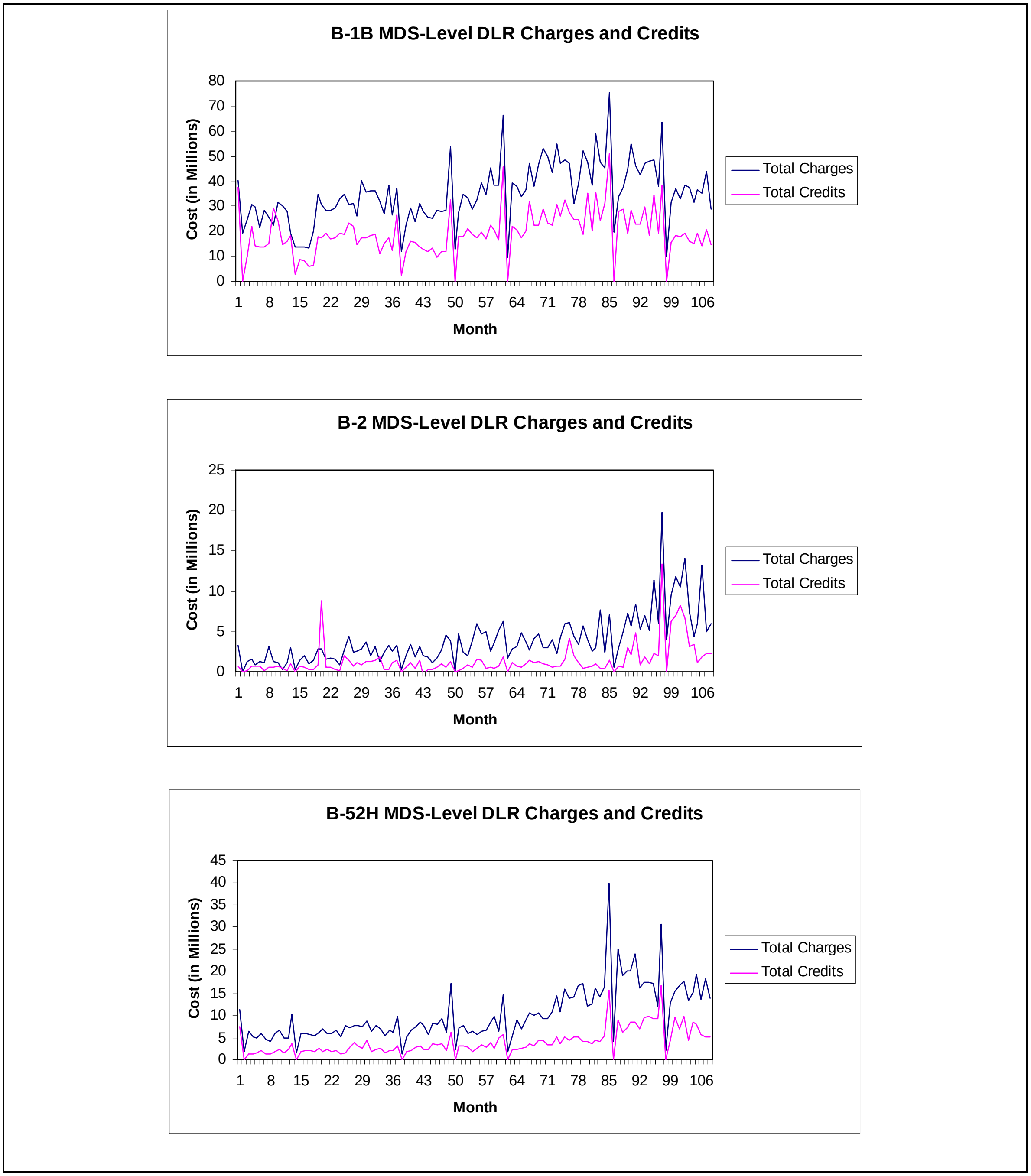

Figure 22: DLR Charges and Credits Over Time

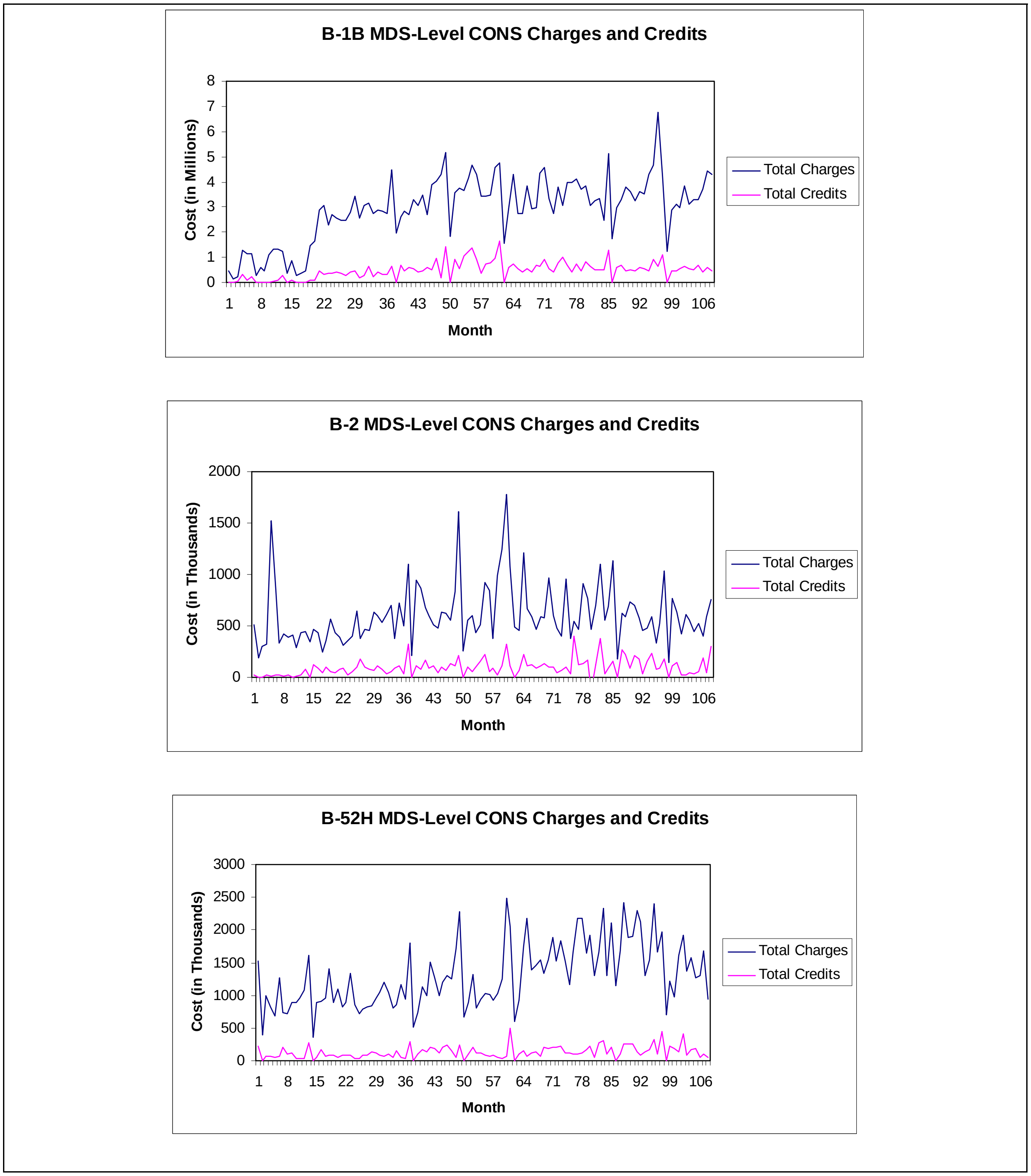

Figure 23: CONS Charges and Credits Over Time

Bibliography

Armstrong, Patrick D. *Developing an Aggregate Marginal Cost Per Flying Hour Model for Air Force's F-15 Fighter Aircraft*. MS Thesis, AFIT/GCA/ENV/06M-01. School of Engineering and Management, Air Force Institute of Technology (AU), Wright-Patterson AFB OH, March 2006 (ADA423137).

Bryant, Michael T. *Forecasting the KC-135 Cost Per Flying Hour: A Panel Data Analysis*. MS Thesis, AFIT/GCA/ENV/07-M2. School of Engineering and Management, Air Force Institute of Technology (AU), Wright-Patterson AFB OH, March 2007.

Congressional Budget Office. "The Long-Term Implications of Current Defense Plans: Detailed Update for Fiscal Year 2007." Washington, D.C., April 2007.

Dawson, Kevin P. and Jeremy A. Howe. "Analyzing Air Force Flying Hour Costs," *Air Force Journal of Logistics*, Summer 2006.

Defense and the National Interest. "Rising Cost of O&M Per Flying Hour." (January 2006) http://www.defense-and-society.org/charts_data/rising_cost_fh_onm.htm.

"Fact Sheet: B-1B Lancer," *Air Force Link*. n. pag. 13 August 2007 http://www.af.mil/factsheets/factsheet_print.asp?fsID=81&page=1.

"Fact Sheet: B-2 Spirit," *Air Force Link*. n. pag. 13 August 2007 http://www.af.mil/factsheets/factsheet_print.asp?fsID=82&page=1.

"Fact Sheet: B-52 Stratofortress" *Air Force Link*. n. pag. 13 August 2007 http://www.af.mil/factsheets/factsheet_print.asp?fsID=83&page=1.

Government Accountability Office. *Trends in Operations and Maintenance Costs and Supporting Services Contracting*. Washington DC: Government Printing Office. GAO-07-631. May 2007.

Government Accountability Office. *Services Need Strategies to Reduce Cannibalizations*. Washington DC: Government Printing Office. GAO-02-86. November 2001.

Government Accountability Office. *Observations on the Air Force Flying Hour Program*. Washington DC: Government Printing Office. GAO/NSIAD-99-165. July 1999.

Government Accountability Office. *Contingency Operations: Defense Costs and Funding Issues*. Washington DC: Government Printing Office. GAO/NSIAD-96-121BR. March 1996.

Hart, Karl and Terry Mitchell. "Aging Aircraft," *Military Aerospace Technology*, April 2003.

Hawkes, Eric M. *Predicting the Cost per Flying Hour for the F-16 using Programmatic and Operational Variables*. MS Thesis, AFIT/GOR/ENC/05-01. School of Engineering and Management, Air Force Institute of Technology (AU), Wright-Patterson AFB OH, June 2005 (ADA436138).

Hebert, Adam J. "When Aircraft Get Old," *Air Force Magazine*, 30-35 (January 2003).

Hildebrandt, Gregory G. *An Estimation of USAF Aircraft Operating and Support Cost Relations*. Santa Monica, CA: RAND, May 1990 (N-3062-ACQ).

Kammerer, Joseph T. "Notes from the: Deputy Assistant Secretary, Cost & Economics." *Air Force Comptroller*, 36-2:19 (April 2002).

Kammerer, Joseph T. "Notes from the: Deputy Assistant Secretary, Cost & Economics." *Air Force Comptroller* (July 2003).

Kiley, Gregory T. *The Effects of Aging on the Costs of Operating and Maintaining Military Equipment*. Washington, DC: Congressional Budget Office, August 2001.

Laubacher, Matthew E. *Analysis and Forecasting of Air Force Operating and Support Cost for Rotary Aircraft*. MS Thesis, AFIT/GCA/ENV/04M-05. School of Engineering and Management, Air Force Institute of Technology (AU), Wright-Patterson AFB OH, March 2004 (ADA423137).

McClave, James T., P. George Benson, and Terry Sincich. *Statistics for Business and Economics* (9th Edition). Upper Saddle River, New Jersey: Pearson Prentice Hall, 2005.

Pyles, Raymond A. *Aging Aircraft: USAF Workload and Material Consumption Life Cycle Patterns*. Santa Monica, CA: RAND, 2003 (MR-1641).

Rose, Pat A. Jr. "Cost Per Flying Hour Factors: A Background and Perspective of How They Are Developed and What They Do," *Air Force Comptroller*, 31-1:4-9 (July 1997).

Scaggs, John. "New System Streamlines Air Force Sustainment Funding," *Air Force Print News Today*. Air Force Materiel Command Public Affairs, WPAFB, OH. 25 September 2006 http://www.af.mail/news/story_print.asp?id=123027864.

Wallace, John M., Scott A. Hauser, and David A. Lee. *A Physics-Based Alternative to Cost-Per-Flying-Hour Models of Aircraft Consumption Costs*, Logistics Management Institute, August 2000 (ADA387273).

Wooldridge, Jeffrey M. *Introductory Econometrics: A Modern Approach* (3rd Edition). Mason OH: Thompson South-Western, 2006.

REPORT DOCUMENTATION PAGE	*Form Approved* *OMB No. 0704-0188*

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
27-03-2008	Master's Thesis	Oct 2006 - Mar 2008

4. TITLE AND SUBTITLE
Forecasting Flying Hour Costs of the B-1, B-2, and B-52 Bomber Aircraft

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Van Dyk, Stefanie L., First Lieutenant, USAF

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Air Force Institute of Technology
Graduate School of Engineering and Management (AFIT/EN)
2950 Hobson Way
WPAFB OH 45433-7765

8. PERFORMING ORGANIZATION REPORT NUMBER
AFIT/GCA/ENV/08-M02

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

10. SPONSOR/MONITOR'S ACRONYM(S)

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

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

13. SUPPLEMENTARY NOTES

14. ABSTRACT
This thesis both evaluates, and presents improvements to, the current method of forecasting flying costs of Air Force aircraft. It uses depot level repairable (DLR) and consumable (CONS) data for the Air Force's bomber platforms: B-1B, B-2, and B-52H. The current forecasting method assumes a proportional relationship between costs and flying hours such that 1) when no hours are flown costs are zero, and 2) a 1% increase in flying hours will increase costs by 1%. The findings of this research indicate that applying log-linear ordinary least squares regression techniques may be an improved fit of flying cost data over the current proportional model; the actual data indicate a non-zero intercept and a less than proportional relationship between costs and flying hours. This research also found that models including factors other than flying hours as independent variables, such as sorties, lagged costs, and fiscal trends, may be more useful than models based solely on flying hours. Finally, this research found that estimating quarterly costs at the base-level may yield more accurate estimates than estimating at the monthly level, or mission design series level.

15. SUBJECT TERMS
cost analysis, cost per flying hour, B-1B, B-2, B-52H, forecasting, OLS regression, O&M costs, bomber, DLR, consumables, cannibalization rates

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Lt Col Eric J. Unger, AFIT/ENV
U	U	U	UU	105	19b. TELEPHONE NUMBER *(Include area code)* 937-255-3636 x 7402 eric.unger@afit.edu