Modeling the Cumulative Impact of Change Orders

By

Karim Ashraf Sabry Iskandar

A thesis submitted in partial fulfillment of

the requirements for the degree of

Master of Science

(Civil and Environmental Engineering)

at the

UNIVERSITY OF WISCONSIN – MADISON 2016

Acknowledgements

Looking back two years from now, I still remember the first steps I undertook when I joined the graduate school in Madison in Summer 2014. Personally, those two years were intellectually challenging and enriching. In the name of all people I have met during the two years, some influential people shall be mentioned.

First of all, I would like to thank my advisor and mentor, Professor Awad S. Hanna, for giving me the opportunity to join the Construction Management Program at UW-Madison. During the past two years, Professor Hanna has provided guidance at key moments throughout this research. I learnt a lot from Professor Hanna's vast expertise in the construction management field. Professor Hanna also taught me how to be persistent, and how to approach things from a broader perspective before getting into details. I am and will be always grateful to Professor Hanna.

Also, I would like to express my gratitude to Professor Wafik Lotfallah for his great support throughout the past two years. I learnt a lot from Professor Wafik's rich knowledge in Mathematics. Furthermore, I would like to thank Professor Wei-Yin Loh, Professor Kevin Packard, Mr. John Gillett, and Xiaomao Li, for the knowledge I gained in statistics so far. In addition, I would like to thank the outstanding, Professor Jeffrey Aiken, for what I learnt from him during the amazing "Legal Aspects of Engineering" course.

Of course, I would also like to thank the Construction Industry Institute (CII) as well as all the mechanical and electrical contractors who participated in this study for the input and the effort that was exerted in collecting the data of this thesis report since the very debut.

Also, I would like to thank my dear friend and my fellow graduate student, Michael W. Ibrahim, for his support during the past two years through the good and bad times. I would also like to thank my fellow graduate students, Hala, Ramy, Dilibie, Tia, and Paul, for the fun and the great time that we spent in the office.

Finally, I would like to end my acknowledgments by reciting my dedication of this thesis to my parents. I am very thankful to my great father, Ashraf Sabry, my loving mother, Sahar Maher, and my dearest brother, Raed.

Karim Iskandar

May 8th, 2016

Abstract

Change orders occur in almost every construction project and regularly cause variations to the contractors' anticipated working conditions, resources, and manner of work completion. Change orders are major source of additional congestion, change in sequence, and loss of momentum in the construction jobsite. They frequently cause unforeseen labor productivity loss, which forces contractors to extend their stays on projects. Contractors encounter a lot of resistance from owners when proving productivity loss attributable to change orders, which may lead to unresolved disputes and lengthy litigations.

Previous researchers attempted to set standards and methods in order to quantify the cumulative impact of changes on labor productivity. Some of the previous studies were based on case studies of two or three projects, others included a larger number of projects and more reliable analysis. Generally, it is very difficult to conclusively determine the exact amount of productivity loss attributable to change orders. As a result, there is a continuous need to enhance and enrich the cumulative impact research field.

This current research is based on a database of one hundred and forty-five mechanical and electrical projects, encompassing two project groups: projects impacted by changes, and projects unimpacted by changes. Using two-sample t-tests and Chi-squared tests, a series of numerical and categorical variables were found to be significant in distinguishing between impacted and unimpacted projects, thus revealing the underlying causes of productivity loss associated with change orders.

Furthermore, sixty-eight impacted projects were used in order to quantify the cumulative impact of changes using linear regression analysis. A series of statistical model selection criteria

were applied in order to carefully identify the best predictive models. Candidate models were statistically diagnosed and thoroughly tested to check their validity. Statistical tests and measures were used in order to check whether there are outlying or influential observations in the models. In addition to that, new projects were collected to verify the future predictive ability of the candidate models. The analysis identified the following six factors as best cumulative impact predictors: percent owner initiated change orders, overmanning, turnover, absenteeism, percent time spent by project manager on project, and productivity tracking. The models developed in this research provide the construction industry with means that could be used during dispute resolutions to support the contractors' calculations and assertions for cumulative impact claims.

Finally, this study incorporates a significant statistical component that highlights the most common challenges that analysts face when building linear regression models, such as multicollinearity and the presence of hidden extrapolations. The models developed in this research were extensively analyzed in full details through various statistical tests and measures in order to avoid misleading and deceptive results.

Table of Contents

List of Tables

List of Figures

Chapter 1. Introduction

1.1. Background

The construction industry is an essential contributor to the economy in the U.S., accounting for approximately 4 percent of the nation's Gross Domestic Product (GDP) in 2015. Around 6.6 million people in the U.S. were employed in the construction area in 2015 (Bureau of Labor Statistics 2016). According to (Statista 2016), the total construction put in place in the U.S. is projected to exceed 1.07 trillion U.S. dollars in 2016 and more than 1.3 trillion U.S. dollars in 2019.

Due to the economic importance of the construction industry, there has been a great deal of research investigating the underlying labor productivity trends within the construction sector (Yi, W. et al. 2013). While some studies reported a steady decline of construction labor productivity since the 1960’s (Business Roundtable 1988; Teicholz 2000), others unexpectedly reported an increase in labor productivity trends (Allmon et al. 2000). Although previous literature includes contradicting conclusions, the construction industry is unarguably suffering from countless labor productivity impediments that hinder its performance.

Change orders are one of the most important concerns related to labor productivity. Previous research has reported that change orders are among the major factors that cause a decreased labor productivity (Hanna et al. 1999 a, b; Hanna 2001; Ibbs 2005). Change orders occur in almost every construction project for several reasons, such as scope additions, design changes and errors, insufficient design, schedule compression, and value engineering. They frequently result in disputes and litigations. Change orders’ effects are complex and can be problematic to contractors, because they have indirect and unforeseen effects on work areas that are distant from the changed work itself. The unforeseen effects consist of unanticipated decreases in labor efficiency due to loss of learning, out-of-sequence work, demobilization and mobilization, schedule acceleration, stacking of trades, among others (Hanna 2001). This indirect efficiency loss is known as the cumulative impact of change orders. Cumulative impact of changes can be referred to as "the costs associated with impact on distance work, and are not readily foreseeable or, if foreseeable, are not ready computable as direct impact costs. The source of such costs is the sheer number and scope of the changes to the contract. The result is an unanticipated loss of efficiency and productivity which usually extends the contractor's stay on the job." (Court Case Reference – *Pittman Construction Co.*, GSBCA Nos. 4897, 4923,81-1 BCA 14,847). Generally, it is very hard to prove and measure the cumulative impact of changes. As a result, contractors often waive their rights for additional compensation for the unforeseen effects of change orders.

This research builds upon previous studies conducted at the University of Wisconsin-Madison in order to assess the the cumulative impact of changes on labor productivity. The ultimate goal of this study is to provide the industry with a means of predicting the amount of labor efficiency loss attributable to the cumulative impact of changes.

1.2. Research Motivation

The construction industry is characterized as labor-intensive, since the workforce is the dominant productive resource (Jarkas 2010). More specifically, the Mechanical Electrical and Plumbing (MEP) trades are classified among the most labor-intensive segments of the construction industry due to their high dependence on the labor component. The performance of MEP trades is a key concern of the construction industry. Poor performance of MEP contractors can be one of the root causes of project failures (Arditi and Chotibhongs, 2005). MEP trades play an important role in every construction project, and their performance in terms of timeliness in completion and competency of the overall project team would determine whether a project would be successful (Chan et al., 2004; Lu et al., 2008). In addition to that, MEP contractors constitute 30% to 40% of the total project cost (Hanna et al. 2005; Hanna et al. 2008). They are last in line and carry delays from preceding trades. Moreover, a small increase or decrease in labor productivity can double or wipe out a contractor’s profit (Hanna et al. 2005; Hanna et al. 2008). Due to the high risks associated with MEP trades, this research will focus on examining the impact of change orders on labor productivity, especially for mechanical and electrical contractors.

1.3. Problem Statement

Contractors frequently end up bearing extra costs resulting from consuming additional resources and spending extra work hours as a result of the cumulative impact of changes. Contractors encounter a lot of difficulties and resistance from owners in proving these extra costs and work hours, which often leads to unresolved disputes and lengthy litigations. Construction contracts include clauses that specify procedures for change orders. However, they do not include explicit methods for quantifying the cumulative impact of changes and that would allow contractors to receive proper compensation. Determining and quantifying the cumulative impact of changes is a difficult task because of the inherently complex nature of construction projects. The interconnected factors that control a project’s outcome make it very challenging to isolate and measure the cumulative impact of changes.

Though previous researchers attempted to set standards and methods to quantify the impact of changes on labor productivity, there is still room for improvement in the change order research field. Some of the previous studies were based on case studies of two or three projects (Thomas and Napolitan 1994), others included larger number of projects and more reliable analysis (Leonard 1988; Hanna 1999a,b; Hanna 2001). In Chapter 2, the detailed shortcomings of previous studies will be highlighted.

1.4. Research Objectives

The goal of this research is to provide the construction industry with a quantitative assessment of the cumulative impact of changes on labor productivity. More specifically, this research aims at:

1. Explaining the underlying and hidden factors that lead to the cumulative impact of change orders.
2. Distinguishing between projects impacted by changes and projects unimpacted by changes
3. Quantifying labor productivity loss at the project level and relate it to several productivity indicators.
4. Quantifying labor productivity loss associated with the cumulative impact of change orders.
5. Providing recommendations to contractors and owners in order to to reduce the impact of changes on projects.

1.5. Research Methodology

This study builds upon the original cumulative impact study that was conducted by the Construction Industry Institute (CII) Research Team 158 in cooperation with the University of Wisconsin-Madison (The CII/Hanna study 2001) entitled: *"Quantifying the Cumulative Impact of Change Orders for Electrical and Mechanical Contractors."* The original study involved a wide array of mechanical/electrical contractors and CII committee members who developed a comprehensive survey that targeted projects that are primarily impacted by change orders. The survey was formed based upon an extensive literature review and a full list of factors that are related to the cumulative impact of changes. The survey was carefully reviewed and tested by many mechanical and electrical contractors across the U.S. The CII Research Team RT 158 collected data regarding impacted and unimpacted projects and developed two models: 1) A model that predicts the probability of a project being impacted by changes; 2) a model that quantifies the cumulative impact of changes based on sample size of forty-two impacted mechanical and electrical projects across the country.

This current research focuses on quantifying the cumulative impact of changes, by developing additional models based on a larger sample size of sixty-eight impacted mechanical and electrical projects, including the CII original forty-two impacted projects. The additional twenty-six impacted projects were collected using the same CII survey (Hanna 2001). In this current research, full model diagnostics and exceptionally careful model selection criteria were adopted, as will be shown in Chapters 5. In addition, the models are thoroughly tested to verify their future performance, as will be shown in Chapter 6.

It is worth mentioning that in the CII original research, an elemental schedule compression technique, “Overmanning”, was defined as: “Hiring more craftsmen to work in a location than what would normally be optimal. It is a means of dealing with schedule compression or acceleration (Hanna 2001).” However, in this research, “Overmanning” is defined based on the jobsite manpower ratios, by examining the ratio of actual peak manpower to actual average manpower. Before proceeding to the next chapters, some of the underlying main concepts behind this research should be first reviewed.

1.5.1. Delta Approach

In this study, efficiency loss associated with the cumulative impact of change orders is defined by the term “Percent Delta.” Percent Delta is a macro-level approach that is able to assess productivity loss at the project level, unlike micro-level measurements that focus on specific project activities. As shown in Figure 1, Percent Delta is an hourly-based measurement metric defined as follows (Hanna et al. 1999 a, b):

$$\%Delta = \frac{Total\ Actual\ Direct\ Labor\ Hours - (Estimated\ Hours + Change\ Order\ Hours)}{Total\ Actual\ Direct\ Labor\ Hours}$$

A positive value of Percent Delta indicates that the actual productivity, assessed at the project level, is less than the estimated productivity. On the other hand, a negative value of Percent Delta indicates that a project has achieved a higher productivity level as compared to the estimated productivity (Hanna et al. 1999 a, b). A vital advantage of the Delta Approach is that it uses hourly rather than monetary measurements, which allows for comparing between projects located in different geographical areas and with various sizes. Another advantage is that it is perfectly suits the purpose of this research, as it accounts for the cumulative effects caused by change orders on later phases of a project. In the context of this research, %Delta denotes the productivity loss associated with change orders. In Chapter 5, linear regression models will be developed in order to predict %Delta.

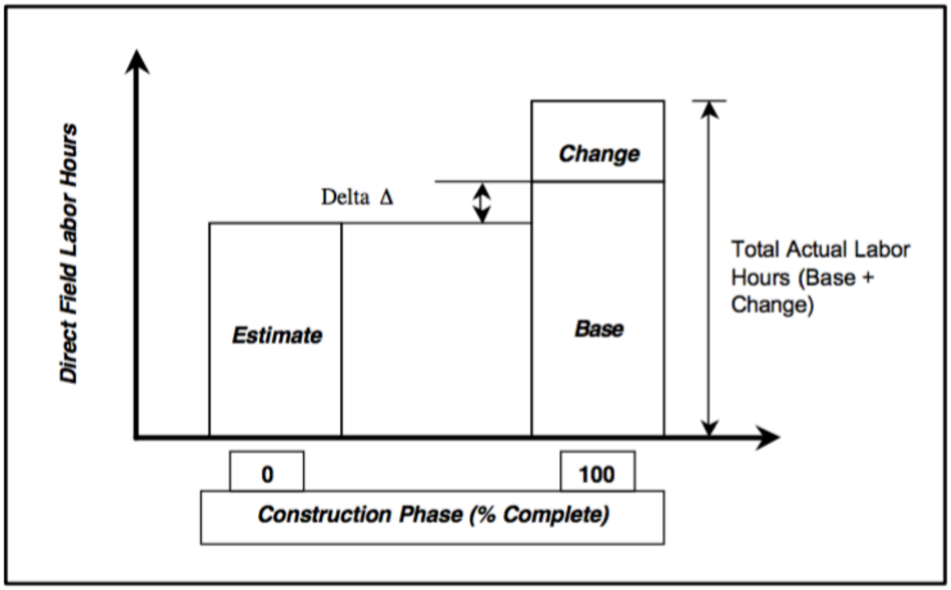

Figure 1. The Delta Approach (Hanna et al. 1999)

1.5.2. Causes of Delta

Delta may result from inaccurate bid estimates. For example, a contractor can incorrectly underestimate the number of work hours that are necessary to accomplish a given project, while in reality, the project actually needs a larger number of work hours. This scenario could result in a positive Delta, part of which would therefore be due to inaccurate contractor estimates. Delta can also result from various productivity-related factors, such as poor contractor performance, adverse weather conditions, among others. Figure 2 shows the possible causes of Delta (Hanna et al. 1999 a, b).

For the purpose of this study, projects were solicited so that change orders were the primary reason for efficiency loss. To ensure that the collected projects do not incorporate inaccurate estimates, a question in the CII survey asked contractors to report the difference between the lowest and the second bidder. If this difference is high, this gives an indication of inaccurate estimates. Furthermore, projects were reviewed with the respondents in order to ensure that efficiency loss was mainly due to change orders.

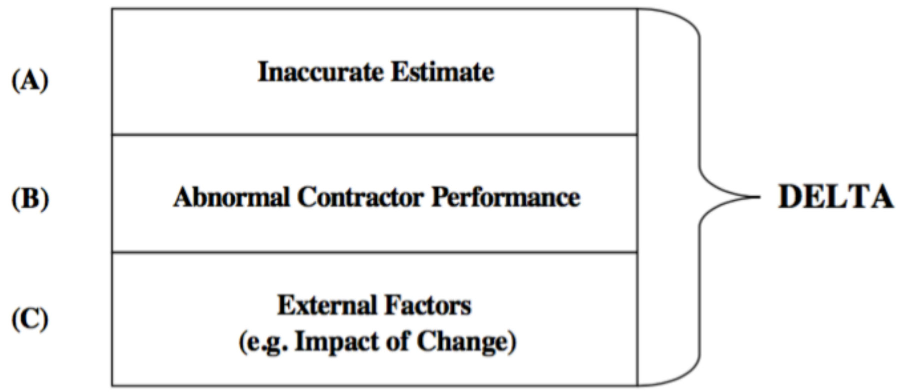

Figure 2. Possible Causes of Delta (Hanna et al. 1999)

1.6. Research Scope

This research is limited to mechanical and electrical projects with Lump Sum contracts and a Design Bid Build (DBB) delivery approach. Lump Sum contracts are essential for applying the Delta Approach, which depends on the contractor's accurate work hour estimates. This study is also limited to projects with sizes larger than 2,000 work hours., as most of the projects in the dataset fall within the range of 2,000 to more than 250,000 work hours. However, the results of this research should be used with caution for projects that are larger than 100,000 work hours. Chapter 3 will discuss the main characteristics of the collected projects in details.

1.7. Thesis Organization

This thesis report consists of seven chapters. Chapter one explains the basic underlying concepts behind this research. Chapter two provides an in-depth literature review in regard to productivity-related factors at the project level, and discusses some of the most noticeable cumulative impact research. Chapter three illustrates some of the main characteristics of the database, and illustrates histograms of Percent Delta and Percent Change for the collected projects. Chapter four reveals the underlying and hidden causes of productivity loss associated with change orders, by statistically examining a series of numerical and categorical variables that would distinguish between impacted and unimpacted projects. Chapter five is the most critical chapter of this thesis report as it focuses on quantifying Percent Delta using linear regression models. In chapter five, a thorough model building process is adopted by applying a series of model selection criteria and detailed model diagnostics. Chapter six involves several statistical procedures in order to verify the predictive ability of the models. Moreover, chapter six examines the performance of the developed models using newly collected projects. Finally, the final research conclusions and recommendations are reported in chapter seven.

Chapter 2. Literature Review

This chapter discusses previous literature pertinent to change orders as well as other productivity-related impediments. Despite the fact that the scope of this research incorporates projects where change orders were the primary reasons for efficiency loss, other productivity-related issues should be thoroughly discussed before proceeding to the next chapters of this research. A practical approach to understanding the effect of change orders on labor efficiency should incorporate a deep knowledge of all factors that are connected to change orders. For example, both excessive and inefficient design would result in an increased amount of change orders in a given project. For work to proceed, the time it takes for contractors to wait for design clarifications and for change orders to be processed decreases labor productivity. Slow response to Request for Information (RFIs) increases the rate of mobilization and demobilization and forces contractors to move workers to other areas where they can produce work, which leads to piecemeal work and loss of learning. Furthermore, contractors may sometimes use scheduling acceleration techniques, such as overmanning, overtime, and shiftwork, in order to recover schedule delays resulting from increased scope. All these consequential effects of change orders are known as the "ripple effect" of changes. In this chapter, we will first begin by discussing the concept of productivity, then we will have a close insight onto change orders and a comprehensive set of other productivity-related factors.

2.1. Productivity Definition

There is no agreement among researchers on a single exact definition of the term "productivity." According to the Concise Oxford Dictionary, productivity is jointly defined by the following three sets of terms: the power of being productive; efficiency; and the rate at which goods are produced. First, the power of being productive is the force behind production itself. Second, efficiency is a measure of how well a given set of factors are utilized. Third, the rate is a measure of the output of the factors of production over a defined period of time. The term "productivity" is generally used to denote a relationship between the outputs and their associated inputs in a production process (Hanna 2001; Yi, W. et al. 2013). In construction, labor productivity is an elemental metric in every project. An example of labor productivity measurement at the task or micro-level would be the labor hours consumed per linear foot of installed electrical conduits, or the labor hours spent for installing one linear foot of ventilation duct. This example expresses productivity as the ratio of input (i.e. the number of actual work hours spent) to output (i.e. the units of work). Lower input to output ratios indicate higher productivity values.

Other productivity definitions rely on a relative measurement of efficiency, such as the definition provided by the American Association of Cost Engineers (AACE) (2013): "productivity is a relative measure of labor efficiency, either good or bad, when compared to an established base or norm." Recalling the delta approach from chapter one, this latter productivity definition could be found very relevant to the approach adopted in this research.

2.2. Productivity-related Factors

In this section, factors that may affect labor productivity at the project level are discussed in detail. As the primary goal of this research is to quantify the cumulative impact of changes, the different types of change orders will be first explained, followed by a brief description of other productivity-related issues.

2.2.1. Changes and Change Orders

A change order is defined as "a written authorization provided to a contractor approving a change from the original plans, specifications, or other contract documents, as well as a change in the cost. With the proper signatures, a change order is considered a legal document." (Means Illustrated Construction Dictionary 2010). Changes can be avoidable or unavoidable. Avoidable changes are preventable, an example of this being the substitution of a type of material with another type to improve quality. Unavoidable changes cannot be foreseen, such as doing rework due to new regulations in a certain work area. The management team should save time and energy by expeditiously agreeing on unavoidable change items, and should direct their effort to resolving issues related to avoidable change items (Hester et al. 1991).

Most construction contracts include a change clause, which enables owners to issue formal changes and allows for contractors to receive equitable adjustments for additional cost or time caused by change orders. In many cases, extra costs and time resulting from the direct impact of changes, or work directly related to changes, are compensable with little debate. On the other hand, the cumulative impact of changes is subject to disputes and disagreements between owners and contractors. This is due to the difficulty in assessing the indirect damages of change orders and their associated reduction in labor productivity for the project as a whole. Previous research has identified several sources of cumulative impact: dilution of supervision, out of sequence work, piecemeal work, mobilization and demobilization, loss of learning curve, trade stacking, change order processing time, request for information processing time, rework, clean-up, and schedule acceleration (Hanna 2001). Generally, it is impractical to consider the effects of a single change in isolation. Therefore, the impact of multiple changes should be studied in tandem in order to determine their overall impact on the job.

2.2.1.1. Types of Change Orders

Change orders can be classified into several categories: Formal changes, constructive changes, and cardinal changes. Formal changes are changes that occur within the scope of the project. With a change clause, an owner has a unilateral right to add or delete work, or both, within the scope of the project. This refers to changes directed by the owner. Constructive changes are other types of changes where an action/inaction of one of the parties requires change in the performance of another party. An example of this type of change is when the designer's or construction manager's conduct requires the contractor to perform work other than what's specified in the original contract. This can be due to defective specifications and/or latent ambiguities, which leads to extra work. In addition to that, changes can sometimes be minor, and are not intended to alter the original time and cost for the project. However, a contractor should promptly submit a claim in order to get compensation for minor changes at a later time (Sweet 1994). In all these cases, the contract terms should control the subsequent cost and time adjustments, as long as the changes are within the scope of the project.

Cardinal Changes, however, are changes outside and far beyond the original scope of the contract, resulting in a materially different undertaking. Despite a change clause, cardinal changes can be considered as a breach of contract by the owner, which entitles the contractor to quit or to continue performing the project. If the contractor decides to continue, he/she should be compensated based on the actual cost of work or services. Elements of a cardinal change should include a radical alteration in the cost of work, the quantity of work, or the character/nature of work.

2.2.1.2. Methods of Quantifying the Cumulative Impact

There are generally two approaches for quantifying the cumulative impact of changes: the micro approach and the retroactive approach. The micro approach is a proactive method where each productivity factor is evaluated separately. It is also called the factor approach. On the other hand, the retroactive approach encompasses methods that evaluate the cumulative impact after the fact. The following methods include both proactive and retroactive methods:

Total Cost Method: This method is accepted by some courts, but it is the least favored. This method relies on subtracting the estimated cost of the project from the actual cost incurred. The resulting difference in that case is assumed to be due to the owner only, which is a skeptical approach that should be used as a last resort. A major disadvantage of this method is that it can reward the contractor's inefficiencies and managerial incompetence. This method should be only used when each of the following four conditions are met: 1) actual damages and nature of loss cannot be determined with reasonable accuracy; 2) the estimated cost of the project was realistic; 3) the contractor's actual costs were reasonable; 4) and the contractor was not responsible for the added costs (Schwartzkopf 1992).

Modified Total Cost Calculations: The rationale behind this method is that the equation of the total cost method is adjusted so that owners are no longer responsible for contractors' performance inefficiencies and errors in bid estimates. The use of this method is strengthened when the cost attributable to the contractor's inefficiencies are accurately proven and subtracted from the equation (Schwartzkopf 1992).

Measured Mile Calculations: The "Gold Standard." This is the most broadly accepted method for calculating lost productivity. In this method, similar activities are compared on impacted and unimpacted segments of the project in order to determine the productivity loss stemming from the impact. The difference in productivity between the impacted and unimpacted periods is considered to be the lost productivity. This method is favored because it only takes into account the claimed impact, an approach that avoids any doubt regarding the validity of cost estimates. A disadvantage of this method is that in highly distressed and troubled projects, it is hard to isolate unimpacted from impacted periods. It is also difficult to find two different periods where identical activities were being performed (Schwartzkopf 1992).

Industry Publications: Industry publications are often used to prove the productivity loss associated with change orders. Courts and boards sometimes accept many reliable industry publications established by recognized researchers and practitioners.

Experts and Consultants: Consultant and expert testimony are often used to prove lost productivity in construction projects. In that case, opinion of experts is not sufficient. Supportive documents incorporating the analysis of actual circumstances and cost data of the project are needed to prove the actual incurred decreased productivity (Schwartzkopf 1992).

2.2.1.3. Previous Cumulative Impact Research

Various studies have been conducted in order to asses the impact of change orders on labor productivity. Some of the them were based on case studies of two or three projects, others included larger number of projects and more reliable analysis. In this section, a brief description of the most noticeable cumulative impact studies is presented.

2.2.1.3.1. Leonard Study

Based on 90 cases of construction disputes, Leonard (1988) developed relationships between productivity loss and percent change for electrical/mechanical work and civil/architectural work. Productivity loss was measured as the ratio of the unproductive labor hours to the actual labor hours spent on the original contract (Leonard 1988). Figure 3 shows the relationships for electrical/mechanical. Similar graphs were also developed for civil/architectural work. The graphs show productivity loss attributable to the following three cases: 1) change orders are the only cause of productivity loss; 2) change orders plus one additional major cause are the source of productivity loss; 3) change orders plus two additional major causes are the source of productivity loss. In the Leonard study, the term “additional major cause” refers to the use of an acceleration technique, adverse weather conditions, or poor site management and logistics (Moselhi et al. 1991 and Leonard 1988). Leonard reported that the cumulative impact of changes result in lost productivity due to out of sequence work, low labor morale and demotivation, lack of engineering support, and the use of acceleration techniques when equitable adjustments are not settled (Schwartzkopf 1997). Although it was the first effort to investigate the cumulative impact of changes, the Leonard method had some drawbacks in that it only considered small and building projects that were already troubled at the dispute stage (Jones 2001; Revay 2002).

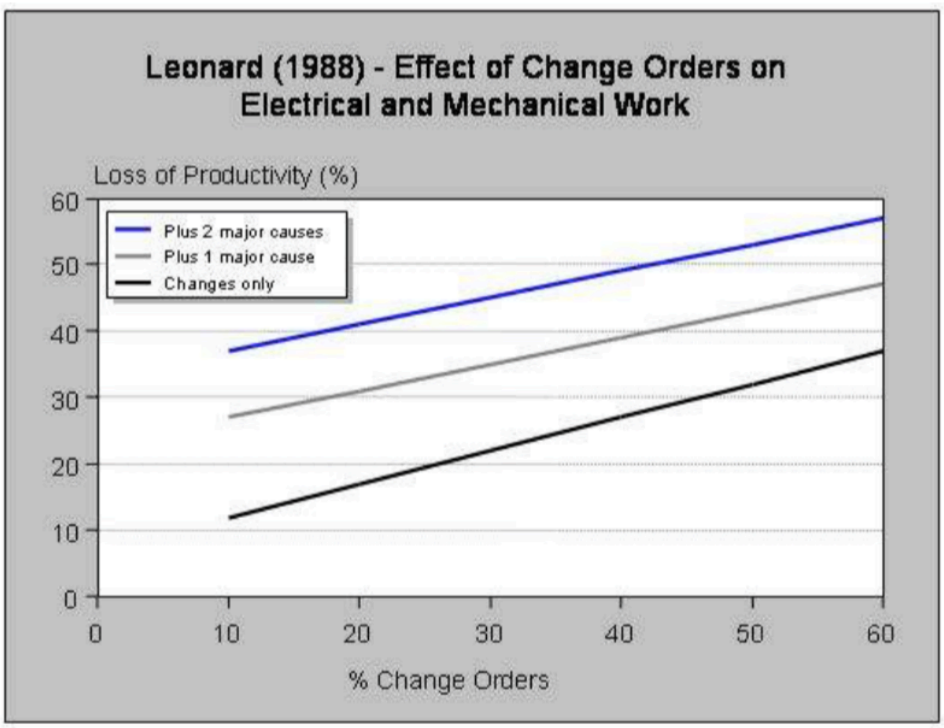

Figure 3. Leonard curves for Mechanical and Electrical Projects. Source: (Leonard 1988)

2.2.1.3.2. CII – Construction Changes and Change Orders: Their Magnitude and Impact

In a CII study, Hester collected field data in order to investigate the impact of changes on various productivity factors, including cost and schedule, while also providing the industry with techniques for effectively managing changes (Hester et al. 1991). Although the study did not provide a quantitative assessment of the impact of changes, several conclusions were reported. Among the conclusions is that it is unpractical to assume that the regular change in the work will only affect the work in the changed area. The effects can extend beyond that area and encompass other work areas (Hester et al. 1991). Also, it was reported that contractors should examine the potential impact of changes on labor productivity for new projects, and consider the impact when pricing their bids. However, this is not practical because contractors often tend to lower their prices in traditional delivery approaches in order to win bids.

2.2.1.3.3. Hanna Methods

In 1999, two studies were completed at the University of Wisconsin-Madison in order to quantify the cumulative impact of change orders on mechanical and electrical construction labor productivity (Hanna et al. 1999a, 1999b). Both studies used the percent delta approach, and the following two predictive models were developed:

<u>Electrical Model:</u>

$$\text{Delta\%Tot} = -22.0 - 0.14*\text{MgrYears} + 6.47*\text{Ln}(\text{EstCO\%Est}) + 9.66*\text{Ln}(\text{EstCO}) - 0.90*(\text{Ln}(\text{EstCO}))^2$$

Where,

- MgrYears = Number of years the Project manager has been in the construction industry
- EstCO%Est = Change Order hours as percent of original estimate hours

- EstCO = Estimate of change order hours

Mechanical Model:

$$\text{Delta\%Tot} = -0.16190 - 0.001534*\text{CHGEST} - 0.000729*\text{NUMCHG} + 0.07934*\text{WTIMING} + 0.000032*\text{CHGEST}*\text{NUMCHG}$$

Where,

- CHGEST = Change order hours as percent of original estimate hours
- NUMCHG = the number of change orders
- WTIMING = The weighted timing of the change order occurrence

Once it is determined that a project is impacted by changes, the two models can be used. However, determining whether a project is impacted was subjective and based on past experience and industry knowledge of contractors who participated in the study. Consequently, the CII in conjunction with the University of Wisconsin-Madison have developed a cumulative impact study where the term "Impact" was quantitatively defined, as will be shown in the upcoming section.

2.2.1.3.4. CII/Hanna Method (2001)

Through the CII and with the support of the Mechanical and Electrical Contracting Foundations in the U.S., a comprehensive study was developed to deeply investigate the cumulative impact of change orders on labor productivity (Hanna 2001). Two models were developed throughout the study, as shown in Table 1. The first model is a logistic regression model that served as a quantitative means for defining the impact, by calculating the probability of a project being impacted by changes. The second model is a linear regression model, which quantitatively predicts the cumulative impact of changes using the Delta approach. In the linear regression model, there is a counterintuitive coefficient sign for an elemental predictor: "Overmanning." In fact, when overmanning occurs on a project, %Delta (productivity loss) is expected to increase. The counterintuitive coefficient sign was interpreted that it was due to multicollinearity.

Table 1. Logistic and Linear Regression Models (CII/Hanna Method 2001)

Logistic Regression Model	Linear Regression Model
Sum of Factors = - 6.997 - 1.0939 MechanicalorElectrical + 3.889 %Change*MechanicalorElectrical - 1.0371Estimated Peak/Actual Peak + 0.6342 ProcessingTime +2.6433 Overmanning +1.1933 Overtime + 1.2048 Peak/Average Manpower + 0.017154 Percent Change Orders related to Design Issues Probability (Impact) = $\frac{e^{Sum\ of\ Factors}}{1+e^{Sum\ of\ Factors}}$	%Delta= 0.36866 +0.11957 PercentChange -0.08065 PM%TimeOnProject -0.16723 %OwnerInitCO -0.09147 Productivity -0.05213 Overmanning +0.022345 ProcessingTime

2.2.2. Design Deficiencies and Inadequate Designer Support

Design errors, omissions, and insufficiencies are common sources of change orders in construction projects. When contractors frequently wait for clarifications and further instructions from designers regarding design deficiencies, work disruptions and efficiency loss occur. Signs of efficiency loss are lengthy change order processing time, excessive Requests for Information (RFI's) and an abundance of clarification requests.

When designers do not provide adequate and immediate support during construction, contractors may encounter rework and out of sequence flow of work. In addition, when designers are constrained with tight design schedules, the process of developing the contract documents may include a lot of errors and omissions. When not adequately coordinated prior to construction, these design errors and discrepancies would result in productivity losses during later phases of a project, due to the time wasted in revising design documents and correcting errors.

Another problem related to design issues, especially in Lump Sum contracts, is that contractors are totally separated from designers and are not sharing their expertise with designers during early phases of a project. Contractors should provide their input and should have their own ingenuity to get the project done. Other project delivery methods, such as the Integrated Project Delivery (IPD) approach, enable a highly cooperative atmosphere for all project parties throughout all project phases. However, this is not of our concern here, as our current research is exclusively conducted for projects delivered using the traditional Design Bid Build (DBB) approach.

2.2.3. Schedule Compression

When a project experiences productivity loss due to change orders, contractors often find themselves employing acceleration techniques in order to achieve the project completion date set forth in the contract. Schedule compression (or acceleration) simply refers to the reduction of time available to complete a project. A considerable number of inefficiencies can result from schedule compression because when a schedule is accelerated, design issues and RFI's would more often occur, materials would be installed at a faster rate, and possibility of errors would highly increase. When schedule compression is significant, the ability of a contractor to adequately plan ahead will be diminished due to extensive deviations in the time schedule and the resulting alterations to the ordinary or planned sequence of work. The top three common methods that are used when facing schedule compression are overtime, shift work, and overmanning.

2.2.3.1. Overtime

Overtime is the work performed beyond the reference schedule of 8 hours per day and 5 days per week. Overtime can occur in a variety of schedules, such as 5 days of 10 working hours per day, 6 days of 10 working hours per day, 7 days of 10 working hours per day, among others. Overtime introduces many issues such as workers' fatigue, low labor morale, higher expenses, and a phenomenon described by the U.S. Army (1979) where workers tend to pace themselves to adjust their work according to longer time spans. Contractors often use overtime in response to an accelerated schedule. Nevertheless, previous researchers reported that productivity is reduced when overtime is used, but the magnitude of reduction varies from study to study.

In 1992, Thomas reported that many previous overtime studies are incomplete and based on opinions rather than factual and reliable data (Thomas 1992; NECA's Overtime study 1969; Howerton 1969). However, the results were generally consistent in that there is about a 10% increase in efficiency losses for each additional 10 hours per week added to the schedule beyond 40 hours. In 1984, the CII sponsored a study of construction overtime ("The Effects" 1988). The report concluded that productivity does not necessarily decrease on an overtime schedule.

Thomas also reported losses of efficiency of 10-15% for 50-hour and 60-hour work weeks (Thomas 1997). Furthermore, (Hanna et al. 2005) developed a model that quantified the amount of productivity loss resulting from extended overtime use. The model was able to predict the amount of loss depending on the average number of work hours per crew member per week as well as the total number of actual work hours expended while using a single specified crew scheduling technique (Hanna et al. 2005). In addition to that, Hanna agreed with previous studies in that 50 hours per week and 60 hours per week overtime schedules yield to lower productivity levels as compared to 40 hours per week schedules (Hanna et al. 2005).

2.2.3.2. Shiftwork

Shift work is defined as the amount of work executed by a second group of craftsmen whose work is performed after the first group of the same trade has retired for the day. The amount of weekly work hours can be doubled when a second shift schedule is used, which can significantly increase the work production. Shift work, however, requires additional costs for extra administration personnel, supervision, and quality and safety control. (Coburn 1997) reported that a significant amount of shift work cost is due to reduced human performance at work. Problems associated with shift work include additional work coordination requirements due to little cooperation between shifts, inconsistent work procedures among different shifts, inefficient communication between crews, and the unavailability of timely administrative decisions away from regular business hours (Penkala 1997).

Shiftwork can lead to positive effects and productivity gains if used on a limited amount. Haring reported that night shift can result in a 20–25% savings in productivity when the amount of shiftwork is 1% of the total project hours (Haring 1981). Furthermore, Hanna reported that productivity savings can reach up to 11% due to shift work use (Hanna el al. 2008). However, Hanna reported that efficiency loss due to shiftwork can reach 17% depending on the amount of shiftwork used, as shown in Figure 4. Negative values for efficiency loss in Figure 4 indicate that there were productivity savings. The X-axis in Figure 5 represents the percent shiftwork, denoted by the total number of shiftwork hours as a percentage of total budgeted work hours for a given project.

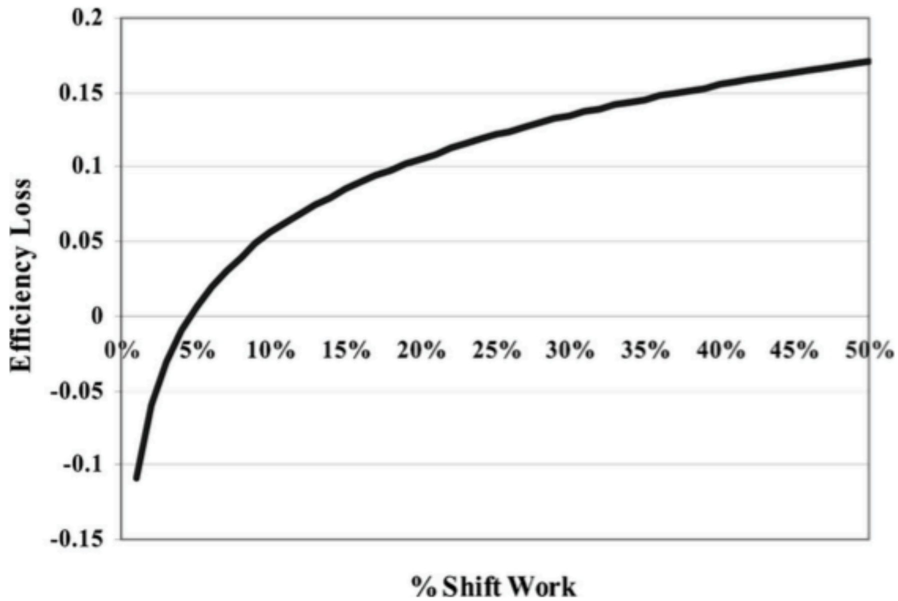

Figure 4. Efficiency Loss and Shiftwork (Hanna et al. 2008)

2.2.3.3. Overmanning

Overmanning can be understood in two different ways. First, overmanning can be defined as a situation where more workers are consumed on a certain job/trade so that the crew size for that job/trade is no longer optimal. The optimal crew size is the minimum number of workers required to perform a task within the allocated period (U.S. Army 1979). Second, according to (Hanna et al. 2005), overmanning can be defined as a situation where the actual peak manpower to the actual average manpower exceeds 1.6. This definition is related to manpower consumption or manpower loading curves.

The simple empirical relationship to manpower loading is that the maximum manpower on a job is 160% of the average manpower requirement. The maximum manpower first occurs when 50% of the project time has elapsed, and the period of maximum manpower occurs for 25% of the project time (Allen 1979; Wideman 1994). The standard resource input profile shown in Figure 6 has a shape of a trapezoid, indicating three typical stages for a planned manpower loading curve: build-up, steady-state, and run-down (Allen 1979; Wideman 1994). The profile indicates that 40% of the total manpower requirements occurs in the first 50% of the time, a further 40% of the total manpower requirements occurs in the next 25% of the time, and the last 20% of the manpower requirements occurs in the last 25% of the time (Allen 1979; Wideman 1994). The profile shown in Figure 5 is applicable at the task, trade, subcontract, or at the whole project level (Wideman 1994).

Problems associated with overmanning are site congestion and dilution of supervision. To preserve the quality of a project when overmanning occurs, more intensive supervision is required.

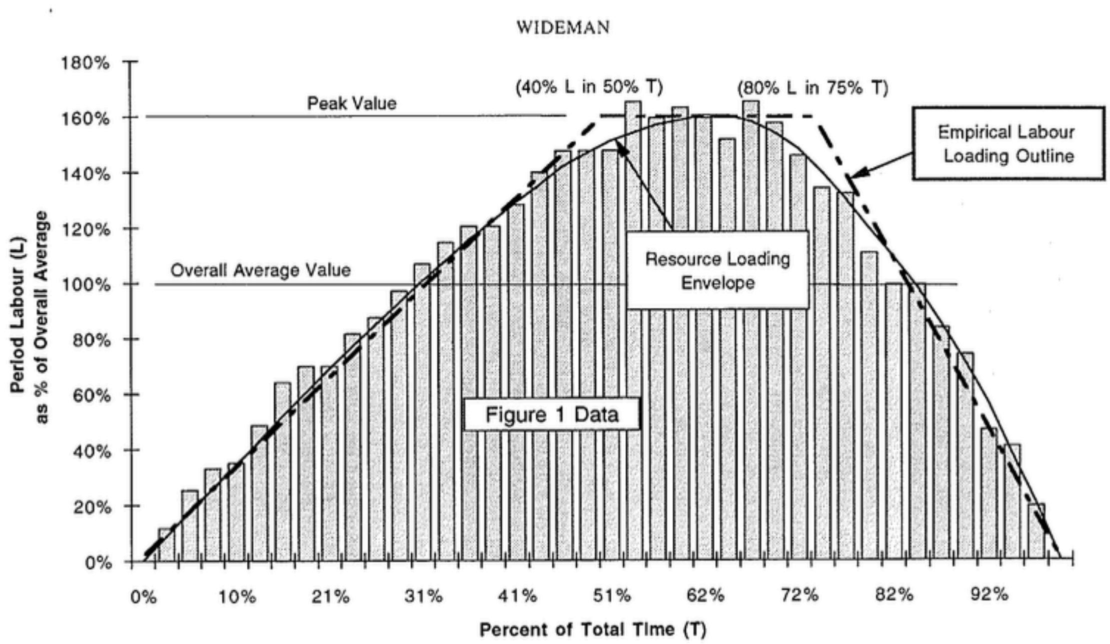

Figure 5. Empirical resource loading input of site manpower (Wideman 1994)

2.2.4. Absenteeism and Turnover

Absence and turnover of workers are critical issues in the construction industry. Reasons for absenteeism and turnover include excessive on-site rework, unsafe working conditions, and high commute time to the worksite, and the availability of overtime work in other projects. Numerous change orders can cause more on-site rework, which lowers labor morale and increases absence and turnover rates. Previous researchers investigated the causes and consequences of absenteeism and turnover. Some researchers examined the role of social norms and absence culture on construction projects (AbouRizk et al. 2010; Ahn et al. 2013; Sichani et al. 2011); others quantified the effects of absenteeism and turnover on labor productivity (Hanna et al. 2003). The literature indicated that high absenteeism and turnover rates can reduce the production of any workforce. (Hinze 1985) noted that reduced production can lead to increased costs for contractors and owners. (Hanna et al. 2005 and Hinze 1985) stated that frequent absence of workers can cause workflow disruptions and reduction in labor productivity.

The Business Roundtable Report (1982) indicated several consequences of excessive absenteeism and turnover, including loss of revenue due to delays, additional administrative work for recruiting and training new employees, lower efficiency of new or inexperienced workers, and lower employee morale. The Business Roundtable (1982) reported that companies lose 12 work hours for each absent worker and 24 hours for every termination, whether quits or firings. Regarding turnover, Ahn noted that frequent involvement of inexperienced replacement workers may increase the likelihood of accidents due to the existing pressure of having to maintain the project time and cost targets (Ahn et al. 2013). It can therefore be concluded that decreased absenteeism and turnover levels result in higher productivity rates and significant reductions in project costs.

2.2.5. Learning Effect

The time and effort expended to complete repetitive activities decrease as the number of repetitions increases. This phenomenon is known as the learning effect. Reasons for the existence of such effect include increased worker familiarization, development of more efficient techniques and methods, and better work coordination throughout repetitive tasks. Having a plenty of change orders can prevent workers from getting better and faster in doing repetitive tasks.

2.2.6. Stacking of Trades

Stacking of trades describes project conditions where multiple trades are working simultaneously in a single work area, resulting in a decreased square footage or component spacing per worker. Trade stacking can occur both under normal and accelerated schedule conditions. Trade stacking results in site congestions, crew interferences, and lower productivity. Change orders can lead to trade stacking. Changed work can involve various tasks at the same time, which results in a situation where multiple trades get in each others way in order to concurrently perform different tasks.

2.3. Summary of Previous Research

This chapter showed that a significant amount of research has been conducted to examine the most critical productivity issues that hinder the performance of construction projects. Table 2 presents a summary of the most noticeable productivity-related studies that were conducted by previous researchers, classified by year and specific research area. The information presented in the table below only considers the reliable studies that provided useful and valuable information.

Table 2. Valuable Information from Previous Researchers

Authors	Year	Research Focus	Remarks
Leonard	1988	Change Orders	Data taken from projects that are in the dispute phase. Curves were developed to relate productivity loss to the amount of change.
Hester et al.	1991	Change Orders	This study did not provide quantitative assessment of the impact of changes. However, it provided important conclusions and best practices.
Hanna and Russell	1999	Change Orders	Two cumulative impact predictive models: one for the electrical trade, and one for the mechanical trade.
Hanna and Russell	2001	Change Orders	A logistic impact model for predicting the probability of a project being impacted by changes and a linear regression model that predicts productivity loss for impacted projects.
Hanna and Loh	2004	Change Orders	A decision tree approach to classify and quantify the cumulative impact of changes.
Thomas	1992	Overtime	The study concluded that there is about 10% increase in efficiency losses for each additional 10 hours per week added to the schedule beyond 40 hours.
Thomas	1997	Overtime	The study reported loss of efficiency of 10-15% for 50-hour and 60-hour work weeks.
Hanna et al.	2005	Overtime	50 and 60-hour work week yield to lower productivity levels as compared to 40-hour work week schedule.
Haring	1981	Shiftwork	Night shift can result in 20-25% saving s in productivity when the amount of shiftwork is 1% of the total project hours.
Hanna et al.	2008	Shiftwork	Productivity savings can reach up to 11% due to shiftwork use. However, efficiency loss due to shiftwork can reach 17% depending on the amount of shiftwork used.
Wideman	1994	Overmanning	Investigate empirical manpower curves: 40% of the total manpower requirements occurs in the first 50% of the project time, a further 40% occurs in the next 25% of the time, and the last 20% of the manpower requirements occurs in the last 25% of the time.
Hanna et al.	2005	Overmanning	Overmanning occurs when the ratio of actual peak manpower to actual average manpower exceeds 1.6.
The Business Roundtable	1982	Absenteeism	Companies lose 12 work hours for each absent worker and 24 hours for each termination. Absenteeism results in loss of revenue due to delays, additional administrative work for recruiting and training more employees,, low labor morale, and low efficiency of new or inexperienced workers.
Ahn et al.	2013	Turnover	Frequent involvement of inexperienced replacement workers may increase the likelihood of accidents due to the existing pressure of having to maintain the project time and cost targets.

2.4. Summary

In this chapter, a comprehensive literature review regarding productivity-related issues was presented. A thorough understanding of those productivity issues is a key step before proceeding to the next chapters. In addition to that, chapter 2 discussed previous cumulative impact research, and highlighted some of the drawbacks associated with previous studies.

Chapter 3. Data Characteristics

This study amalgamates a total of one hundred and forty-five projects, including two project groups: impacted projects and unimpacted projects. Out of the one hundred and forty-five projects, sixty-eight projects were impacted. Projects were collected from contractors randomly selected from the National Electrical Contractors Association (NECA) directory and the Mechanical Contractors Association of America (MCAA) directory. The survey used to collect the data was established during the development of the CII's famous cumulative impact research (Hanna 2001). The sample size used in this current research is fairly large, which allowed for a robust statistical analysis, as will be shown in Chapters 4, 5, and 6. In this Chapter, a brief description of the mechanical and electrical projects included in the dataset are presented.

3.1. Project Size

In this study, project size is defined as the actual work hours utilized at project completion, including change order hours. By combining all 145 projects, it was found that the total actual labor hours of the entire dataset are equal to 6,811,909 work hours. Figures 6 and 7 show the histograms of project size for mechanical and electrical projects, respectively. The mean project size for mechanical projects is 34,990 work hours, while the mean project size for electrical projects is 47,380 work hours. Although some projects have large sizes of up to 260,00 work hours, most of the projects in the dataset have project sizes that fall within the range of 2,000 to 100,000 work hours. Consequently, results of this research should be used with caution for projects exceeding 100,000 work hours.

It is worth mentioning that 68% of the projects have exceeded their original expected duration, while only 32% of the projects have met or completed the required work ahead of schedule. In construction, the majority of projects fail to meet their schedule targets due to the productivity impediments that are addressed in this current study.

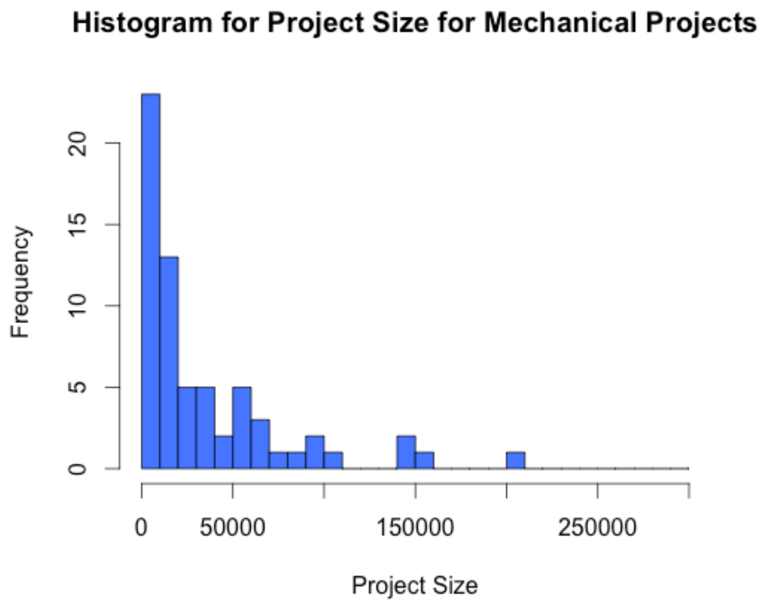

Figure 6. Distribution of Project Size for Mechanical Projects

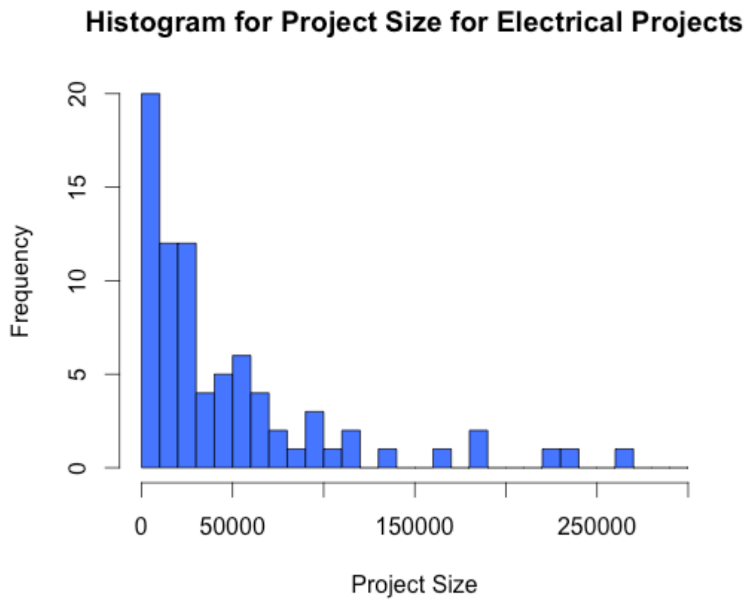

Figure 7. Distribution of Project Size for Electrical Projects

3.2. Percent Delta

Figures 9 and 10 show the histograms of Percent Delta for mechanical and electrical projects, respectively. The values shown in Figures 8 and 9 include both impacted and unimpacted projects. The mean percent delta for mechanical projects is 12.12%, while the mean percent delta for electrical projects is 10.76%. Figure 10 shows the histogram of percent delta for all electrical and mechanical projects combined. The distribution shows a fairly even split between impacted and unimpacted projects. In the next chapter, we will discuss the exact percent delta threshold used to separate impacted from unimpacted projects.

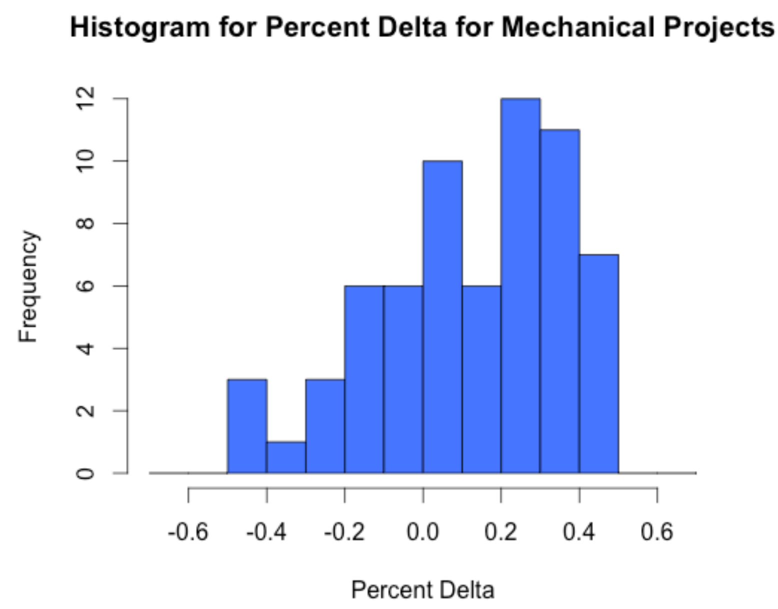

Figure 8. Distribution of Percent Delta for Mechanical Projects

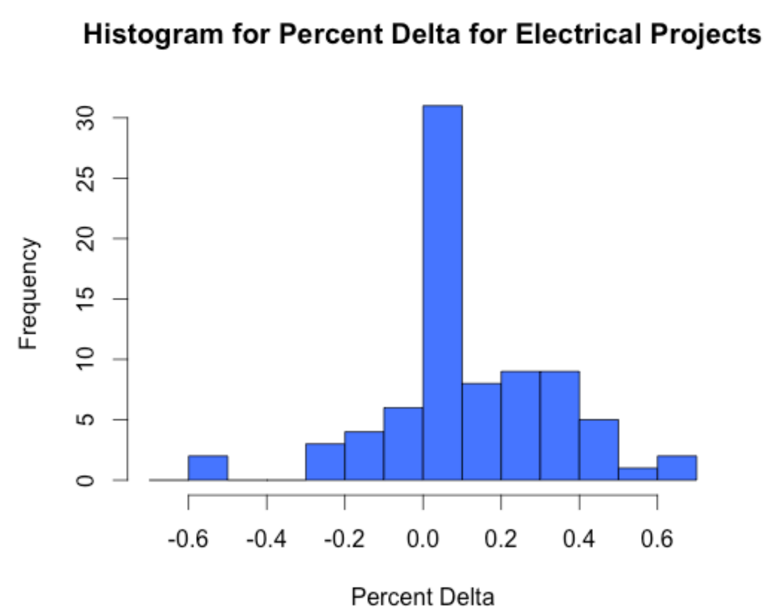

Figure 9. Distribution of Percent Delta for Electrical Projects

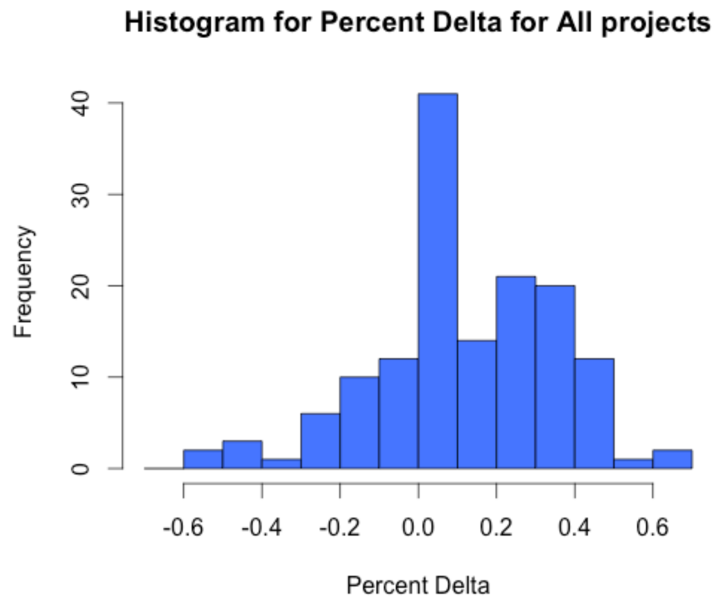

Figure 10. Distribution of Percent Delta for All Projects

3.3. Percent Change

Percent change is defined as the total owner approved change order labor hours plus the absolute value of the total credit change order hours, divided by the original budgeted labor hours (Hanna 1999 a, b; Hanna 2001). The absolute value is used for credit change order hours in order to obtain the true amount of changes that occur in a given project, whether there are scope additions or deletions. Below is the equation that defines percent change (Hanna 2001):

$$\%Change = \frac{(Total\ Change\ Orders\ Approved\ Hours) + (|Total\ Change\ Orders\ Credit\ Hours|)}{Original\ Budgeted\ Labor\ Hours} * 100$$

Using the above mentioned equation, it was found that the percent change for all 145 projects combined is 41.5%. Figures 11 and 12 show the histograms of percent change for mechanical and electrical projects, respectively. The mean percent change for mechanical projects is 28.7%, while the mean percent change for electrical projects is 25.6%. Projects with percent change higher than 90% were excluded from this study. Projects with extremely high values for percent change represent very unique and abnormal cases, where cardinal change may have occurred.

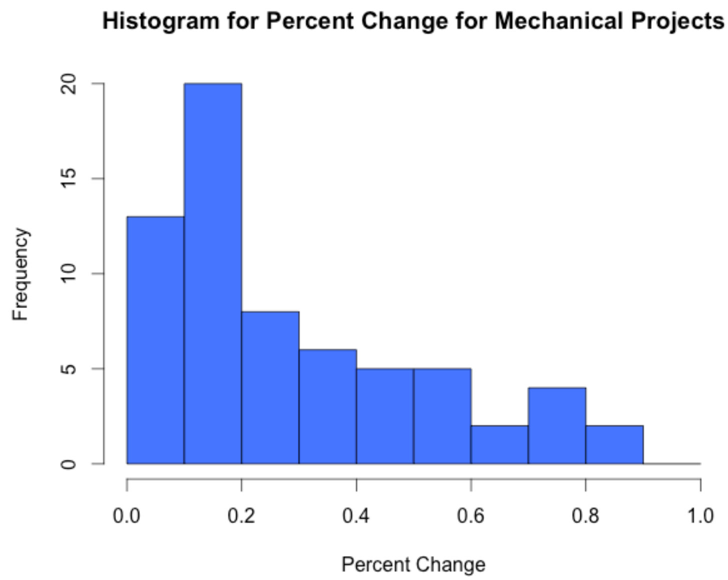

Figure 11. Percent Change for Mechanical Projects

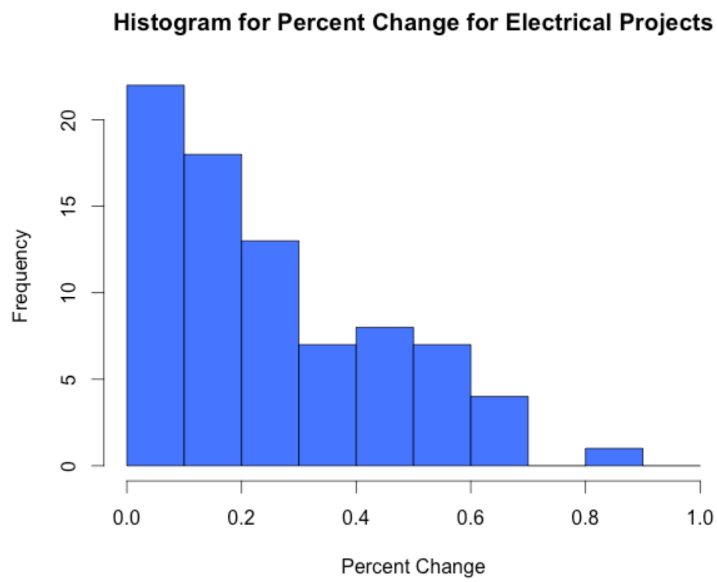

Figure 12. Percent Change for Electrical Projects

3.4. Reasons for Change Orders

Change orders can occur due to various reasons. The pie chart shown in Figure 13 shows the distribution of the projects included in the dataset with respect to the reasons for change order work. The pie chart shows that 39% of change orders were due to scope additions, 26% were due to design changes, and 18% were due to design errors. Change orders seldom occur as a result of value engineering, as shown in the pie chart below.

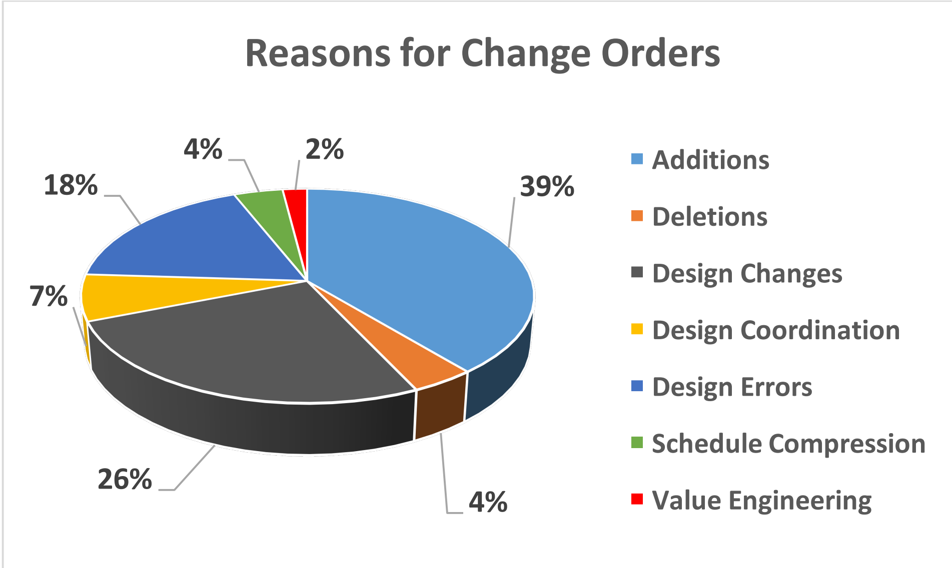

Figure 13. Reasons for Change Orders

3.5. Mechanical and Electrical Project Data

The 145 projects included in the current dataset are geographically located in more than 25 States in the U.S. Of the 145 projects, 80 are electrical and 65 are mechanical. Figures 14 and 15 show the distribution of electrical and mechanical projects for impacted and unimpacted projects, respectively. For impacted projects, the distribution in Figure 14 shows an even split between mechanical and electrical work. For unimpacted projects, 39% are mechanical, while 61% are electrical, as shown in Figure 15.

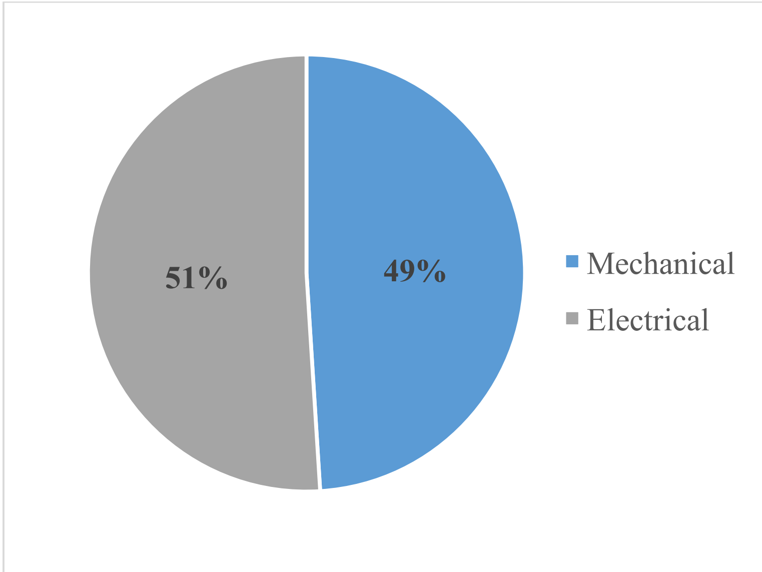

Figure 14. Distribution of Impacted Mechanical and Electrical Projects

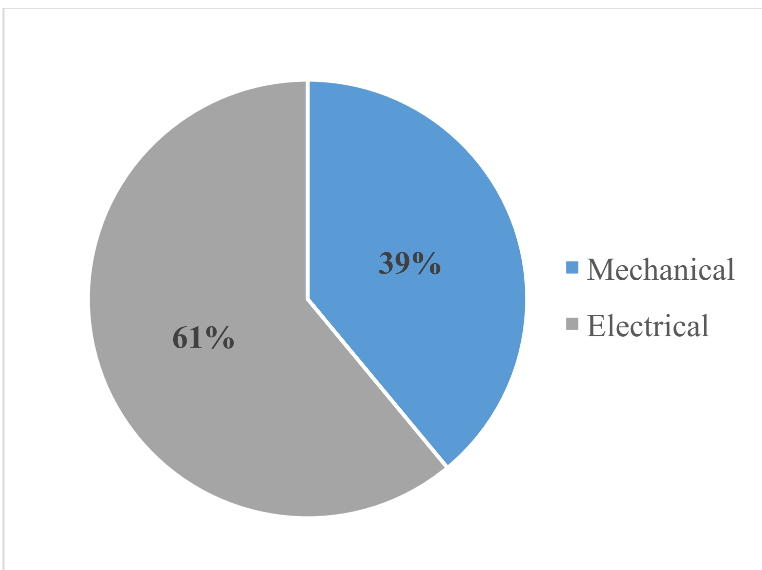

Figure 15. Distribution of Unimpacted Mechanical and Electrical Projects

3.6. Type of Owner

Figures 16 through 19 show the distribution of projects by owner type for four categories of projects: Mechanical impacted projects, mechanical unimpacted projects, electrical impacted projects, and electrical unimpacted projects. The majority of projects belong to the private sector, and the ratio of private to public projects is almost constant across all four categories. More specifically, 74% of the mechanical impacted projects are private, while 26% are public, as shown in Figure 16. For the other three categories, two-thirds of the projects are private, while one-third are public, as shown in Figures 17, 18, and 19.

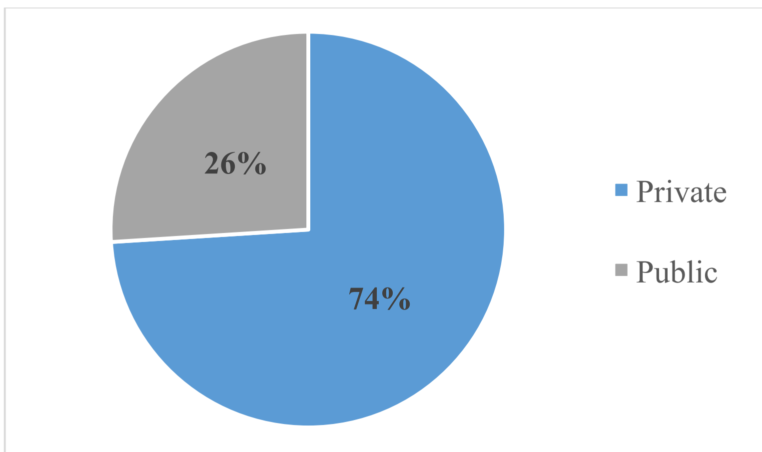

Figure 16. Distribution of Mechanical Impacted Projects by Owner Type

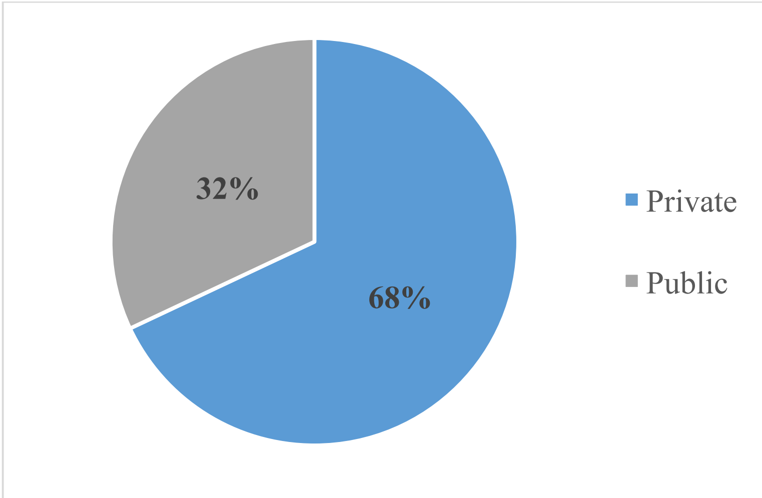

Figure 17. Distribution of Mechanical Unimpacted Projects by Owner Type

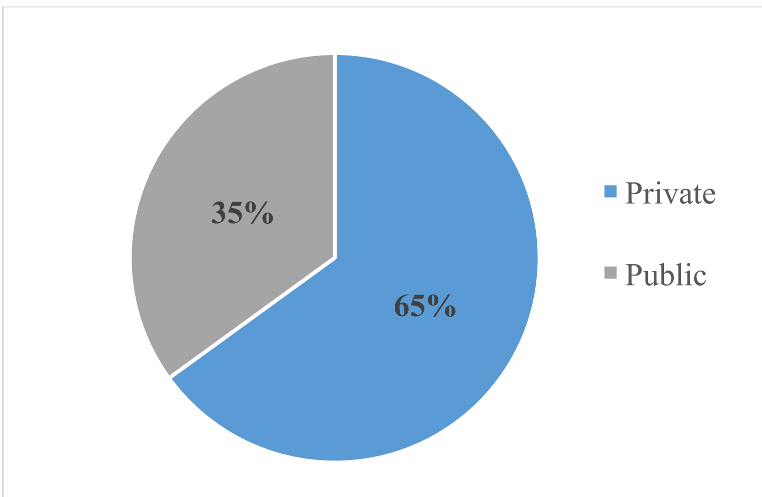

Figure 18. Distribution of Electrical Impacted Projects by Owner Type

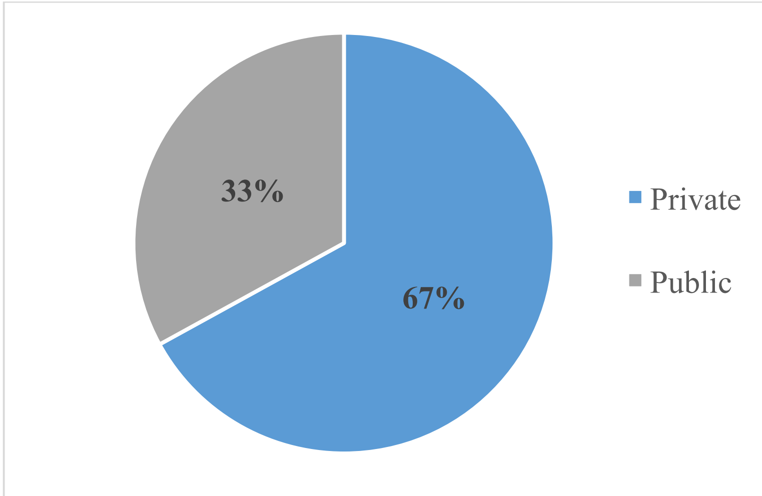

Figure 19. Distribution of Electrical Unimpacted Projects by Owner Type

3.7. Construction Work Types

The dataset used in this thesis report includes all types of construction work. Figure 20 shows that almost half of projects involved new construction work, whereas 28% of the projects were addition/expansion work and 24% of the projects were renovation/rehabilitation work.

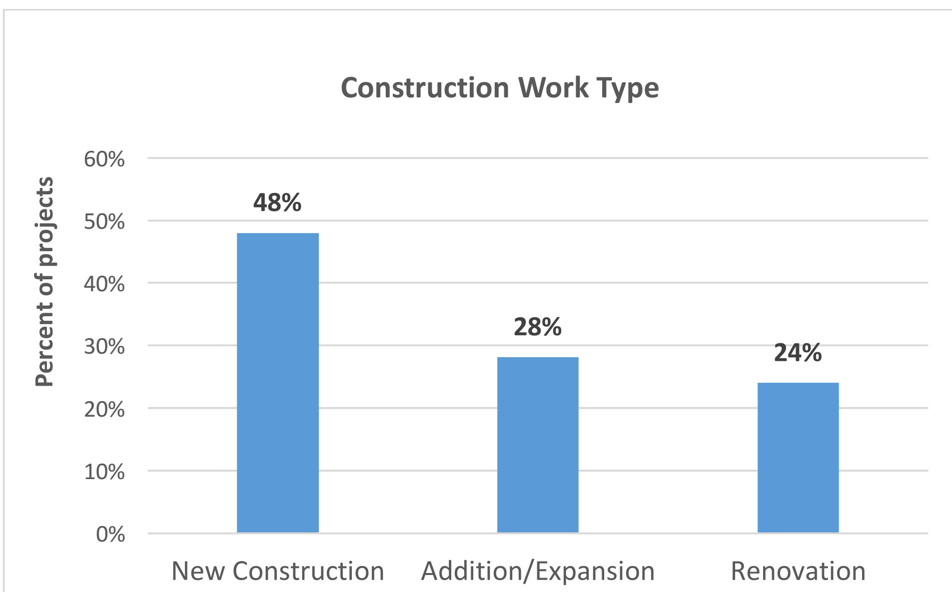

Figure 20. Construction Work Type

3.8. Project Types and Previous Work Experience

Three-thirds of the contractors reported that they had past work experience with project owners, as shown in Figure 21. The distribution shown in Figure 21 was the same for impacted as well as for unimpacted projects, indicating that past work experience with owners is not a significant factor in determining whether a project is impacted by changes. Figure 22 shows that for the majority of projects, owners had previous work experience regarding the same project type. The dataset incorporates a wide variety of project types, including commercial, institutional, industrial, manufacturing, power plants, residential, and maintenance projects. The distribution of project types included in the dataset had a similar pattern as compared to the distribution of the entire mechanical and electrical industries. For mechanical projects, 45% of the projects were industrial, 20% were commercial, and 15% were institutional, as shown in Figure 23. For electrical projects, around 40% of the projects were commercial, 30% were industrial, and 10% were institutional, as shown in Figure 24.

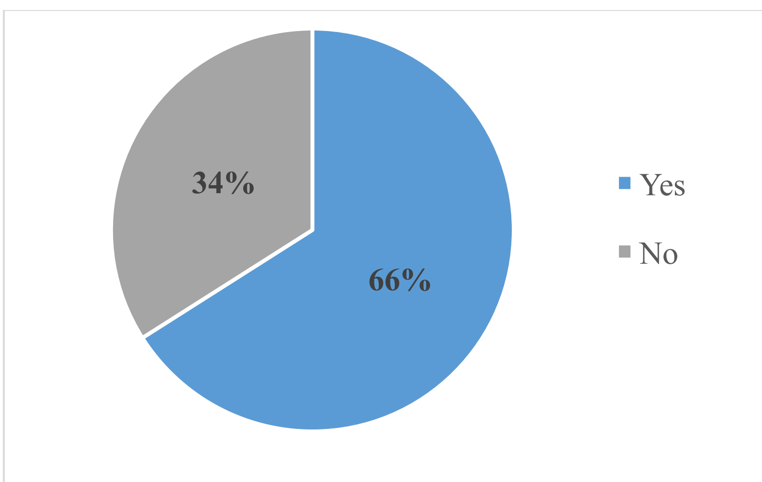

Figure 21. Previous work experience between Owner and Contractor

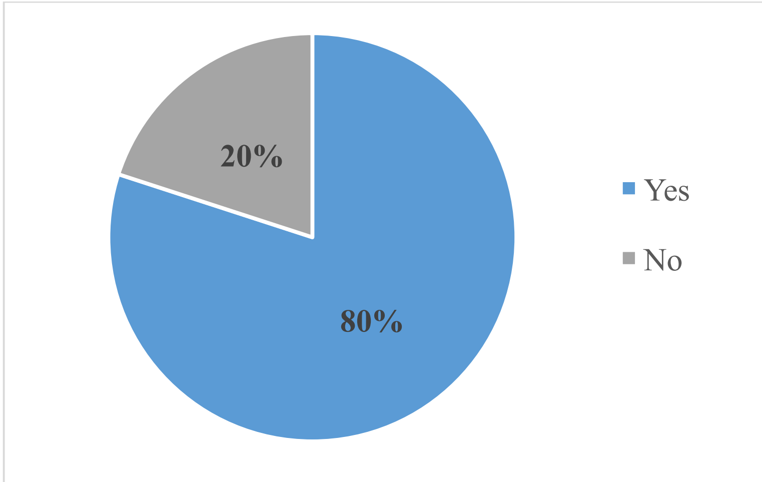

Figure 22. Owners' Experience with Type of Project

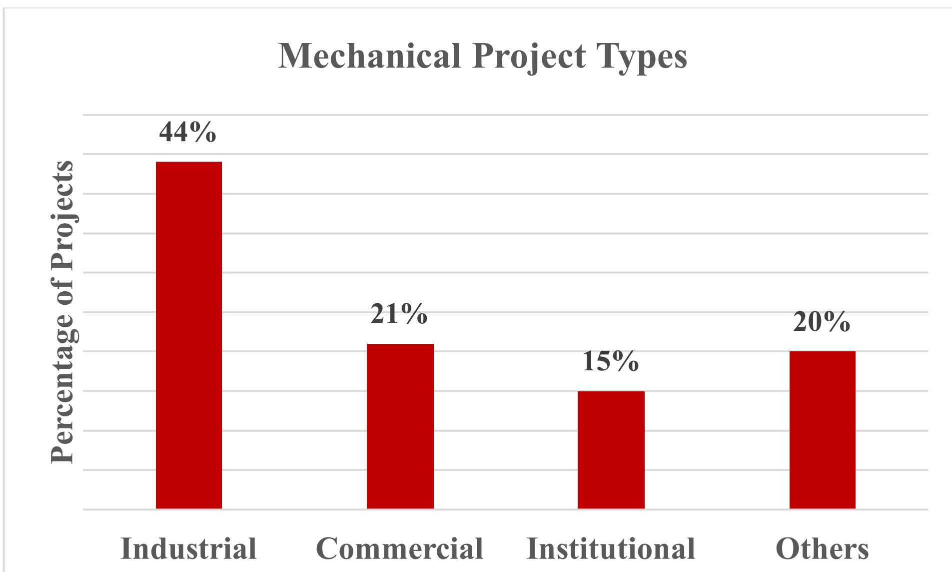

Figure 23. Distribution of Project Types for Mechanical Projects

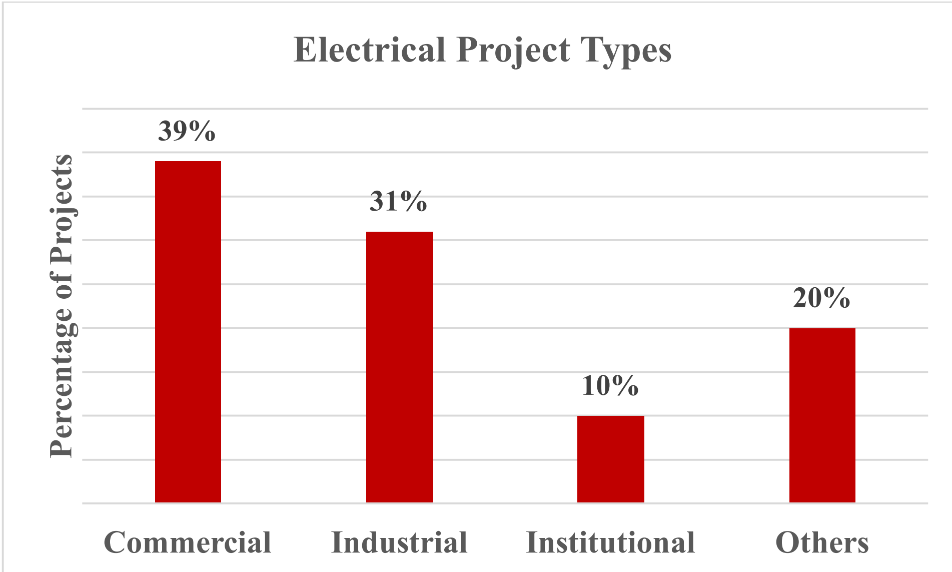

Figure 24. Distribution of Project Types for Electrical Projects

3.9. Summary

This chapter provided a detailed description of the one hundred and forty-five projects included in the data. The data encompasses two groups of projects: projects impacted by changes and projects unimpacted by changes. The distribution of mechanical and electrical projects for impacted projects is very close to the distribution for unimpacted projects. The data includes both public and private projects, located in over twenty-five States in the U.S., with sizes ranging from 2,000 work hours to more than 250,000 work hours. The dataset also incorporates a wide variety of project types, including commercial, institutional, industrial, manufacturing, power plants, residential, and maintenance projects.

Chapter 4. Productivity Factors

The first part of this chapter will highlight major features and characteristics that distinguish between impacted and unimpacted projects. Impacted projects are projects where change orders had a negative effect on labor efficiency. Some project features can determine whether a project is impacted or unimpacted by changes, such as the amount of change itself. Unimpacted projects are projects where change orders did not affect labor efficiency. In that case, either the project had no signs of being impacted by changes, or the contractor was so competent that he/she has managed to overcome the negative impact of changes orders. %Delta for impacted projects have positive values, indicating an overall project efficiency that is lower than the original expectations in the estimate. Although the survey has only targeted projects with accurate bid estimates, a tolerance of ± 5% for %Delta was considered in order to account for the contractor's estimating ability. Consequently, in the context of this thesis report, impacted projects have %Deltas higher than 5%. %Delta for unimpacted projects have negative values, indicating an overall project efficiency that is higher than the original expectations in the estimate. It should be noted that projects having %Delta less than 5% were re-classified as unimpacted.

The second part of this chapter will focus on relationships between efficiency loss and different productivity factors for impacted projects only. This part will examine trends related to productivity loss and will disclose some of the significant factors that hinder the performance of contractors and impede their project efficiency.

4.1. Impacted and Unimpacted Projects

Unimpacted projects were found to have greater performance in regard to many productivity indicators as compared to impacted projects. This first part of the chapter will examine those productivity factors/variables that were found significant in distinguishing between impacted and unimpacted projects. The Welch's two-sample t-test was used for numerical variables, while the Chi-squared test for independence was used for categorical variables.

4.1.1. Manpower Ratios

Four manpower ratios were examined for all projects: estimated average over actual average manpower (EA_A), estimated peak over actual peak manpower (EA_P), estimated peak/average over actual peal/average (EA_PA), and actual peak to actual average manpower (P/A). The Welch's two-sample t-test was used to test the null hypothesis that states that the two groups of projects (impacted and unimpacted) have equal means regarding each of the four manpower ratios. It was found that the mean of each of four manpower ratios was significantly different for impacted projects as compared to unimpacted projects.

4.1.1.1. Estimated Average over Actual Average Manpower (EA_A)

When (EA_A) is lower than 1, this indicates that the actual average manpower has exceeded the average manpower that was originally estimated for the project. Unimpacted projects have shown more desirable values for (EA_A), with a mean value of 1.04, indicating that, on average, unimpacted projects haven’t exceeded the average manpower that was originally estimated. The mean value of (EA_A) for impacted projects was found to be equal to 0.80, indicating that on average, the actual average manpower for impacted projects was higher than what was anticipated in the estimate. The Welch's two-sample t-test resulted in a significantly small p-value of 0.005229, showing strong evidence that the mean of (EA_A) for impacted projects is lower than the mean of (EA_A) for unimpacted projects.

4.1.1.2. Estimated Peak over Actual Peak Manpower (EA_P)

When (EA_P) is lower than 1, this indicates that the actual peak manpower has exceeded the peak manpower that was originally estimated for the project. Unimpacted projects have shown more desirable values for (EA_P), with a mean value of 1.07, indicating that, on average, unimpacted projects haven't exceeded the peak manpower that was originally estimated. The mean value of (EA_P) for impacted projects was found to be equal to 0.77, indicating that on average, the actual peak manpower for impacted projects was higher than what was anticipated in the estimate. The Welch's two-sample t-test resulted in a significantly small p-value of 0.006081, showing strong evidence that the mean of (EA_P) for impacted projects is lower than the mean of (EA_P) for unimpacted projects.

4.1.1.3. Estimated Peak/Average Over Actual Peak/Average Manpower (EA_PA)

Similar to the previous two manpower ratios, the ratio of (estimated peak/average) to (actual peak/average) manpower (EA_PA) was lower for impacted projects as compared to unimpacted projects. The Welch's two-sample t-test resulted in a small p-value of 0.024, showing strong evidence that the mean of (EA_PA) for impacted projects is lower than the mean of (EA_PA) for unimpacted projects.

4.1.1.4. Actual Peak Over Actual Average Manpower (P/A)

According to (Hanna 2005), when the ratio of actual peak manpower to actual average manpower in a given project exceeds 1.6, overmanning is considered to occur. In this research, the 1.6 threshold is used to distinguish between projects that experienced overmanning (P/A ratio higher that 1.6) and projects that haven't experienced overmanning (P/A ratio lower than 1.6). The Welch's two-sample t-test resulted in a small p-value of 0.025, showing strong evidence that the mean of (P/A) for impacted projects is higher than the mean of (P/A) for unimpacted projects. Table 3 summarizes the results for all manpower ratios.

Table 3. Manpower Ratios

Manpower ratio/ Values for the two project groups	Mean Impacted	Mean Unimpacted	P-value
Estimated Average over Actual Average (EA_A)	0.80	1.04	0.005
Estimated Peak over Actual Peak (EA_P)	0.77	1.07	0.006
Estimated Peak/Average Over Actual Peak/Average (EA_PA)	0.92	1.03	0.024
Actual Peak Over Actual Average (P/A)	2.16	1.94	0.025

4.1.2. Percent Change for Impacted and Unimpacted Projects

The amount of change was found to be a significant factor in distinguishing between impacted and unimpacted projects. Projects with high percent change values are most likely to be impacted by change orders. Table 4 shows that the mean value of (Percent Change) for impacted projects is equal to 30.2%, while the mean value of (Percent Change) for unimpacted projects is equal to 23.4% The Welch's two-sample t-test resulted in a p-value of 0.04933, showing evidence that the mean (Percent Change) for impacted projects is higher than the mean (Percent Change) for unimpacted projects.

Table 4. Percent Change for Impacted and Unimpacted Projects

Percent Change/ Values for the two project groups	**Mean Impacted**	**Mean Unimpacted**	**P-value**
Percent Change	30.2%	23.4%	0.04933

Figure 25 shows the distribution of projects by (Percent Change) and (Impact). It was found that 31% of impacted projects had a (Percent Change) higher than 40%, while 22% of unimpacted projects had (Percent Change) higher than 40%. In addition, 17% of impacted projects had a (Percent Change) lower than 10%, while 31% of unimpacted projects had a (Percent Change) lower than 10%.

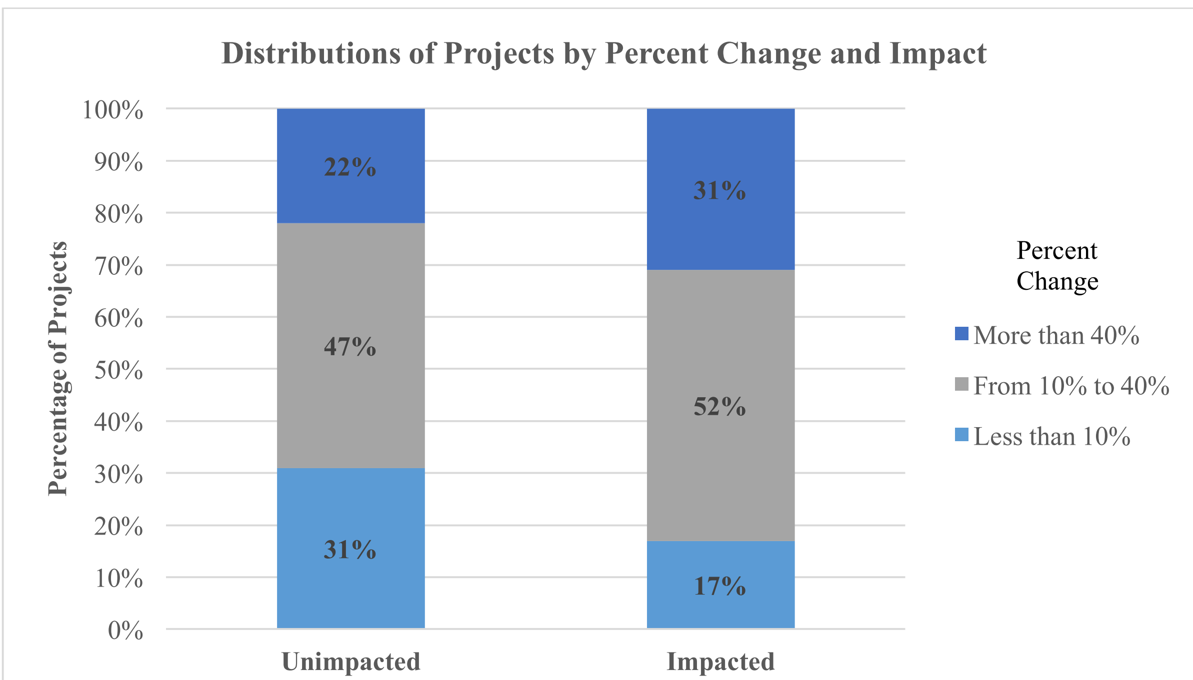

Figure 25. Distribution of Projects by Percent Change and Impact

4.1.3. Architect/Engineer (AE's) coordination prior to construction

Inadequate A/E's design coordination and review prior to project commencement can lower labor efficiency during construction. When design defects are left undetected during the early phases of a project, work interruptions are more likely to occur during construction. A/E's should carefully review their design documents and coordinate potential conflicts as early as possible in order to ensure an efficient work flow during construction and avoid rework and delays. The Chi-squared test for independence was conducted to test the null hypothesis that states that "Impact" is independent from "Architect/Engineer's coordination prior to construction." With a p-value of 0.004, the null hypothesis was rejected, indicating that whether a project is impacted or unimpacted (i.e. "Impact") highly depends on the AE's level of coordination for design issues prior to construction. Figure 26 shows that 56% of impacted projects reported inadequate AE coordination prior to construction, while 31% of unimpacted projects reported inadequate AE coordination prior to construction.

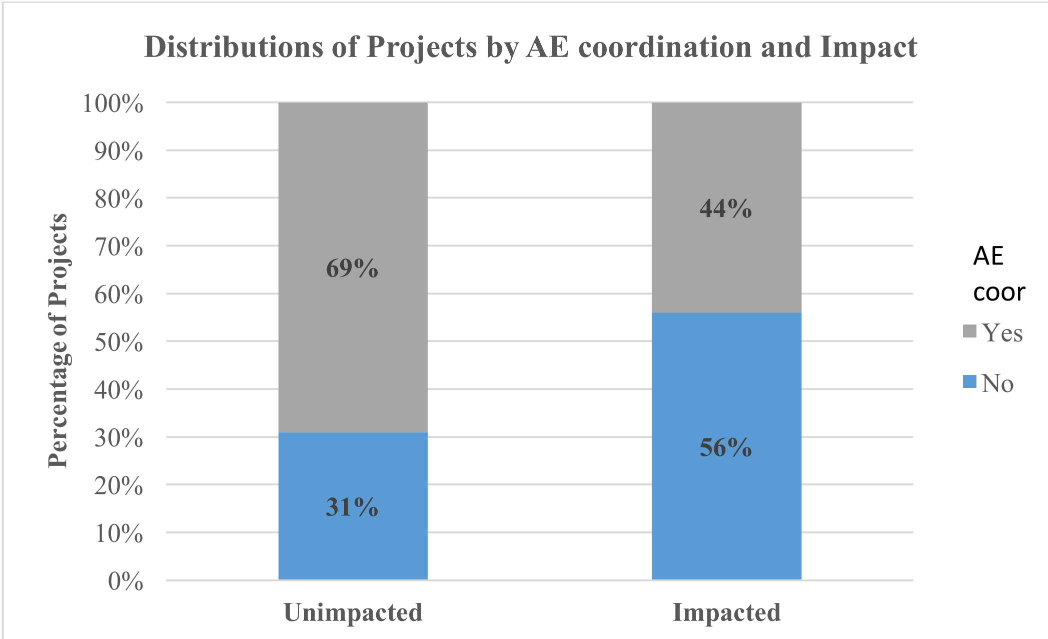

Figure 26. Distribution of Projects by AE coordination and Impact

4.1.4. Architect/Engineer (AE's) support during construction

AEs should promptly respond to the Requests for Information (RFI's) submitted by contractors. AEs should quickly resolve any design issues/ambiguities and promptly provide contractors with the necessary information throughout project execution. Inadequate AE support during construction and slow response to the contractors' clarification requests can cause project delays. The Chi-squared test for independence showed that "Impact" depends on "AE support during construction." With a p-value of 0.0018, the null hypothesis was rejected, indicating that whether a project is impacted or unimpacted (i.e. "Impact") highly depends on the AE's support in resolving design issues during construction. Figure 27 shows that 47% of impacted projects reported adequate AE support during construction, while 74% of unimpacted projects reported adequate AE support during construction.

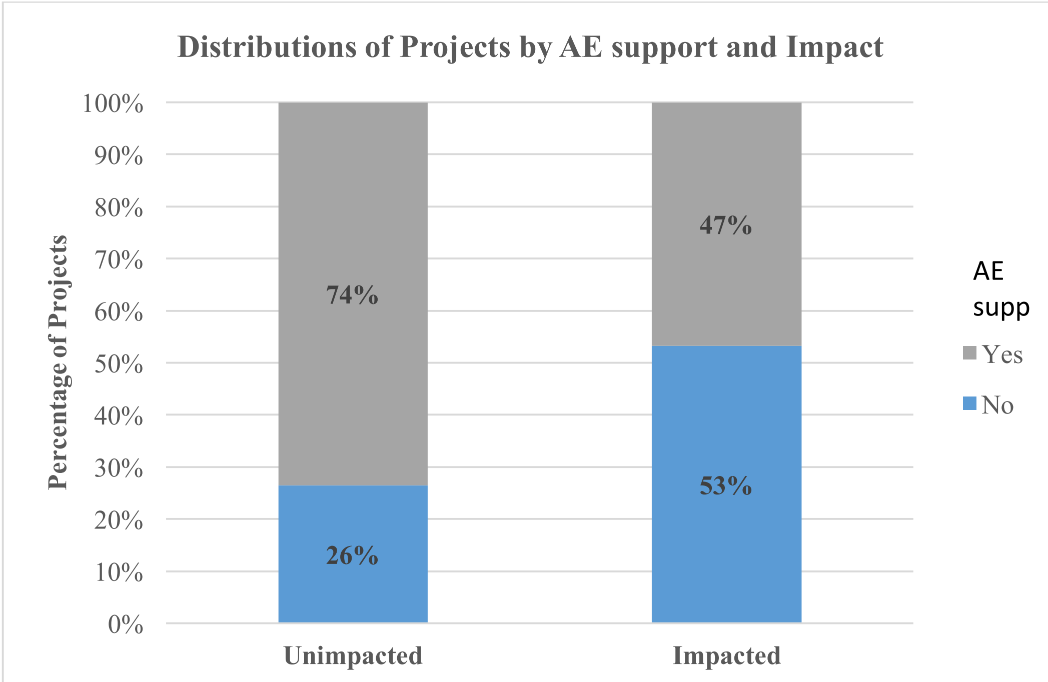

Figure 27. Distribution of Projects by AE support and Impact

4.1.5. Change Order Processing Time

In the context of this research, the change order processing time is defined on a five-level scale: (1= 1 to 7 days; 2= 8 to 14 days; 3= 15 to 21 days; 4= 22 to 28 days; and 5= Greater than 28 days). When the time between the initiation of the change order and the owner's approval of the change order increases, it is likely that the contractor has already executed and finished the change order work before even getting the owner's official approval for the change order. Therefore, longer change order processing time is disfavored by contractors as they often fail to get full compensation for the changed work. The Chi-squared test for independence was conducted to test the null hypothesis that states that "Impact" is independent from "Change Order Processing Time." With a p-value of 0.002, the null hypothesis was rejected, indicating that whether a project is impacted or unimpacted (i.e. "Impact") highly depends on the change order processing time. The average change order processing time for impacted projects is 4.21, while the average change order processing time for unimpacted projects is 3.26. Figure 28 shows that 25% of unimpacted projects have a change order processing time less than 7 days, while only 6% of impacted projects have a change order processing time less than 7 days.

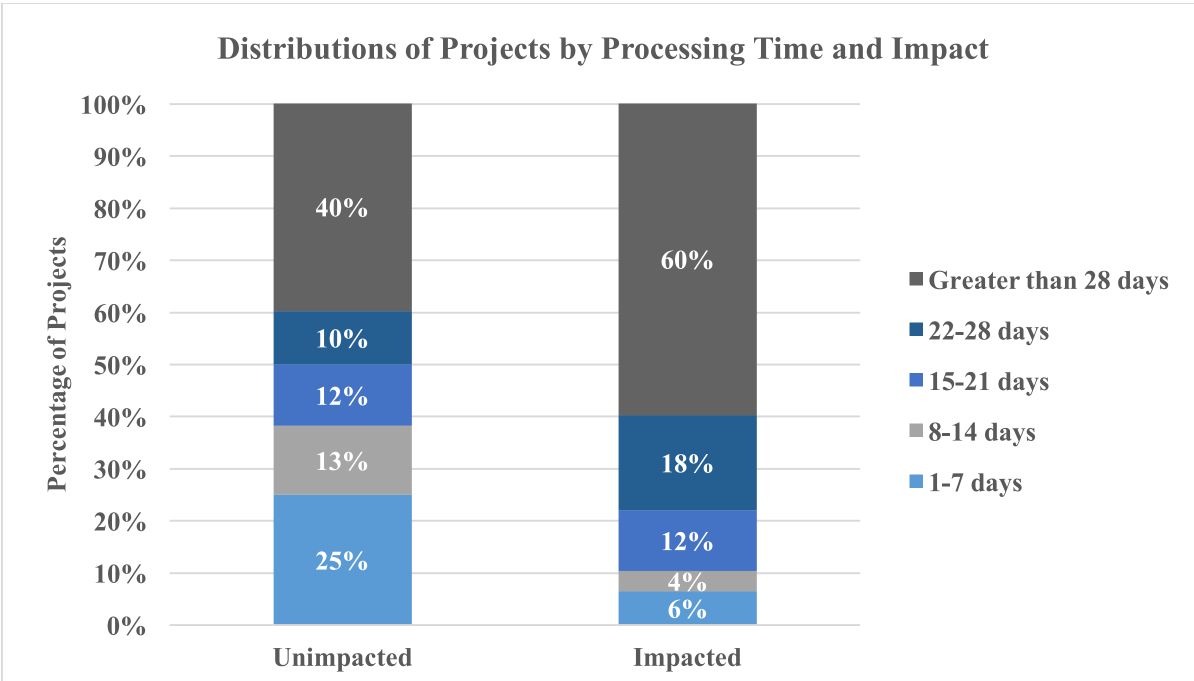

Figure 28. Distribution of Projects by Processing Time and Impact

4.1.6. Manpower shortage during peak

Labor availability is crucial for the successful performance of MEP contractors, as they highly depend on the labor component. Manpower shortages can impose a substantial risk on project performance as they may reduce the work pace and cause delays. Figure 29 shows that almost half of impacted projects reported manpower shortages during peak, while 29% of unimpacted projects reported manpower shortages during peak. The Chi-square test for independence was conducted to test the null hypothesis that states that “Impact” is independent from “Manpower shortage during peak.” With a p-value of 0.028, the null hypothesis was rejected, indicating that whether a project is impacted or unimpacted (i.e. "Impact") highly depends on manpower shortage during project peak.

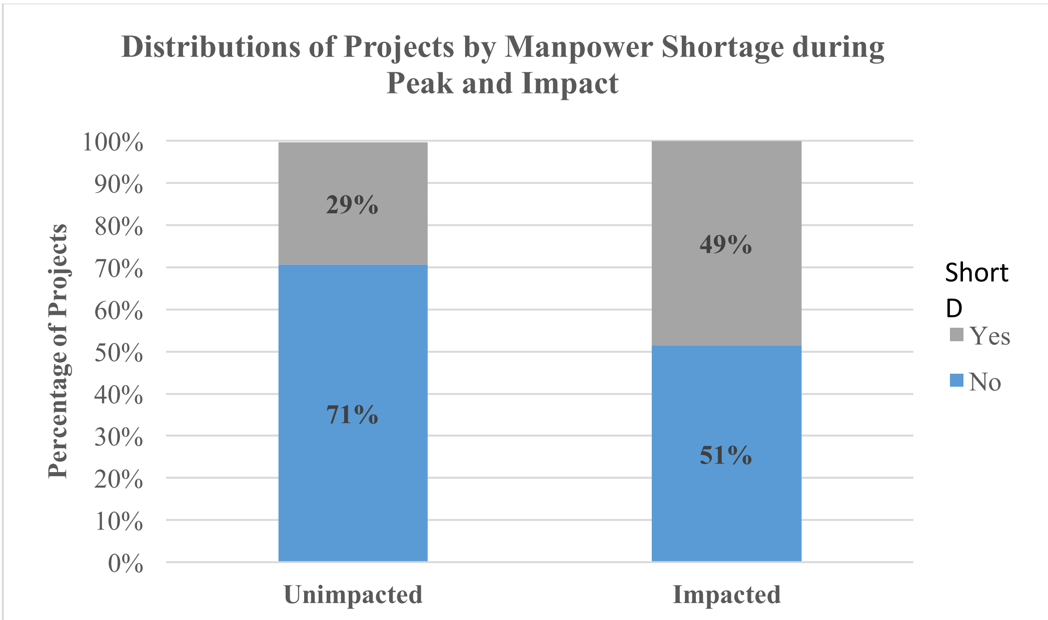

Figure 29. Distribution of Projects by Manpower Shortage at Peak and Impact

4.1.7. Overtime for Change Orders

As was stated in the literature review, overtime is a common acceleration technique that may be used by contractors to overcome the delays caused by change orders. However, overtime introduces many issues such as workers' fatigue, low labor morale and higher project expenses. Figure 30 shows that 73% of impacted projects have used overtime to complete change order work, while 59% of unimpacted projects have used overtime to complete change order work. Although the Chi-squared test indicated a relatively high p-value of 0.11, it is obvious from Figure 30 that overtime was more frequently used in impacted projects than in unimpacted projects.

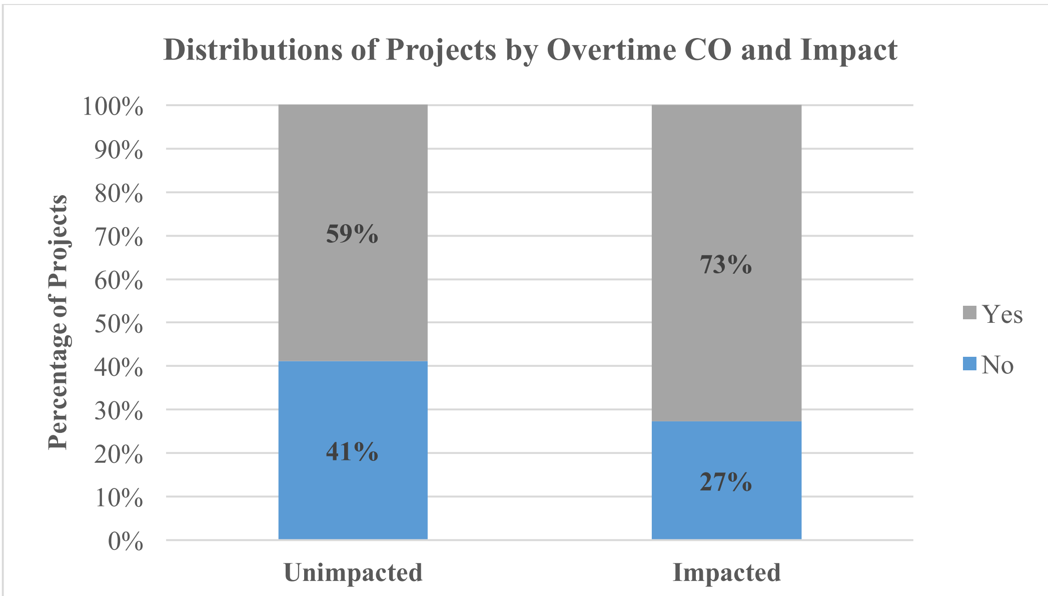

Figure 30. Distribution of Projects by Overtime and Impact

4.1.8. Shiftwork for Change Orders

As was previously discussed in the literature review, shiftwork is often used by contractors to overcome the delays that may result from change order work. Shiftwork often requires additional supervision staff and quality and safety control personnel. Excessive use of shiftwork results in inconsistent work procedures among different shifts and inefficient communication between different crews, which may lower productivity. Figure 31 shows that 22% of impacted projects have used shifts to complete change order work, while only 9% of unimpacted projects have used shifts to complete change order work. The Chi-squared test indicated a low p-value of 0.05, which means that “Impact” and “Whether overtime was used to complete change order work” highly depend on one another.

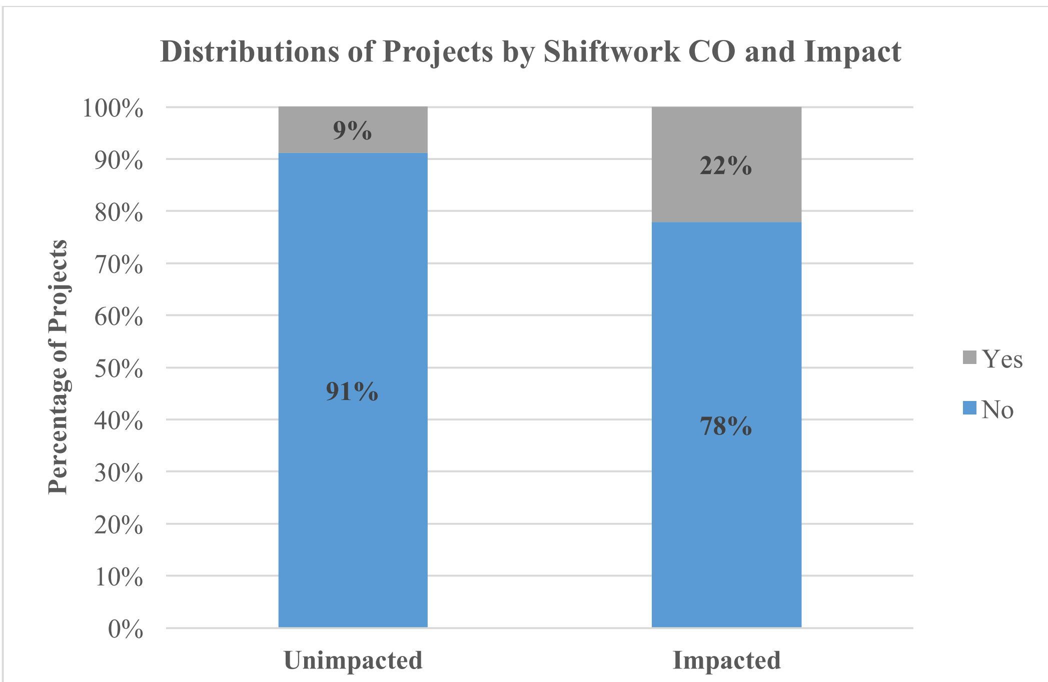

Figure 31. Distribution of Projects by Shiftwork and Impact

4.1.9. Absenteeism and Turnover

In the context of this study, turnover is defined as the percent of craftsmen hired to replace those who have left of the total number of craftsmen employed on the project. Absenteeism is defined as the percent of craftsmen who failed to appear for work of the total number of craftsmen employed on the project. Absenteeism and turnover rates are measured on a four-level scale: (1= 0-5%; 2= 6-10%; 3= 11-20%; and 4= Greater than 20%). Figures 32 and 33 show the distribution of absenteeism and turnover rates for both impacted and unimpacted projects. The mean absenteeism and turnover rates were higher for impacted projects as compared to unimpacted projects. More specifically, the mean percent absenteeism for impacted and unimpacted projects are 1.61 and 1.30, respectively. The mean percent turnover for impacted and unimpacted projects are 1.74 and 1.50, respectively. In general, it is not very common for a project to have an absenteeism rate of more than 20%. However, when absenteeism reaches this level, its effects on project performance are detrimental.

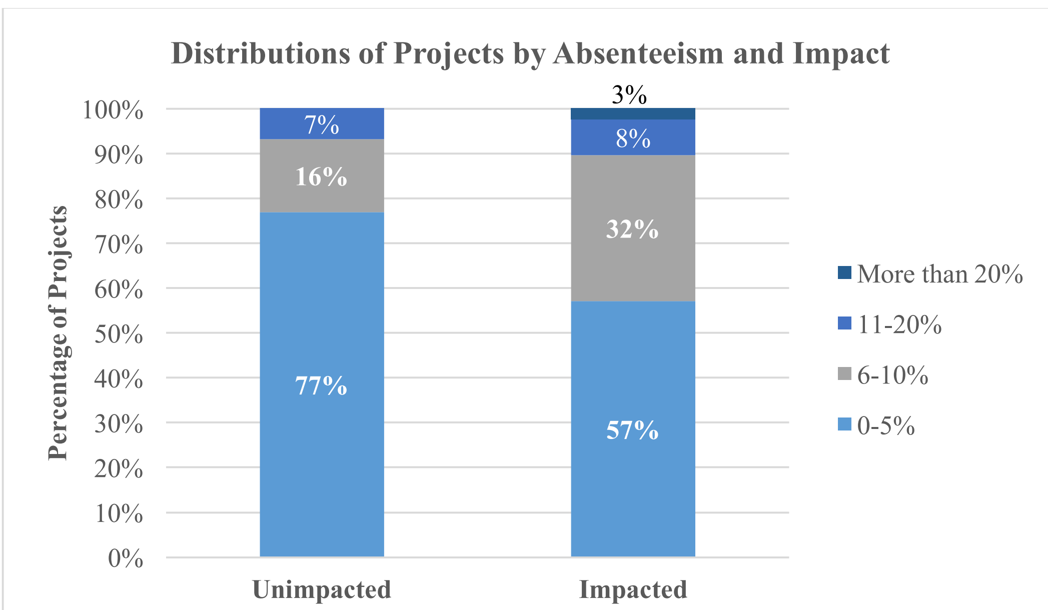

Figure 32: Distribution of Projects by Absenteeism and Impact

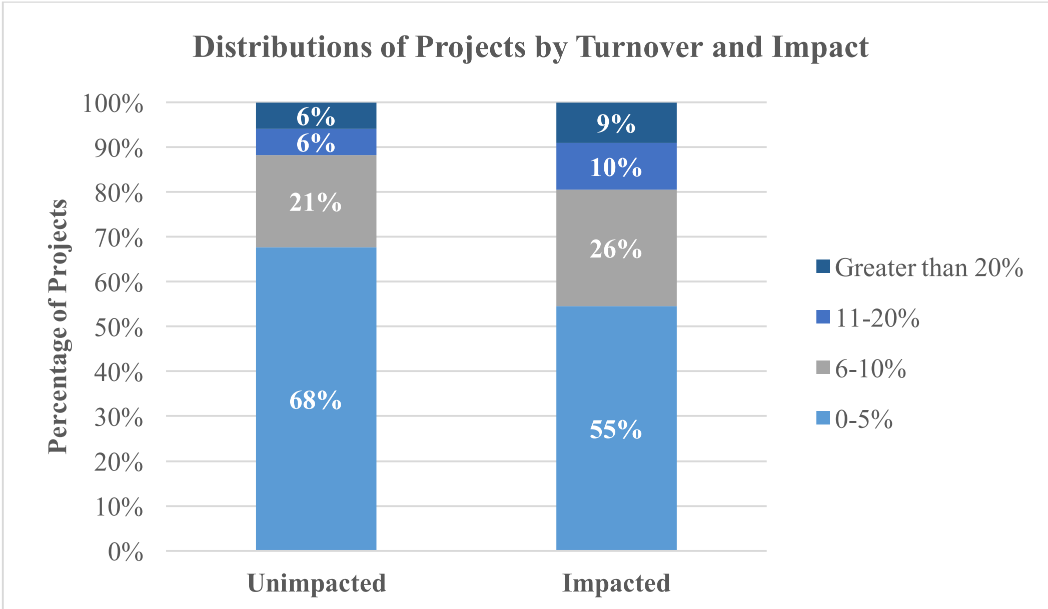

Figure 33: Distribution of Projects by Turnover and Impact

4.2. Efficiency loss and Productivity factors

In this part of the chapter, our focus is directed to impacted projects only. The amount of efficiency loss, expressed by %Delta, is determined based on changing the levels/values of various productivity factors. Productivity factors that showed significant relationships with %Delta are reported in this section. Those productivity factors can serve as productivity indicators, alerting contractors for potential efficiency loss throughout project execution.

4.2.1. Productivity Tracking

Monitoring the progress of work is among the essential means for pinpointing jobsite issues and taking corrective actions whenever required throughout project execution. Different management techniques can be used to monitor the progress on a project. Those include tracking percentage complete, tracking the installed quantities, and computing actual and earned work hours. Figure 34 shows the effect of productivity tracking on %Delta. The mean %Delta for projects where productivity was tracked (*Productivity =1*) equals to 23%, while the mean %Delta for projects where productivity was not tracked (*Productivity =0*) equals to 32%. Projects where productivity has been tracked showed better performance as compared to projects where productivity was not tracked.

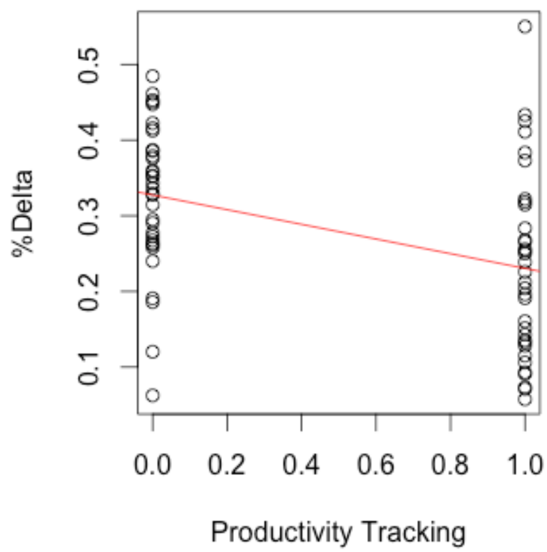

Figure 34. %Delta and Productivity Tracking

4.2.2. Percent Time Spent by Project Manager on a Project

When a project manager handles more than one project at the same time, it is harder to equally monitor all projects. The absence of a project manager on a project even for a short period of time can result in lower efficiency. Consequently, %Delta is expected to decrease when the percent of time spent by the a (PM) on a project increases. Figure 35 shows the relationship between %Delta and the percent of time spent by the Project Manager (PM) on a project. The two variables have a negative correlation coefficient of -0.32. A project manager focusing on one project at a time can continuously ensure efficient communication between project participants. Project managers should also provide adequate leadership and should empower their employees throughout project execution. In addition to that, responsiveness and quick decision making are crucial for project success.

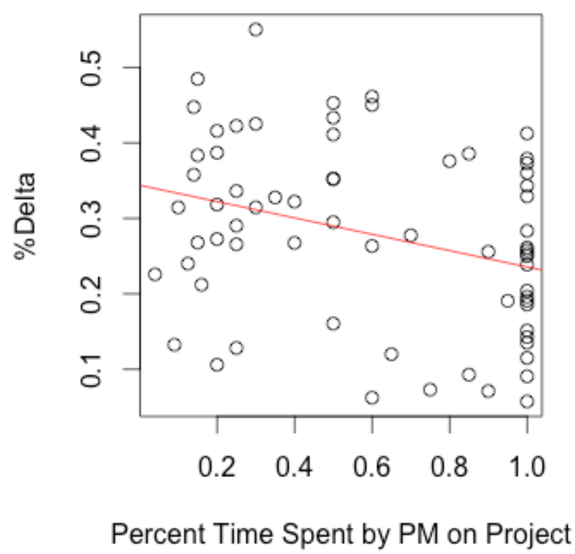

Figure 35. %Delta and Percent of Time Spent by PM on a Project

4.2.3. Architect/Engineer (A/E's) Coordination and Support

Figure 36 shows the effect of poor A/E coordination prior to construction on %Delta. The mean %Delta for projects with inadequate A/E coordination equals to 30%, while the mean %Delta for projects with adequate A/E coordination equals to 24%. In addition, Figure 37 shows the effect of inadequate A/E support during construction on %Delta. The mean %Delta for projects with inadequate A/E support equals to 31%, while the mean %Delta for projects with adequate A/E support equals to 23%. It can therefore be concluded that adequate A/E support and coordination of design issues can reduce the amount of productivity loss.

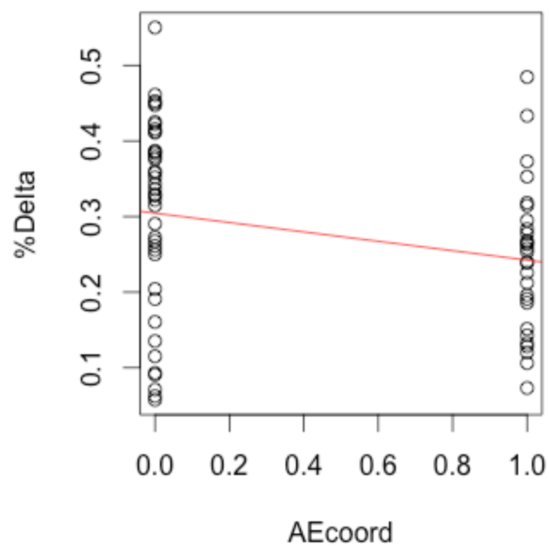

Figure 36. %Delta and AE coordination prior to construction

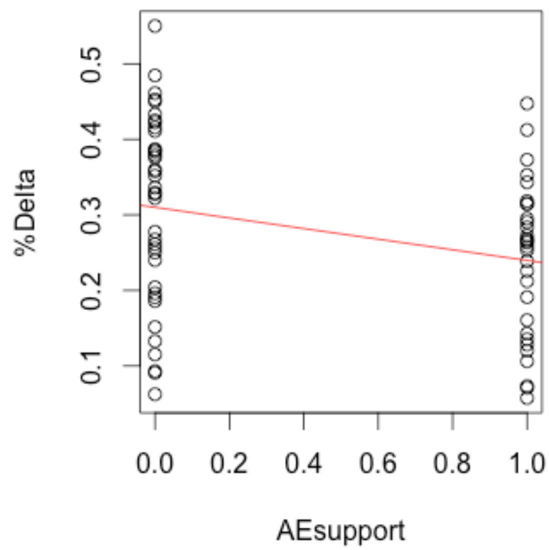

Figure 37. %Delta and AE support during construction

4.2.4. Percent Owner Initiated Change Orders

In the context of this research, the percent of change orders initiated by the owner is defined as the number of owner initiated change items divided by the total number of change items. Figure 38 shows the linear relationship between %Delta and percent of change orders initiated by the Owner *(%OwnerInitiatedCO).* The two variables have a negative correlation coefficient of -0.54. When the percent of change orders initiated by the owner increases, contractors are more likely to be receive full compensation for the additional work hours and resources used for the changed work. Consequently, %Delta is expected to decrease when most of the change orders are directed by the owner.

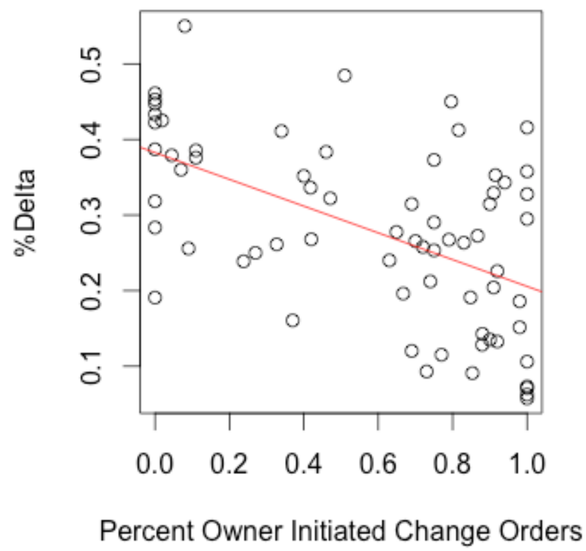

Figure 38. %Delta and Percent Owner Initiated Change Orders

4.2.5. Overmanning

In this research, overmanning is computed based on the actual jobsite manpower ratios. Overmanning occurs when the ratio of actual peak manpower to actual average manpower exceeds 1.6 (Hanna 2005). Figure 39 shows the effect of overmanning on %Delta. Projects experiencing overmanning (*Overmanning =1*) are more likely to face site congestions and dilution of supervision. Therefore, they tend to have higher %Delta as compared to projects not experiencing overmanning. The mean %Delta for projects experiencing overmanning is 29%, while the mean %Delta for projects not experiencing overmanning is 23%.

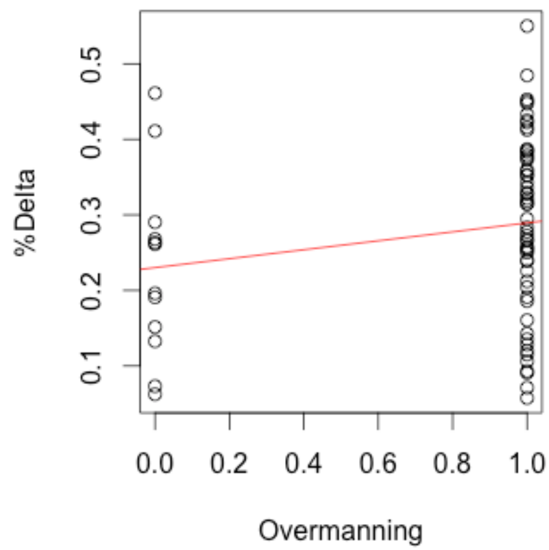

Figure 39. %Delta and Overmanning

4.2.6. Manpower Shortage at Peak

When electrical and mechanical contractors face labor shortages during construction or at project peak, productivity will certainly decrease. Figure 40 shows the effect of manpower shortage at project peak on labor efficiency. Figure 40 shows that the mean %Delta equals to 31% for projects that reported manpower shortages during construction, and the mean %Delta equals to 24% for projects with adequate manpower availability. Consequently, the amount of efficiency loss is expected to be higher for projects experiencing situations of labor shortages.

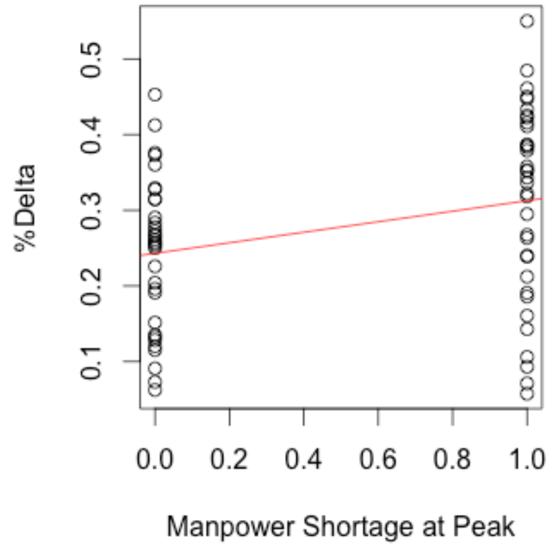

Figure 40. %Delta and Manpower Shortage at Peak

4.2.7. Turnover and Absenteeism

Change orders can cause more field rework, which lowers labor morale and increases absence and turnover rates. Excessive absenteeism and turnover usually result in lower workforce production, increased costs for owners and contractors, and reduced labor efficiency. Figures 41 and 42 show the effect of turnover and absenteeism on labor productivity, respectively. When the level of turnover is 4 (i.e. greater than 20%), %Delta increases. Similarly, when the level of absenteeism is 4 (i.e. greater than 20%), which rarely occurs, %Delta significantly increases. As a result, the amount of efficiency loss is expected to increase for projects experiencing high turnover and absenteeism rates.

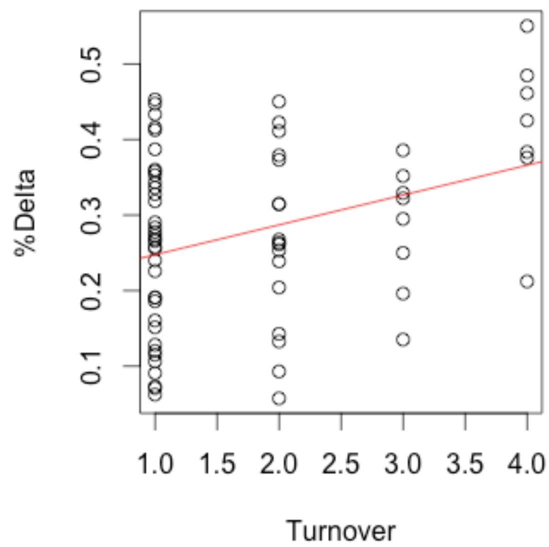

Figure 41. %Delta and Turnover

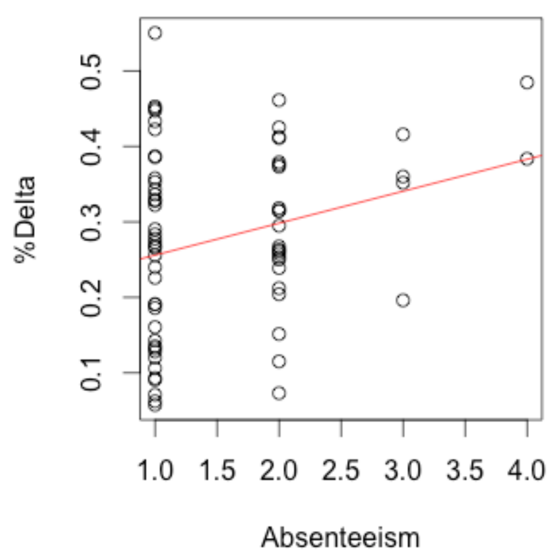

Figure 42. %Delta and Absenteeism

4.3. Summary of Significant Productivity Factors

Tables 5 and 6 summarize the results of the t-tests, Chi-squared tests, and simple linear relationships that were performed in this chapter. Some variables were found significant in distinguishing between impacted and unimpacted projects, while other variables showed significant relationships with the amount of productivity loss. This chapter provides contractors with a set of productivity indicators that should be monitored throughout project execution in order to avoid potential efficiency losses and achieve higher productivity levels.

Table 5. Summary of Factors Distinguishing between Impacted and Unimpacted Projects

Factor	**Interpretation:** **… is more likely to be impacted**	**P-value**
Percent Change	A project with higher percent change …	0.049
Estimated Average over Actual Average (EA_A)	A project with lower Estimated Average over Actual Average Manpower …	0.005
Estimated Peak over Actual Peak (EA_P)	A project with lower Estimated Peak over Actual Peak Manpower …	0.006
Estimated Peak/Average Over Actual Peak/Average (EA_PA)	A project with lower Estimated Peak/Average Over Actual Peak/Average Manpower ratio…	0.024
Actual Peak Over Actual Average (P/A)	A project with higher Actual Peak Over Actual Average ratio…	0.025
Overmanning	A project with Actual Peak to Actual Average Manpower ratio higher than 1.6…	0.000
AE coordination prior to construction	A project with inadequate AE coordination prior to construction…	0.004
AE support during construction	A project with inadequate AE support during construction …	0.0018
Processing Time	A project with longer change order processing time …	0.002
Manpower Shortage during Peak	A project experiencing manpower shortage during peak …	0.028
Absenteeism	A project with higher absenteeism rates …	0.02
Turnover	A project with higher turnover rates …	0.09
Overtime	A project where overtime was used to complete change order work …	0.11
Shiftwork	A project where shiftwork was used to complete change order work …	0.05

Table 6. Summary of Simple Linear Models with %Delta

Equation: **%Delta = …**	**Interpretation:** **On Average …**	**P-value**
0.38258 -0.17707 Percent Owner Initiated Change Orders	A project with higher Percent Owner Initiated Change Orders has lower %Delta	0.00
0.34387 -0.10829 Percent of Time spent by PM	A project with higher percent of time spent by PM has lower %Delta	0.008
0.30450 -0.06217 AE coordination prior to construction	A project with adequate AE coordination prior to construction has lower %Delta	0.03
0.30991 -0.07030 AE support during construction	A project with adequate AE support during construction has lower %Delta	0.01
0.32740 -0.09700 Productivity Tracking	A project where productivity is tracked has lower %Delta	0.0005
0.23015 +0.05920 Overmanning	A project with Actual Peak Over Actual Average ratio higher than 1.6 has higher %Delta	0.08
0.21357 +0.04231 Absenteeism	A project with higher absenteeism rates has higher %Delta	0.03
0.20816 +0.03943 Turnover	A project with higher turnover rates has higher %Delta	0.005
0.24311 +0.06953 Manpower shortage at peak	A project experiencing labor shortages during construction has higher %Delta	0.01

Chapter 5. Predicting Percent Delta

This chapter is the most critical amongst all chapters of this thesis report. The ultimate goal of this chapter is to build a linear regression model in order to quantify the amount of productivity loss associated with change orders (%Delta) for impacted projects. Throughout the model building process, candidate models are diagnosed in order to check their validity and whether they violate any of the linear regression model assumptions. Several model selection criteria were used in this chapter in order to derive the best candidate models and avoid overfitting. The model diagnostics involve a lot of statistical terms, tests, and measures, that quantitatively assess the influence of deleting each project/data point on the candidate regression functions. Appendix A includes a brief description of the statistical terms and measures used in this chapter. Furthermore, this chapter statistically addresses some of the major challenges that analysts face when building linear regression models, such as hidden extrapolations, multicollinearity, and non-constancy of variance of the error terms.

5.1. Stepwise Regression

One of the most commonly used model selection methods is stepwise regression. Stepwise regression develops a sequence of regression models at different steps. In each step, a certain criterion is used to determine whether to add or drop a variable. There are two types of stepwise regression: forward stepwise regression and backward stepwise regression. Forward stepwise regression starts with no variables, and adds a variable at each step, while checking whether a previously added variable should be dropped. On the other hand, backward stepwise regression starts with all variables, and drops a variable at each step, while checking whether a previously dropped variable should be added. Backward elimination and forward selection differ from backward stepwise regression and forward stepwise regression. In backward elimination, once a variable is dropped from the model, there is no chance that it can be added again. Similarly, in forward selection, once a variable is added to the model, there is no chance that it can be dropped later in the process. In this chapter, the variables there were included throughout the model selection process are: percent change, overmanning, change order processing time, overtime, shiftwork, percent owner initiated change orders, percent of time spent by the project manager on the project, manpower shortage at peak, productivity tracking, turnover, and absenteeism.

5.1.1. Criteria used for Stepwise Regression

Among the criteria used to compare different models within stepwise regression procedures or to test whether to add or drop a variable at each step are the “Akaike’s Information Criterion (AIC)” and the “Bayesian Information Criterion (BIC).” AIC and BIC are criteria for penalizing models with larger number of predictors. When the sample size is greater than 8, the penalty for BIC is larger than AIC. Therefore, BIC tends to favor more parsimonious models. Parsimonious models are models that accomplish a desired level of explanation with as few predictors as possible. The absolute values of AIC and BIC are meaningless, but the relative values between models are meaningful. A preferred model has a minimum AIC or BIC.

5.1.2. Testing whether to Include Interaction Terms

Our first task is to determine which stepwise method has the highest out-of-sample prediction accuracy. Moreover, we will determine whether including interaction terms would provide better predictions for %Delta. Among the important interaction terms that were tested are the interaction effect between change order processing time and percent change, the interaction effect between percent change and productivity tracking, and the interaction effect between percent owner initiated change orders and productivity tracking. Interactions can result from a situation where the effect of a first variable on %Delta is modified by changing the level of a second variable. One should not confuse between correlation and interaction. If two variables are uncorrelated, they can still interact and affect %Delta.

To decide whether we should include interaction terms and to determine the most appropriate stepwise method, cross-validation was performed for each of the four stepwise regression methods listed in Table 7. It should be noted that for each stepwise method, cross-validation was performed twice: once when interaction terms are not included, and once when interaction terms are included. Therefore, there were a total of eight cross-validation runs, each run pertaining to a different variable selection procedure.

For each of the eight runs shown in Table 7, the data was randomly divided into 5 parts, where each part served as a testing set once, and the remaining 4 parts were used as a training set to develop a training model. For each of the eight runs, 5 training models and 5 testing sets were developed. The 5 training models were generated based on the stepwise procedure used in the run. Within each run, the 5 training models had different variables in order to allow for the stepwise method to fit each of the 5 folds as better as possible. One should not confuse between cross-validating a method and cross-validating a model. In this section, we are cross-validating different methods. The stepwise method with the smallest out-of-sample prediction errors will be applied to the entire data to derive the best regression model.

Table 7 also shows the “Prediction Mean Square Error” or Prediction MSE for each of the eight runs. The run with the smallest Prediction MSE indicates the best stepwise method to be used as well as whether to include interaction terms or not. The Prediction MSE of each run is calculated using the following two steps:

Step one: Compute the sum of the squared differences between the actual %Delta values and the %Delta values predicted from the different training models in the run.

Step two: Divide the quantity in step 1 by the number of observations (68).

Table 7. Cross-Validating Different Variable Selection Procedures

Method/Measurement	Without Interaction Terms			With Interaction Terms		
	Sum of the Squared Prediction Errors	**Prediction MSE**	**Procedure Ranking**	**Sum of the Squared Prediction Errors**	**Prediction MSE**	**Procedure Ranking**
Forward Stepwise Regression - AIC	0.460	0.00676	**#1**	0.477	0.00701	**#3**
Forward Stepwise Regression – BIC	0.467	0.00686	**#2**	0.599	0.00880	**#8**
Backward Stepwise Regression – AIC	0.530	0.00779	**#5**	0.510	0.00750	**#4**
Backward Stepwise Regression - BIC	0.577	0.00848	**#6**	0.584	0.00858	**#7**

5.1.3. The Top-Ranked Variable Selection Procedure

Table 7 showed the ranking of each variable selection procedure. The top two procedures consisted of cases where interaction terms were excluded. Consequently, it was found that including interactions would reduce the prediction accuracy of the regression models. The "Forward Stepwise Regression - AIC" without interactions was ranked first; it had the smallest Prediction MSE (0.00676). The "Forward Stepwise Regression - BIC" without interactions was ranked second; it had the second smallest Prediction MSE (0.00686). As was expected, the AIC criterion had a slightly higher out-of-sample prediction accuracy as compared to the BIC criterion. The AIC criterion penalizes the number of parameters less strongly as compared to the BIC criterion. Therefore, AIC allows for more predictors to be included in the model, while BIC tends to favor more parsimonious models or models with less predictors. The BIC criterion searches for the true model that delivers the same amount of explanation for %Delta, while using less predictors.

5.1.4. Deriving Model 1 Using the Top-Ranked Stepwise Regression Procedure

According to Table 7, "Forward Stepwise Regression - AIC" without interactions was found to be the best method for model selection. The method resulted in the following multiple regression model:

Model 1:

%Delta = 0.33466 – 0.14550 %OwnerInitiatedCO – 0.08564 Productivity + 0.02182 Turnover – 0.08426 PM%TimeOnProject + 0.05611 Overmanning + 0.02426 Absenteeism

Model 1 has a coefficient of multiple determination (R^2) of 60.27%, a fairly high value for the type of data analyzed. The R^2 adjusted for the model is 56.36%, which didn't vary much from R^2. A problem with R^2 is that it never decreases as the number of predictors increases. Therefore, the R^2 adjusted is an alternative measure that takes into account the number of predictors through the degrees of freedom associated with the error sum of squares (SSE) and the total sum of squares (SSTO). A brief description of SSE and SSTO can be found in Appendix A. Unlike R^2, R^2 adjusted can indeed decrease while the number of predictors increases. Therefore, R^2 adjusted is a more reliable measure as compared to R^2. The null hypothesis that states that all regression parameters of Model 1 are equal to zero was rejected with a p-value of 1.152e-10.

It is worth mentioning that the error sum of squares (SSE) for Model 1 is equal to 0.380. The SSE for the model was expected to be lower than the sum of squared prediction errors indicated in Table 6 for "Forward Stepwise Regression without Interaction Terms – AIC." This is due to the fact that in Table 6, the data was split into 5 folds, resulting in a sum of the squared prediction errors (0.460) that is higher than the actual SSE (0.380) of the model fit from the entire sample size.

Table 8 illustrates a detailed description of the variables included in the model. The model should be only applied within the range specified for each variable, as shown in Table 9. Table 10 shows the model summary, including the coefficient of each predictor and its associated p-value. For Model 1, all p-values are considerably less than 0.1, except for absenteeism.

Table 11 shows the ANOVA table for Model 1. The ANOVA table shows the sequential sum of squares, which denote the reduction in the error sum of squares or the increase in the regression sum of squares when one or more variables are added to the model. As shown in

Table 11, a reduction in the error sum of squares of 0.28145 is attributable to "%OwnerInitiatedCO." Given that "%OwnerInitiatedCO" is already included in the model, an extra reduction in the error sum of squares of 0.12987 occurs when "Productivity" is added to the model. Therefore, "%OwnerInitiatedCO" and "Productivity" achieved a combined overall reduction of 0.41132 in the error sum of squares. By summing up all 6 sequential sum of squares that correspond to the 6 predictors included in Model 1, the regression sum of squares SSR is obtained. SSR denotes how much variation in %Delta Model 1 has explained. The leftovers are what Model 1 was unable to explain, denoted by the error/residual sum of squares SSE. The total sum of squares SSTO is the total variation in %Delta, which is equivalent to the sum of SSE and SSR. The SSR, SSE, SSTO for Model 1 were found to be equal to 0.57734, 0.38063, and 0.95797, respectively. It should be noted that R^2 for Model 1 can be obtained by dividing SSR by SSTO. The higher the ratio of SSR to SSTO, the more the amount of variation in the response a model can explain. A brief description of these statistical terminologies are included in Appendix A.

Table 8. Definitions of Variables included in Model 1

Predictor name	Definition of variable
%OwnerInitiatedCO	-Numerical (in Decimals): The number of owner initiated change items divided by the total number of change items.
Productivity	-Binary: 1 = The contractor was tracking productivity 0 = The contractor was not tracking productivity
Turnover	The percent of craftsmen hired to replace those who have left of the number of craftsmen employed on the project. -Ordinal: 1= 0 to 5% 2= 6 to10% 3 = 11 to 20% 4 = Greater than 20%
PM%TimeOnProject	-Numerical (in Decimals): Percent of time by the Project Manager on the project
Overmanning	-Binary: 1 = Ratio of Actual Peak to Actual Average Manpower > 1.6 0 = Ratio of Actual Peak to Actual Average Manpower < 1.6
Absenteeism	The percent of craftsmen who failed to appear for work of the number of craftsmen employed on the project. -Ordinal: 1= 0 to 5% 2= 6 to10% 3 = 11 to 20% 4 = Greater than 20%

Table 9. Description of Predictors for Model 1

Predictor name	Min.	Max.	Mean	Median	Standard Deviation
%OwnerInitiatedCO	0	1	0.58	0.72	0.366
Productivity	0	1	-	0	0.503
Turnover	1	4	1.79	1	1.015
PM%TimeOnProject	0.04	1	0.59	0.60	0.349
Overmanning	0	1	-	1	0.384
Absenteeism	1	4	1.54	1	0.741

Table 10. Model 1 Summary

Predictor name	Coefficient	P-value
Intercept	0.33466	3.87e-11 ***
%OwnerInitiatedCO	-0.14550	1.40e-06 ***
Productivity	-0.08564	4.56e-05 ***
Turnover	0.02182	0.05429.
PM%TimeOnProject	-0.08426	0.00363 **
Overmanning	0.05611	0.03010 *
Absenteeism	0.02426	0.10747

Table 11. ANOVA Table for Model 1

Sample size (n) = 68 Number of parameters (p) = 7					
Source of Variation	**SS**	**df**	**MS**	**F-statistic**	**P-value**
Regression	0.57734	p-1 = 6	0.09622	15.42	1.152e-10
%OwnerInitiatedCO	0.28145	1	0.28145		
Productivity given (%OwnerInitiatedCO)	0.12987	1	0.12987		
Turnover given (%OwnerInitiatedCO, Productivity)	0.06551	1	0.06551		
PM%TimeOnProject given (%OwnerInitiatedCO, Productivity, Turnover)	0.05705	1	0.05705		
Overmanning given (%OwnerInitiatedCO, Productivity, Turnover, PM%TimeOnProject)	0.02681	1	0.02681		
Absentee given (Overmanning, %OwnerInitiatedCO, Productivity, Turnover, PM%TimeOnProject)	0.01665	1	0.01665		
Error	0.38063	n-p = 61	0.00623		
Total	0.95797	n-1 = 67	0.01429		

5.2. Best Subsets Algorithms

Unlike stepwise regression, "Best Subsets Algorithms" are computer-search procedures for variable selection that reduce the number of all possible regression models into a smaller group of "best possible models." The algorithm used in this thesis report specifies the best 3 models for each possible number of parameters. For example, when fixing the number of parameters to 2, the best three models including the intercept and one predictor are reported. Similarly, when fixing the number of parameters to 3, the best three models including the intercept and two predictors are reported, and so on. The resulting models pertaining to different number of parameters are compared to one another simultaneously, and the optimal model can be revealed based on different criteria, including R^2, R^2 adjusted, SSE, MSE, Mallows' C_p, and the AIC/BIC. The model with a relatively high R^2 and R^2 adjusted, and a relatively low BIC, SSE, MSE, and Mallows' C_p is chosen for final consideration. A desirable Mallows' C_p value should be equal or slightly lower than the number of parameters. Appendix A includes a brief description of these statistical terms.

It is important to mention that the best subsets algorithm returns separate best models for all parameter sizes; it doesn't make any difference whether the AIC or BIC is used. On the contrary, stepwise procedures identify one single best model and depend on whether the AIC or BIC is used, as was shown in section 5.1 of this chapter.

5.2.1. Determining the Optimal Model

In this section, our goal is to select the optimal model using the Best Subsets Algorithm. Table 12 and Figures 43 through 45 show the results of the Best Subset Algorithm. By examining Table 12 and Figure 44, it can be noticed that model number 13 has a BIC value of (-34.53), which is the lowest amongst all best subset regression models. Although model number 27 has the highest R^2 amongst all models, R^2 adjusted for model number 13 and model number 27 are very close to one another; they are equal to 55.18% and 56.16%, respectively. R^2 adjusted for model 13 is relatively high as compared to all other best subset regression models in Table 11. Figure 43 shows a plot of R^2 and R^2 adjusted for the best subset regression models. The plot shows little to no improvement in R^2 adjusted for cases where the number of parameters are greater than 6.

For model number 13, the error sum of squares (SSE) and the Mean Square Error (MSE) were found to be equal to 0.3973 and 0.0064, respectively. SSE is a measure of the discrepancy between the actual data and the predicted data. A small SSE corresponds to a high R^2, which indicates a tight fit of the model to the data. MSE, on the other hand, is an unbiased estimator of the error variance. The lower the SSE and MSE, the better the model is. A brief description of SSE and MSE can be found in Appendix A. Figure 44 also shows a plot of SSE for the best subset regression models. The plot shows no significant reduction in SSE for cases where the number of parameters are greater than 6.

Mallows' C_p is a model selection criterion used to avoid overfitting. For Mallows' $C_{p,}$ we seek models with two conditions: 1) small C_p values; and 2) C_p values near the number of parameters. For the six-parameter model number 13, C_p is equal to 5.24; a relatively small value that is slightly lower than the number of parameters "6", which is very desirable. Figure 45 shows the C_p values for all best subset regression models. The C_p line shown in Figure 45 denotes the optimum cases where C_p equals to the number of parameters p. Models that fall above the C_p line will tend to have substantial bias. On the other hand, models with C_p near the number of parameters have small bias, which is the case for the six-parameter model number 13. A brief description of Mallows' C_p can be found in Appendix A.

Table 12. Best Subset Models for Different Number of Parameters

Model No.	*p*	RSQ	RSQ_A	SSE	MSE	Cp	BIC
1	2	0.2938	0.2831	0.6765	0.0103	40.2883	-15.215
2	2	0.167	0.1544	0.798	0.0121	59.0159	-3.9845
3	2	0.1122	0.0987	0.8505	0.0129	67.1126	0.35
4	3	0.4294	0.4118	0.5467	0.0084	22.268	-25.4901
5	3	0.3786	0.3595	0.5953	0.0092	29.7683	-19.6921
6	3	0.3424	0.3222	0.63	0.0097	35.113	-15.8428
7	4	0.4977	0.4742	0.4811	0.0075	14.1698	-29.9505
8	4	0.4936	0.4699	0.4851	0.0076	14.7809	-29.3925
9	4	0.4751	0.4505	0.5028	0.0079	17.5149	-26.9509
10	5	0.5573	0.5292	0.4241	0.0067	7.375	-34.3138
11	5	0.542	0.5129	0.4388	0.007	9.6409	-31.9968
12	5	0.5283	0.4984	0.4519	0.0072	11.6579	-29.9987
13	**6**	**0.5853**	**0.5518**	**0.3973**	**0.0064**	**5.2428**	**-34.5342**
14	6	0.5776	0.5435	0.4047	0.0065	6.3794	-33.2838
15	6	0.5711	0.5365	0.4109	0.0066	7.334	-32.2511
16	**7**	**0.6027**	**0.5636**	**0.3806**	**0.0062**	**4.6755**	**-33.2268**
17	7	0.5985	0.559	0.3846	0.0063	5.2942	-32.5134
18	7	0.5903	0.5501	0.3924	0.0064	6.4956	-31.1495
19	8	0.6132	0.568	0.3706	0.0062	5.1265	-30.8265
20	8	0.6071	0.5613	0.3764	0.0063	6.0186	-29.7729
21	8	0.6069	0.561	0.3766	0.0063	6.0536	-29.7318
22	9	0.6205	0.569	0.3636	0.0062	6.0458	-27.9058
23	9	0.6137	0.5613	0.3701	0.0063	7.0447	-26.7045
24	9	0.6132	0.5608	0.3705	0.0063	7.1141	-26.6218
25	10	0.6207	0.5619	0.3633	0.0063	8.0109	-23.7287
26	10	0.6206	0.5618	0.3634	0.0063	8.0211	-23.7162
27	10	0.6205	0.5616	0.3636	0.0063	8.0458	-23.6862

In Table 12, it should be noted that the seven-parameter model number 16 is actually Model 1, which was selected using the AIC stepwise regression in section 5.1 of this chapter. Model 1 (or the best subset regression model number 16) has a C_p of 4.67, which is less than the number of parameter 7. Nevertheless, the six-parameter model number 13 is better than the seven-parameter model number 16 in regard to C_p. Model number 13 has a C_p much closer to the number of parameters as compared to model number 16. Model number 16 may tend to overfit the data, and is more biased to the sample in predicting %Delta as compared to model number 13. Finally, it should be noted that model number 13 was also detected in section 5.1 of this chapter using the BIC stepwise regression, which is a method that tends to select more parsimonious models. Finally, R^2 adjusted for model number 13 and model number 16 are very close to one another; they are equal to 55.18% and 56.36%, respectively. Model number 13 is more desirable than model number 16 because it is able to explain the same amount of variation in %Delta with less predictors.

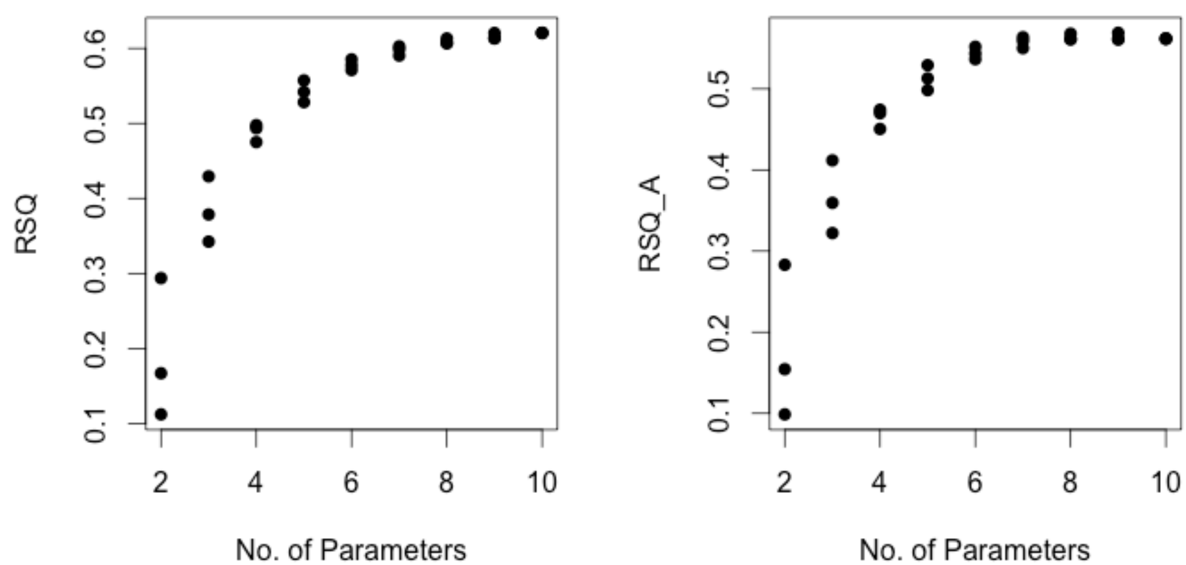

Figure 43. R^2 and R^2 adjusted for the Best Subset Models

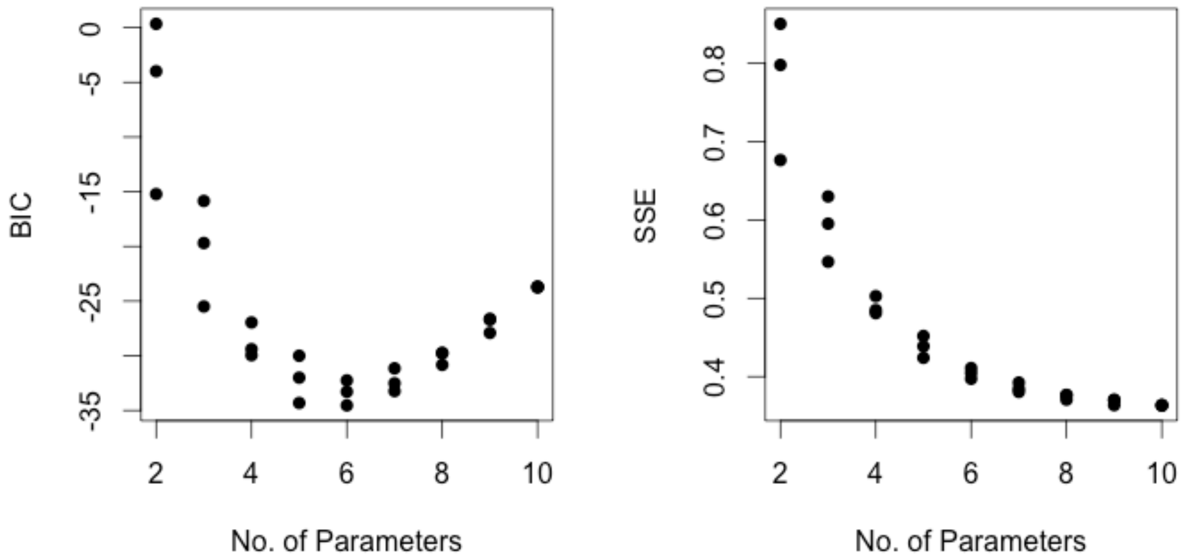

Figure 44. BIC and SSE for the Best Subset Models

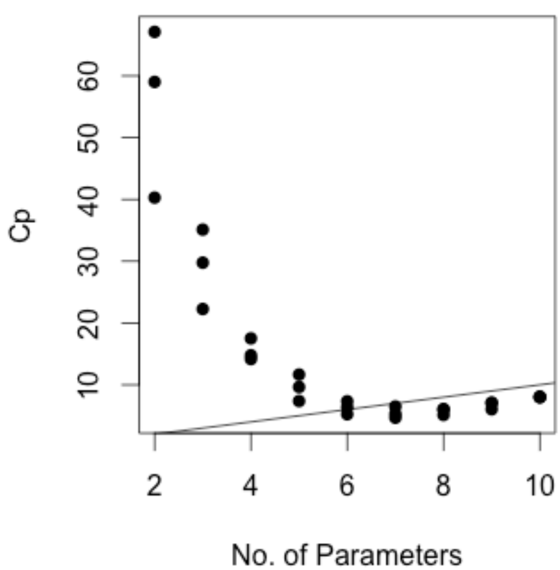

Figure 45. C_p values for the Best Subset Models

5.2.2. Analyzing the Optimal Best Subset Regression Model

Based on the output of the Best Subsets Algorithm, the six-parameter model number 13 was identified as optimal. Following is the multiple regression equation for the model:

Model 2:

%Delta = 0.359326 -0.144190 %OwnerInitiatedCO -0.088081 Productivity + 0.030369 Turnover -0.082290 PM%TimeOnProject + 0.052126 Overmanning

Model 2 has a coefficient of multiple determination (R^2) of 58.53%, a fairly high value for the type of data analyzed. The R^2 adjusted for the model is 55.18%, which didn't vary much from R^2. The null hypothesis that states that all regression parameters of Model 2 are equal to zero was rejected with a p-value of 9.01e-11.

Table 13 shows the description of the variables included in Model 2. The model should be only applied within the variable ranges stated in Table 14. Table 15 shows the model summary, including the coefficient of each predictor and its associated p-value. All p-values were less than the 0.05 threshold, resulting in the rejection of the null hypothesis that states that a given model coefficient is equal to zero.

Table 16 shows the ANOVA table for Model 2. The SSR, SSE, SSTO for Model 2 were found to be equal to 0.56069, 0.39729, and 0.95798, respectively. It should be noted that R^2 for Model 2 can be obtained by dividing SSR by SSTO. A brief description of these statistical terminologies are included in Appendix A.

Table 13. Definitions of Variables included in Model 2

Predictor name	**Definition of variables**
%OwnerInitiatedCO	-Numerical (in Decimals): The number of owner initiated change items divided by the total number of change items.
Productivity	-Binary: 1 = The contractor was tracking productivity 0 = The contractor was not tracking productivity Note: The contractor could use one of the following methods: -Track percent complete by earned value -Track percent complete by actual work-hours -Track percent complete by actual installed quantities
Turnover	The percent of craftsmen hired to replace those who have left of the number of craftsmen employed on the project. -Ordinal: 1= 0 to 5% 2= 6 to10% 3 = 11 to 20% 4 = Greater than 20%
PM%TimeOnProject	-Numerical (in Decimals): Percent of time by the Project Manager on the project
Overmanning	-Binary: 1 = Ratio of Actual Peak to Actual Average Manpower > 1.6 0 = Ratio of Actual Peak to Actual Average Manpower < 1.6

Table 14. Description of Predictors for Model 2

Predictor name	Min.	Max.	Mean	Median	Standard Deviation
%OwnerInitiatedCO	0	1	0.58	0.72	0.366
Productivity	0	1	-	0	0.503
Turnover	1	4	1.79	1	1.015
PM%TimeOnProject	0.04	1	0.59	0.60	0.349
Overmanning	0	1	-	1	0.384

Table 15. Model 2 Summary

Predictor name	Coefficient	P-value
Intercept	0.359326	4.43e-13 ***
%OwnerInitiatedCO	-0.144190	2.09e-06 ***
Productivity	-0.088081	3.39e-05 ***
Turnover	0.030369	0.00331 **
PM%TimeOnProject	-0.082290	0.00489 **
Overmanning	0.052126	0.04508 *

Table 16. ANOVA Table for Model 2

Sample size (n) = 68 Number of parameters (p) = 6					
Source of Variation	**SS**	**df**	**MS**	**F-statistic**	**P-value**
Regression	0.56069	p-1 = 5	0.112138	17.50	9.01e-11
%OwnerInitiatedCO	0.28145	1	0.28145		
Productivity given (%OwnerInitiatedCO)	0.12987	1	0.12987		
Turnover given (%OwnerInitiatedCO, Productivity)	0.06551	1	0.06551		
PM%TimeOnProject given (%OwnerInitiatedCO, Productivity, Turnover)	0.05705	1	0.05705		
Overmanning given (%OwnerInitiatedCO, Productivity, Turnover, PM%TimeOnProject)	0.02681	1	0.02681		
Error	0.39729	n-p = 62	0.00640		
Total	0.95797	n-1 = 67	0.01429		

5.3. Constancy of Variance of the Error Terms

One of the most important diagnostics for regression models is the residual plot. Figures 46 and 47 show the residuals plotted against the fitted/predicted values of %Delta for Model 1 and Model 2, respectively. It can be noticed that for both models, the residuals fall within a horizontal band centered around zero. The absence of any systematic patterns in the residual plots indicates that both models are appropriate linear regression functions. The residual plots also show that the residual variance is constant throughout the different levels of the fitted values, which reinforces the reasonableness of the multiple regression models. In addition to that, no outliers are detected from the residual plots.

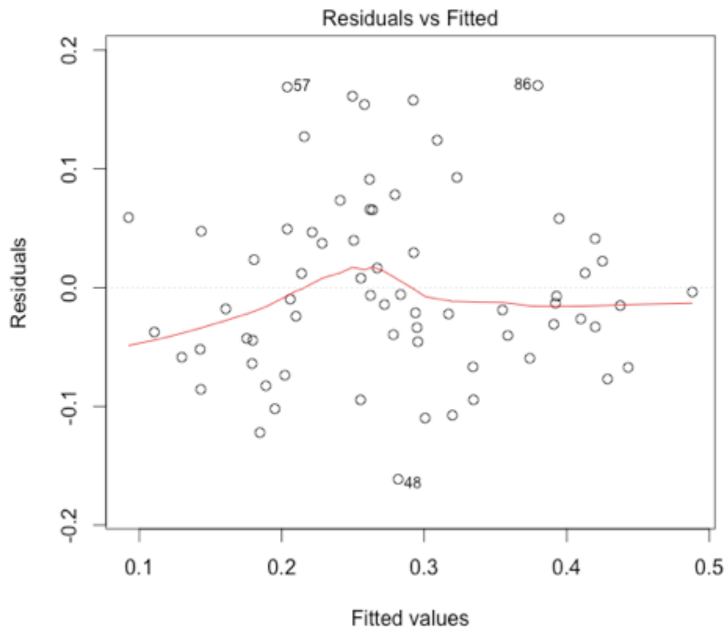

Figure 46. Residuals Against Fitted Values for Model 1

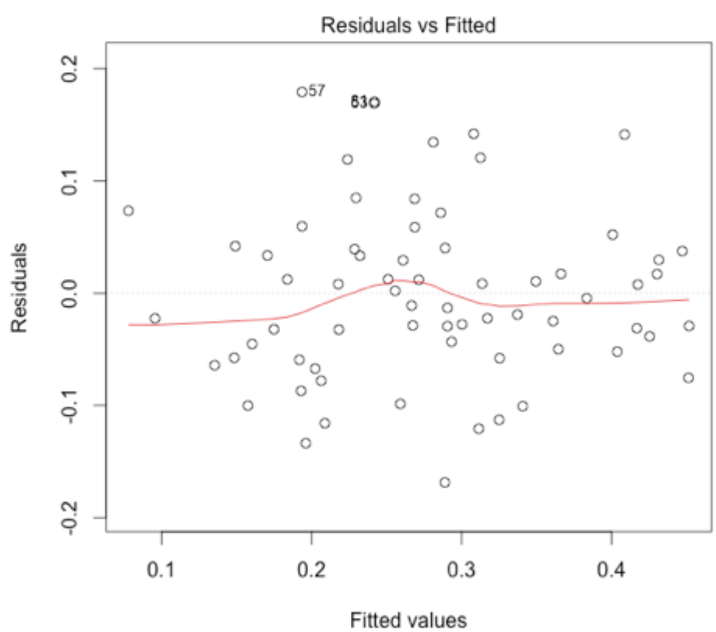

Figure 47. Residuals Against Fitted Values for Model 2

Although the residual variance was found to be constant across the different levels of the fitted values of %Delta, plots of the residuals against some of the predictors were also developed. Figures 48 and 49 show the plots of the residuals against "*%OwnerInitiatedCO*" for Model 1 and Model 2, respectively. The plots provide sufficient support for the assumption of constancy of error variance in relation to "*%OwnerInitiatedCO.*" To further support this, the Brown-Forsythe Test was used to check whether the error variance increases or decreases with the different levels of "*%OwnerInitiatedCO.*" The data was split into two equally sized groups, where the first group has relatively low values regarding the predictor "*%OwnerInitiatedCO*", and the second group has relatively high values regarding the same examined predictor. The test then checks whether the error variance varies across the two groups. By controlling the level of risk α at 0.05, the null hypotheses stating that the error variance is constant across the levels of "*%OwnerInitiatedCO*" for Model 1 and Model 2 were retained with p-values of 0.4342 and 0.3418, respectively.

The same test was used for "*PM%TimeOnProject.*" Figures 50 and 51 show the plots of the residuals against "*PM%TimeOnProject*" for Model 1 and Model 2, respectively. The null hypotheses stating that the error variance is constant in relation to "*PM%TimeOnProject*" for Model 1 and Model 2 were retained with p-value of 0.7331 and 0.9552, respectively. The Brown-Forsythe Test for "*PM%TimeOnProject*" provided further evidence for the appropriateness of both Model 1 and Model 2. A brief description of the Brown-Forsythe Test can be found in Appendix A.

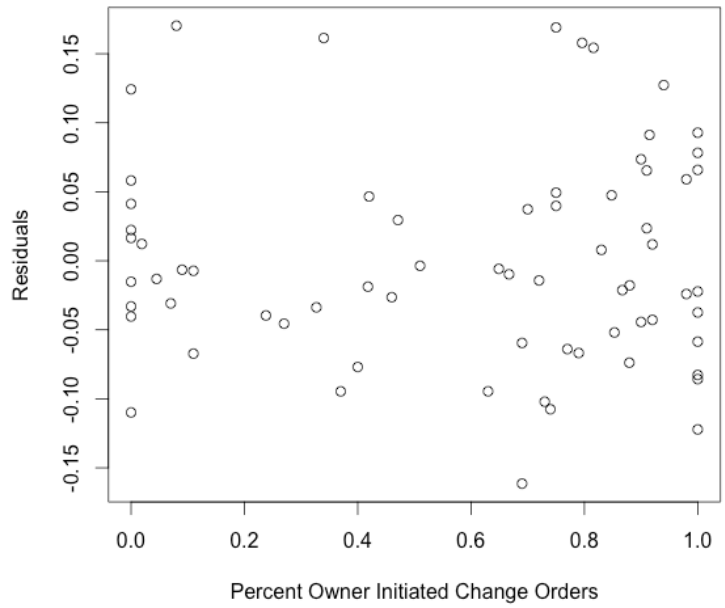

Figure 48. Plot of Residuals Against "%OwnerInitiatedCO" for Model 1

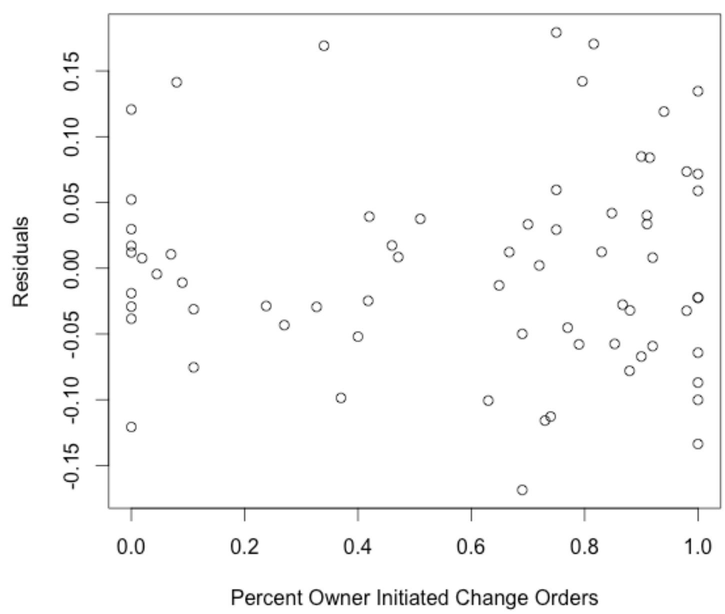

Figure 49. Plot of Residuals Against "%OwnerInitiatedCO" for Model 2

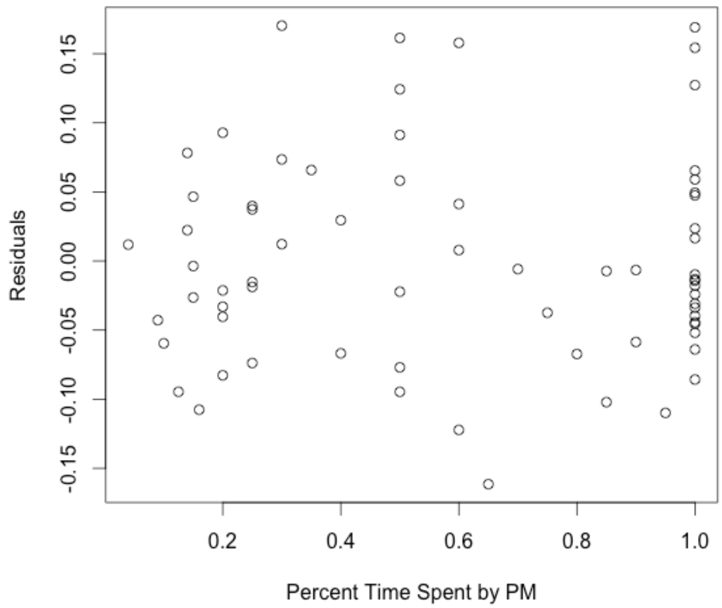

Figure 50. Plot of Residuals Against "PM%TimeOnProject" for Model 1

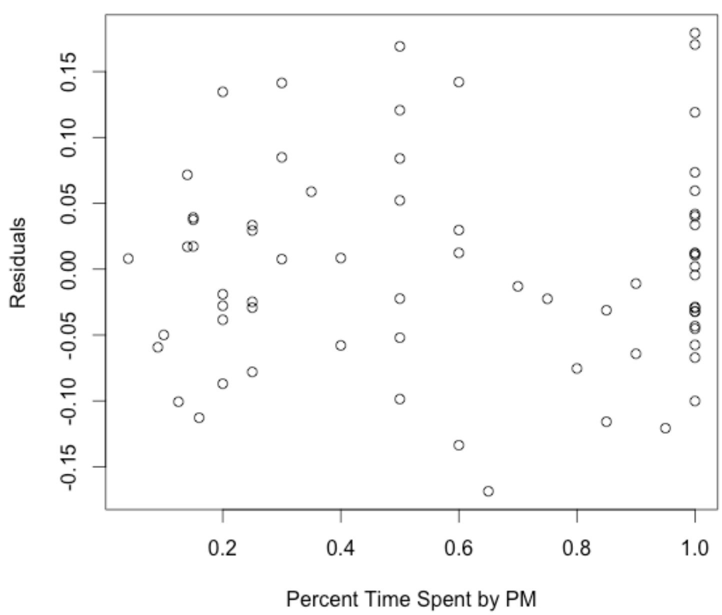

Figure 51. Plot of Residuals Against "PM%TimeOnProject" for Model 2

5.4. Correlation Test for Normality

In order to examine the normality of the residuals, plots of the residuals against their expected values under normality were developed for Model 1 and 2, as shown in Figures 52 and 53. The coefficient of correlation calculated from the plots are compared to a critical value that depends on the sample size and the level of risk α. The critical values for the coefficient of correlation for different sample sizes and for a given α level can be obtained from (Looney and Gulledge 1985). For a sample size of 68 projects and by controlling the α risk at 0.05, the critical value for the coefficient of correlation is 0.983 (Looney and Gulledge 1985). From the plots, the actual coefficients of correlation for Model 1 and Model 2 are equal to 0.986 and 0.990, respectively. Although the distribution of the residuals of Model 2 is much closer to a normal distribution as compared to Model 1, the actual coefficients of correlation exceed the critical value for both models. Consequently, we can conclude that the distributions of the error terms for Model 1 and Model 2 do not depart from a normal distribution. A brief description of how to get the expected values of the residuals under normality can be found in Appendix A.

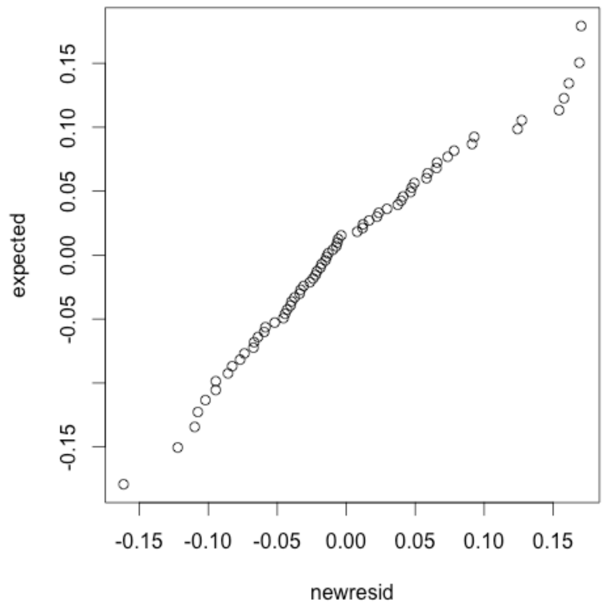

Figure 52. Plot of the Residuals vs. their Expected Values under Normality for Model 1

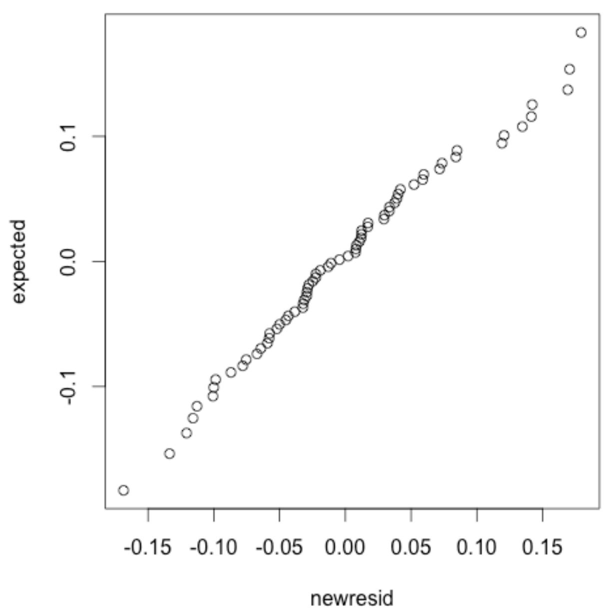

Figure 53. Plot of the Residuals vs. their Expected Values under Normality for Model 2

5.5. Statistical Check for Multicollinearity

In linear regression models, multicollinearity is said to exist when the predictors are highly correlated among themselves. Among the negative effects of multicollinearity are imprecise estimates regarding the true values of the regression coefficients. In other words, when the collected data or the sample is replaced, the estimated regression coefficients will tend to vary, which makes it difficult to detect the true regression model. Moreover, in the case of multicollinearity, the value of a regression coefficient no longer corresponds to the change in the response resulting from a unit change in a given predictor when all other predictors are held constant. Multicollinearity results in a problem where the value of the regression coefficient and the sequential sum of squares associated with any variable depend on which variables are included in the model and which variables are omitted. As a result, multicollinearity may lead to regression coefficients with counterintuitive signs, or regression coefficients that are not statistically significant although a definite relation exists between the dependent variable and the predictors.

A common method for examining multicollinearity is the use of Variance Inflation Factors (VIF). VIF is a measure of how much the variances of the regression coefficients are inflated as compared to when the variables are uncorrelated. By regressing a given predictor (X_k) on all other predictors, the coefficient of multiple determination R_k^2 can be obtained and VIF_k can be calculated through the following equation:

$$VIF_k = (1 - R_k^2)^{-1}$$

When R_k^2 equals to zero (i.e. a predictor X_k is uncorrelated with all other predictors), VIF_k equals to 1, and the variance of the regression coefficient for the examined predictor is not inflated, indicating no signs of multicollinearity. A rule of thumb is when a VIF is greater than

10, multicollinearity is said to affect the estimated regression coefficients. Tables 17 and 18 show VIF for each predictor included in Model 1 and Model 2, respectively. It should be noted that in order to calculate VIF, it is important to work with the standardized regression model. Consequently, the regression coefficient values indicated in Tables 17 and 18 result from the regression models that were fitted using the standardized %Delta and the standardized independent variables.

All VIFs for the predictors included in Model 1 and Model 2 are all near 1, indicating that there is no multicollinearity in the data. The average VIF for Model 1 and Model 2 equal to 1.13349 and 1.035269, respectively. This means that the expected sum of the squared errors in the least squares standardized regression coefficients is only 1.13349 larger (almost equal) as it would be if the independent variables were uncorrelated for Model 1. The same concept applies for Model 2. The average VIFs for both models are significantly lower than 10 and almost equal to 1, showing no signs of multicollinearity.

Table 17. Variance Inflation Factors for Model 1

Predictor Name	**b_k (Standardized regression coefficients)**	**VIF_k**	**k**
%OwnerInitiatedCO	**-0.445**	**1.064032**	**1**
Productivity	**-0.360**	**1.036148**	**2**
Turnover	**0.185**	**1.369044**	**3**
PM%TimeOnProject	**-0.246**	**1.017986**	**4**
Overmanning	**0.180**	**1.011116**	**5**
Absenteeism	**0.151**	**1.302611**	**6**

Table 18. Variance Inflation Factors for Model 2

Predictor Name	b_k (Standardized regression coefficients)	VIF_k	k
%OwnerInitiatedCO	**-0.441**	**1.063107**	**1**
Productivity	**-0.371**	**1.030044**	**2**
Turnover	**0.258**	**1.065410**	**3**
PM%TimeOnProject	**-0.241**	**1.016081**	**4**
Overmanning	**0.167**	**1.001705**	**5**

Another approach to diagnosing the presence of multicollinearity is the use of the Condition Index ($\emptyset$) (Belsley 1991). The Condition Index is based on the eigenvalues of the matrix $\boldsymbol{X'X}$, where $\boldsymbol{X}$ is the design matrix. A brief description of the Condition Index and the design matrix $\boldsymbol{X}$ can be found in Appendix A. A rule of thumb is when the Condition Index falls within the range of 30 to 100, there is evidence of moderate to strong multicollinearity. The largest Condition Index is called the Condition number. Tables 19 and 20 show that all Condition Indices are below 30, indicating that there are no signs of multicollinearity in Model 1 or Model 2.

The largest condition index is associated with the smallest eigenvalue of the matrix $\boldsymbol{X'X}$. Low eigenvalues indicate the presence of linear dependence among the variables. Tables 19 and 20 also show the proportion of variation in the coefficients that is affected by the linear dependence associated with each Condition Index, or each eigenvalue. The first Condition Index equals to 1 because it is equivalent to dividing the maximum eigenvalue by itself. The variance proportion indicates which coefficients are damaged by the linear dependence in the design matrix. For Model 1 and Model 2, none of the variance proportions in the regression coefficients are highly associated with the lowest eigenvalue (i.e. the highest condition index). In general, multicollinearity spotted when two or more variables have large proportions of variance that correspond to large condition indices. It should be noted that in order to compute the variance proportions, the predictors were only scaled to unit length, but not centered. If the predictors were centered, the condition indices and the variance proportions become remarkably smaller than the values shown in Tables 19 and 20.

Table 19. Variance Decomposition Proportions for Model 1

Condition Index		Intercept	*%Owner InitiatedCO*	*Productivity*	*Turnover*	*PM%Time OnProject*	*Overmanning*	*Absentee*
1	1.000	0.002	0.006	0.009	0.005	0.006	0.005	0.004
2	3.515	0.002	0.000	0.906	0.016	0.001	0.010	0.021
3	4.003	0.000	0.375	0.043	0.158	0.032	0.003	0.039
4	4.840	0.000	0.203	0.006	0.017	0.785	0.010	0.007
5	5.253	0.002	0.149	0.001	0.023	0.019	0.615	0.083
6	7.327	0.002	0.080	0.024	0.735	0.023	0.008	0.690
7	11.460	0.992	0.188	0.011	0.047	0.134	0.350	0.156

Table 20. Variance Decomposition Proportions for Model 2

Condition Index		Intercept	*%Owner InitiatedCO*	*Productivity*	*Turnover*	*PM%Time OnProject*	*Overmanning*
1	1.000	0.003	0.009	0.013	0.008	0.009	0.006
2	3.321	0.004	0.017	0.965	0.007	0.009	0.023
3	3.868	0.000	0.437	0.001	0.311	0.007	0.008
4	4.487	0.001	0.130	0.003	0.042	0.804	0.045
5	5.116	0.000	0.167	0.014	0.340	0.002	0.579
6	10.054	0.993	0.240	0.004	0.293	0.170	0.338

5.6. Understanding the Difference between Outlying Cases and Influential Cases

In regression analysis, some cases can be detected as outlying or extreme. However, a key point to understand in linear regression is that not all outlying cases are influential. Some outliers could have dramatic influence on the fitted regression model, while others do not. The outlying cases should be studied carefully in order to decide whether they should be eliminated from a model. If they are kept in the model, their influence should be quantified in order to determine whether the model should be revised through remedial measures.

Observations can be outlying with respect to the Y value (%Delta) or the X values (the predictors). Cases could be also outliers with respect to both Y and X values. In order to detect outliers with respect to Y, the studentized residuals and the deleted studentized residuals will be used. On the other hand, the hat matrix leverage values are typically used in order to detect outliers with respect to X. Appendix A includes a brief description of these statistical terms.

After spotting outliers, several statistical measures can be used to detect whether the outlying cases are influential. Those include DFFITS, Cook's Distance, and DFBETAS. Appendix A includes a brief description of these statistical measures.

5.6.1. Checking Outlying Y Observations

Before checking whether there are outlying Y observations, it is very important to distinguish between the "error" and the "residual." The error is the deviation of an observed value from the true or expected value of the predicted %Delta. On the other hand, the residual is the deviation of an observed value from the predicted value of the "estimated" regression model. The residual plots that were developed in section 5.3 of this chapter were based on the ordinary residuals of the model. Some refinements can be made to the ordinary residuals to increase the efficiency of the analysis. The key motivation of these refinements is that the ordinary "residuals" can have different variances, even when the assumption of the constancy of "errors" holds.

5.6.1.1. Studentized Residuals

A first refinement to the ordinary residuals are the studentized residuals. Studentized residuals account for the sampling errors of the residuals. They are calculated by dividing each ordinary residual by its estimated standard deviation. Studentized residuals are scaled to be "t-like"; they have a variance of 1, unlike the ordinary residuals. A rule of thumb is when an observation has an absolute studentized residual that is exceedingly greater than 2, the observation is considered as an outlying Y observation. Consequently, a studentized residual can be considered as an outlier statistic.

Table 21 shows the highest studentized residuals for Model 1 and Model 2. For both models, 95% of the studentized residuals fell within the range of ± 2. The rest of the studentized residuals have absolute values very close to 2, indicating that there are no outliers. Usually, outliers are detected when they fall far beyond the range of ± 2. In the upcoming section, a more reliable technique will be used in order to detect outlying Y observations.

Table 21. Highest Studentized Residual values for Model 1 and Model 2

Statistical Measure	Value	Case Number
Highest studentized residual for Model 1	2.37	Case number 68
Highest studentized residual for Model 2	2.29	Case number 45

5.6.1.2. Studentized Deleted Residuals

A second refinement to ordinary residuals is to compute the ith residual when the regression model is fitted on all observations except the ith one (i.e. the ith case is deleted). Therefore, these refined residuals are called "deleted residuals." The reason for this refinement is that if an observation i is far outlying in Y, the fitted regression based on all observations including the ith observation will tend to be closer to the actual value of the response variable and will not reveal that the ith observation is outlying. On the other hand, if the regression function is fitted based on all observations excluding the ith observation, the residual for the ith observation will tend to be larger, thus disclosing the outlying Y observation (Kutner et al.

2004).

The studentized deleted residual is computed by dividing the deleted residual by its estimated standard deviation. The deleted studentized residuals follow a t distribution with $n - p - 1$ degrees of freedom. In addition, the studentized deleted residuals are more sensitive to outliers as compared to studentized residuals. Appendix A shows how the studentized deleted residuals can be calculated without fitting a new regression line for each deleted observation. It should be noted that the studentized deleted residual is often called R-Student. If the model is appropriate and no cases are outlying based on a change in the model, the R-student will follow the t distribution with $n - p - 1$ degrees of freedom.

In order to check whether any of the 68 residuals is an outlier, the Bonferroni test procedure can be used in order to adjust for 68 two-tailed tests. The $(\frac{0.05}{2})$ α risk will therefore be adjusted to $(\frac{0.05}{2\times 68})$. For Model 1, the T-critical value is 3.558921; it was obtained from the t-distribution using the Bonferroni adjusted α risk and the $(n - p - 1) = (68 - 7 - 1)$ degrees of freedom. Similary, the T-critical value for Model 2 is 3.555783, which was obtained using the same adjusted Bonferroni α risk, but with $(n - p - 1) = (68 - 6 - 1)$ degrees of freedom. By examining the R-student values for Model 1 and Model 2, it was found that none of them exceeded the critical T-values, indicating no signs of outliers.

Table 22 shows the highest R-student for Model 1 and Model 2. For Model 1, case number 68 had the largest R-student among all 68 cases. The R-student for case number 68 for Model 1 is 2.46, which is still lower than the critical value of Model 1 (3.558921), indicating that it is not an outlier. For Model 2, case number 45 had the largest R-student among all 68 cases. The R-student for case number 45 for Model 2 is 2.38, which is still lower than the critical value of Model 2 (3.555783), indicating that it is not an outlier.

Although the observations with the largest R-student values were not identified as outliers, we might still need to consider whether they influence the linear regression functions. Section 5.6.3 of this chapter will highlight the statistical measures that were used in order to detect whether any of the observations affected or influenced Model 1 or Model 2.

Table 22. Highest R-Student values for Model 1 and Model 2

Statistical Measure	Value	Case Number
Highest R-student for Model 1	2.46	Case number 68
Highest R-student for Model 2	2.38	Case number 45

5.6.2. Checking Outlying X Observations

In this section, our focus is to determine whether there are outliers with respect to the X values (or predictors). An outlying X observation has extreme values regarding the independent variables (i.e. predictors). Similar to outlying Y observations, outlying X observations could also influence the fitted regression model.

5.6.2.1. Leverage Values in the Hat Matrix

The hat matrix $\boldsymbol{H}$ plays a vital role in linear regression. The hat matrix is solely function of the X values (i.e. the predictor values). In fact, any predicted %Delta is a linear combination of all the 68 actual %Deltas. The vector of predicted %Delta values $\hat{\boldsymbol{Y}}$ can be expressed as linear combinations of the actual %Delta values $\boldsymbol{Y}$ through the following equation:

$$\hat{\boldsymbol{Y}} = \boldsymbol{HY}$$

Appendix A includes a brief description of the hat matrix. The diagonal elements of the hat matrix are called leverages. Each of the 68 observations has a leverage value that falls between 0 and 1. An observation with high leverage value is considered distant from the center of all X observations (Kutner et al. 2004). It should be noted that the sum of all 68 leverages equals to the number of parameters p. The average or the mean leverage is thereby equal to p/n.

A rule of thumb is when an observation has a leverage value greater than twice the mean leverage value, that observation is considered to be an outlying X observation. For Model 1, twice the mean leverage equals to $\frac{2x7}{68} = 0.205$. For Model 2, twice the mean leverage equals to $\frac{2x6}{68} = 0.176$. Figures 54 and 55 show the leverage values of all 68 observations for Model 1 and Model 2, respectively. Most of the observations have low leverage values below 0.2 in both models. For Model 1, cases 13 and 35 were found to have leverages slightly higher than the threshold 0.205. More specifically, the leverages for cases 13 and 35 for Model 1 are 0.230 and 0.237, respectively. For Model 2, only case 6 had a leverage of 0.195, which is slightly higher than the threshold for Model 2 (0.176).

Although cases 13 and 35 in Model 1 and case 6 in Model 2 have leverage values slightly higher than the calculated thresholds, they are not considered as outlying X observations. They have very low leverage values centered around 0.2. Recalling that the maximum leverage value is 1, these cases are considered having normal X or predictor values. In addition to that, Figures 54 and 55 show that there is a tiny gap between these cases and the rest of the cases, indicating that they are close to the centroid of the X observations. Despite the fact that these cases are not identified as outliers, we will still examine whether they are influential in section 5.6.3 of this chapter.

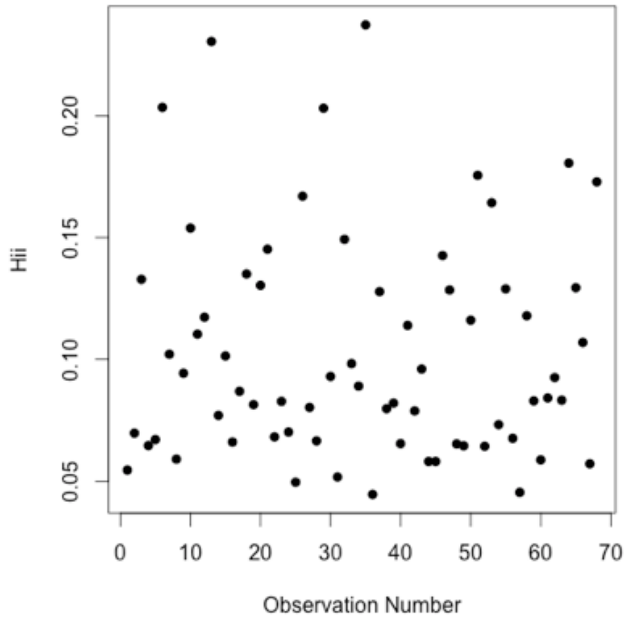

Figure 54. Leverage Values for Model 1

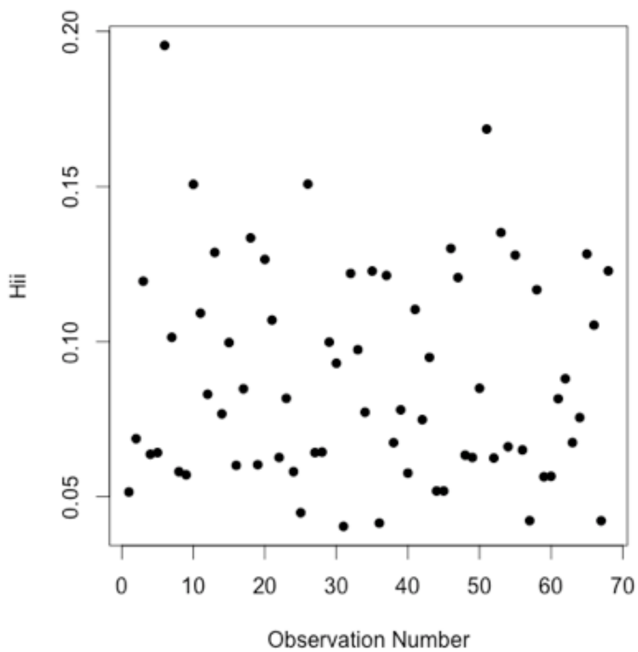

Figure 55. Leverage Values for Model 2

5.6.2.2. Checking Hidden Extrapolations

Hidden extrapolations involve observations with X values outside the range of the data used to build a regression model. Hidden extrapolation appears when the data is not enough to generate robust regression models, or when an analyst is trying to build a regression model using too many variables as compared to the number of data points. The results are impractical regression models that have substantial gaps in the range of X values (i.e. predictors). Hidden extrapolations are dangerous and can exist in models with high R^2 and R^2 adjusted.

Our biggest threat is to obtain impractical regression models that yield negative values of %Delta. Impacted projects only include projects with positive %Delta. Therefore, any inferences made from the regression model should only yield positive %Deltas. Predicting new cases that are extreme in regards to all their X values, but within the range of each individual predictor, is a very important step before accepting a regression model. For Model 1, the smallest possible prediction for %Delta is 6.53%. For Model 2, the smallest possible prediction is 7.51%. Consequently, both models are adequately built and can never result in negative %Delta values.

5.6.3. Identifying Influential Cases

The results of sections 5.6.1 and 5.6.2 showed that there is no substantial evidence for the presence of any outlying Y or X observations. Nevertheless, we will still examine whether any of the 68 observations is influential (i.e. has an influence on or impacts Model 1 or Model 2).

5.6.3.1. DFFITS – Influence on one fitted value

The DFFITS is a quantity that can be computed for each of the 68 observations. It is based on the difference between the fitted value for the ith case when all 68 cases are used in fitting the regression equation and the fitted value when the ith case is excluded from the fitting process (Kutner et al. 2004). Appendix A includes a detailed description of DFFITS.

Interestingly, Appendix A shows how DFFITS can be computed for each observation without having to fit new regression functions each time an observation is deleted.

A rule of thumb is that for large samples (larger than 100 observations), a case could be influential when DFFITS exceeds $2\sqrt{\frac{p}{n}}$. For small to medium-sized samples, a case could be influential when DFFITS exceeds 1. Given that our sample size is medium-sized, a threshold of 1 was adopted for DFFITS.

Recalling the results of Model 1 in sections 5.6.1 and 5.6.2, it was found that case number 68 is a possible outlying Y observation and cases 13 and 35 are possible outlying X observations. For cases 13 and 35, it was found that they have a DFFITS of 0.02 and 0.21, respectively. These values are significantly smaller than 1, showing that the presence of cases 13 and 35 did not do major changes in the regression model. On the contrary, Figure 56 shows that case number 68 had a DFFITS of 1.12; the highest amongst all observations in Model 1. Nevertheless, this DFFITS value is very close to 1; it may not require revising the regression model or taking a remedial action. In section 5.6.4, we will prove that observation number 68 is not influential enough to be removed from the analysis; it did not really affect the fitted values of the regression model.

Recalling the results of Model 2 in sections 5.6.1 and 5.6.2, it was found that case number 45 is a possible outlying Y observation and case 6 is a possible outlying X observation. The DFFITS values for those two observations are 0.20 and 0.55, respectively. According to the guideline of 1, it is concluded that those observations are not influential. All observations of Model 2 had DFFITS lower than 1, as shown in Figure 57.

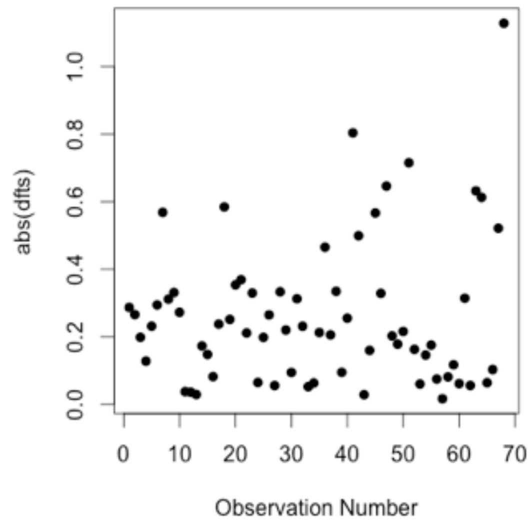

Figure 56. DFFITS absolute values for Model 1

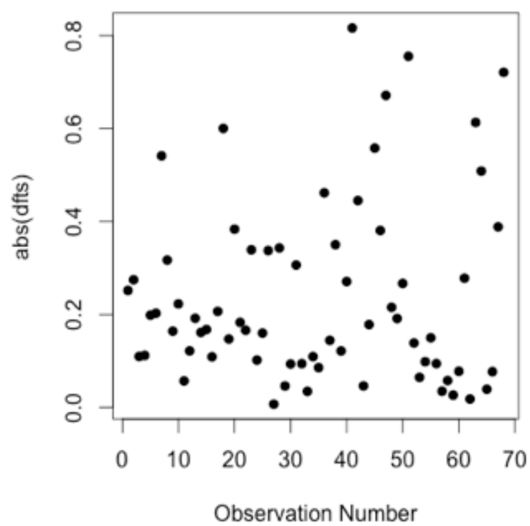

Figure 57. DFFITS absolute values for Model 2

5.6.3.2. Cook's Distance – Influence on all fitted values

Unlike DFFITS, Cook's distance is an aggregate measure that accounts for the influence of deleting the ith case on the fitted values of all 68 observations at a time (Kutner et al. 2004). Cook's distance can be easily calculated through an algebraic equation without fitting a new regression equation each time a case is omitted. Appendix A includes a detailed description on how to calculate Cook's distance. In order to identify influential points, the percentile value of Cook's distance should be calculated from the F $(p, n - p)$ distribution. A rule of thumb is that if the percentile value of a given observation in the F-distribution is less than 20%, that observation has little influence on all 68 fitted values.

Figures 58 and 59 show the Cook's distance values of all 68 observations for Model 1 and Model 2, respectively. For Model 2, the percentile value of all Cook's distance were below 1%, indicating that none of the points in Model 2 are influential. For Model 1 and as shown in Figure 59, there is a substantial gap between observation number 68 and the rest of the observations, indicating that observation number 68 may have an influence on the fitted values. Cook's distance for observation 68 equals to 0.167. By using the F distribution with degrees of freedom $(7, 68 - 7)$, it was found that the percentile value of Cook's distance for observation 68 in Model 1 equals to 1%, which is significantly lower than 20%. This indicates that observation 68 is not influential enough according to Cook's distance.

It should be noted that DFFITS for observation number 68 was also the highest among all observations for Model 1. Both DFFITS and Cook's distance showed a little evidence that observation 68 in Model 1 may be problematic. Although observation 68 had low DFFITS and Cook's distance values, observation 68 for Model 1 will be closely examined in section 5.6.4 of this chapter in order to confirm that it is not influential enough to be omitted from the analysis.

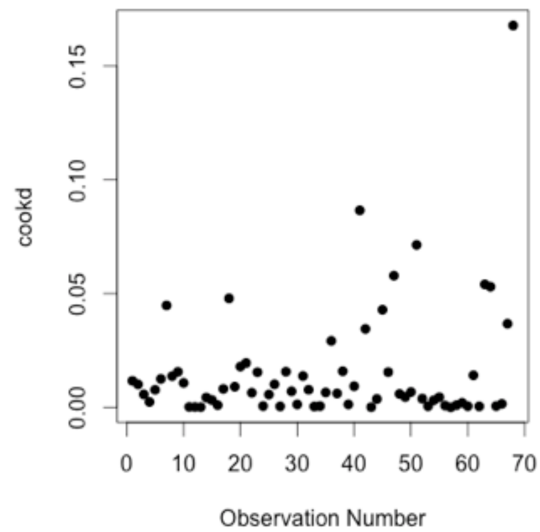

Figure 58. Cook's Distance for Model 1 Observations

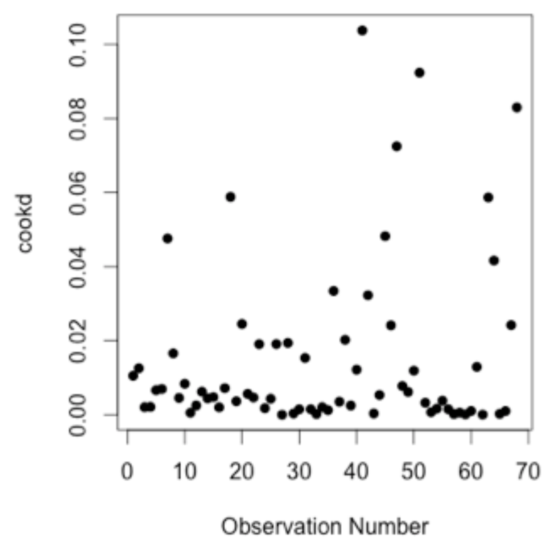

Figure 59. Cook's Distance for Model 2 Observations

5.6.3.3. DFBETAS – Influence on Regression Coefficients

DFBETAS is another influence measure that focuses on the effect of removing an ith case on each of the regression coefficients of a regression model (Kutner et al. 2004). By considering the same threshold used for DFFITS, a case is considered to be influencing a given regression coefficient when the absolute DFBETAS value is greater than 1 for that coefficient.

Table 23 shows the maximum DFBETAS values of all regression coefficients for Model 1 and Model 2. None of the DFBETAS exceeded 1, indicating that the estimated regression coefficients of both Model 1 and Model 2 are close to the reality, as they did not vary by removing any of the 68 observations.

Table 23. Maximum DFBETAS for Model 1 and Model 2 Regression Coefficients

Model/Factor	Intercept	*%Owner Initiated CO*	*Productivity*	*Turnover*	*PM%Time OnProject*	*Overmanning*	*Absentee*
Model 1 *Maximum Absolute DFBETAS*	0.50 (Case 51)	0.41 (Case 7)	0.32 (Case 41)	0.80 (Case 68)	0.31 (Case 63)	0.61 (Case 41)	0.60 (Case 68)
Model 2 *Maximum Absolute DFBETAS*	0.52 (Case 51)	0.40 (Case 51)	0.32 (Case 41)	0.44 (Case 68)	0.35 (Case 63)	0.65 (Case 41)	**N/A** (Predictor does not exist in Model 2)

5.6.4. Examining Potential Influential Case for Model 1

For Model 1, observation 68 was tagged by DFFITS and Cook's distance for further examination. However, DFFITS and Cook's distance did not suggest that observation 68 should be omitted: DFFITS was very close to the threshold value 1, and Cook's distance had a very low percentile value in the F distribution. Both measures did not conclusively identify observation 68 as influential, but just pointed out that it requires closer attention.

To check the influence of observation 68, Model 1 was developed twice: once when case 68 is included, and once when case 68 is excluded. The %Delta of each of the 68 observations was predicted twice: once from the model made using all observations, and once from the model made using all observations except observation number 68. For each of the 68 observations, the difference between the two predicted %Deltas was reported.

The differences in predictions in %Delta were found to be very low, indicating that observation 68 is not influential enough to be excluded from Model 1. Of the 68 observations, 61 observations had differences less than 5% in %Delta. Moreover, the mean prediction difference was 2.87%, which is negligible. It was therefore concluded that observation 68 should be kept in the analysis as it did not affect the estimated regression coefficients or the inferences made from Model 1.

5.7. Summary

In this chapter, various model selection criteria were used in order to select the best candidate predictive regression models. A careful model building process was applied in order to avoid misleading models. Detailed model diagnostics were performed in order to check the linear model assumptions, such as the normality and the constancy of variance of the error terms. Furthermore, various challenges that analysts face when building regression models were statistically tested, such as multicollinearity and the existence of outlying and influential cases. In the next chapter, a set of validation procedures will be used in order to check the performance of the candidate models.

Chapter 6. Model Validation Procedures

In this chapter, the future predictive ability of the models developed in chapter 5 will be tested using cross-validation metrics as well as future observations. In cross-validation, several metrics can be reported in order to compare the performance of different models when predicting unseen data. In addition to that, sixteen new projects were collected to verify the performance of the models during actual usage. The sixteen projects were screened to confirm that they are impacted by changes. A detailed description of the screening procedure for the new observations can be found in Appendix B.

6.1. Crucial Notes regarding Linear Regression

One of the threatening phenomena in linear regression is "Overfitting." Overfitting occurs when a model tends to memorize the data rather than explaining a true generalized trend, resulting in a deceptive and misleading model. Overfitting generally occurs in models that are excessively complicated with too many predictors relative to the number of observations. A model with high R^2 or even R^2 adjusted can still suffer from Overfitting. An extreme example of this phenomenon is a model having a number of predictors equal to the number of data points/observations. This model could perfectly predict the data, but would drastically fail when predicting unseen data.

6.2. Cross-Validation Metrics

One way to detect overfitting and determine the prediction accuracy of regression models is to use 5-fold cross validation. In cross-validation, training models are used to predict unseen data. The unseen data are generated by randomly splitting the entire dataset into five parts, where each part is predicted from the remaining four parts. The prediction errors resulting from this process can therefore be assessed using different metrics. The model with the lowest prediction errors has the highest prediction power for future observations. It should be noted that cross-validation was performed twice: once for Model 1, and once for Model 2. Tables 24 through 28 show the actual %Delta versus predicted %Delta values for the 5 folds of the two cross-validation processes for Model 1 and Model 2.

Table 24. Actual vs. Predicted %Delta for the First Testing Set

First 20%					
	Actual %Delta	Predicted Model 1 %Delta	Predicted Model 2 %Delta	Deviation Model 1	Deviation Model 2
1	26.56%	21.85%	22.37%	4.71%	4.19%
2	32.77%	26.00%	27.00%	6.77%	5.77%
3	46.11%	39.48%	42.47%	6.63%	3.64%
4	43.33%	28.77%	29.15%	14.56%	14.18%
5	26.34%	25.71%	25.40%	0.63%	0.94%
6	25.00%	28.12%	28.72%	-3.12%	-3.72%
7	21.22%	30.49%	32.71%	-9.27%	-11.49%
8	32.92%	25.10%	29.89%	7.82%	3.03%
9	25.33%	20.40%	19.37%	4.93%	5.96%
10	13.53%	16.45%	20.89%	-2.92%	-7.36%
11	55.00%	34.14%	39.83%	20.86%	15.17%

Table 25. Actual vs. Predicted %Delta for the Second Testing Set

Second 20%					
	Actual %Delta	Predicted Model 1 %Delta	Predicted Model 2 %Delta	Deviation Model 1	Deviation Model 2
1	12.84%	21.73%	22.12%	-8.89%	-9.28%
2	38.69%	46.30%	47.19%	-7.61%	-8.50%
3	37.59%	45.41%	46.45%	-7.82%	-8.86%
4	31.45%	39.54%	38.81%	-8.09%	-7.36%
5	10.61%	20.20%	20.56%	-9.59%	-9.95%
6	26.75%	34.77%	34.07%	-8.02%	-7.32%
7	24.01%	36.44%	37.18%	-12.43%	-13.17%
8	31.83%	39.31%	37.65%	-7.48%	-5.82%
9	42.24%	47.43%	49.09%	-5.19%	-6.85%
10	25.58%	28.37%	29.03%	-2.79%	-3.45%
11	19.08%	32.88%	34.25%	-13.80%	-15.17%
12	27.76%	30.20%	31.05%	-2.44%	-3.29%
13	28.35%	28.86%	29.56%	-0.51%	-1.21%
14	32.21%	30.35%	32.38%	1.86%	-0.17%
15	35.77%	30.04%	30.71%	5.73%	5.06%
16	44.73%	46.92%	47.79%	-2.19%	-3.06%

Table 26. Actual vs. Predicted %Delta for the Third Testing Set

Third 20%					
	Actual %Delta	Predicted Model 1 %Delta	Predicted Model 2 %Delta	Deviation Model 1	Deviation Model 2
1	7.32%	12.03%	10.08%	-4.71%	-2.76%
2	16.06%	26.07%	26.48%	-10.01%	-10.42%
3	35.17%	43.31%	40.43%	-8.14%	-5.26%
4	38.56%	38.74%	41.30%	-0.18%	-2.74%
5	14.29%	15.21%	16.77%	-0.92%	-2.48%
6	13.27%	18.24%	20.35%	-4.97%	-7.08%
7	36.00%	40.40%	35.09%	-4.40%	0.91%
8	29.49%	31.32%	31.57%	-1.83%	-2.08%
9	12.02%	28.56%	29.35%	-16.54%	-17.33%
10	37.29%	20.03%	18.69%	17.26%	18.60%
11	19.62%	20.82%	17.91%	-1.20%	1.71%
12	27.26%	30.11%	31.11%	-2.85%	-3.85%
13	42.51%	40.99%	41.62%	1.52%	0.89%
14	45.01%	29.03%	30.94%	15.98%	14.07%

Table 27. Actual vs. Predicted %Delta for the Fourth Testing Set

Fourth 20%					
	Actual %Delta	Predicted Model 1 %Delta	Predicted Model 2 %Delta	Deviation Model 1	Deviation Model 2
1	26.79%	21.27%	22.11%	5.52%	4.68%
2	48.46%	47.12%	43.73%	1.34%	4.73%
3	20.42%	18.37%	17.69%	2.05%	2.73%
4	45.29%	38.66%	39.12%	6.63%	6.17%
5	9.31%	20.39%	21.36%	-11.08%	-12.05%
6	29.03%	23.69%	24.56%	5.34%	4.47%
7	18.59%	20.91%	21.49%	-2.32%	-2.90%
8	31.46%	23.86%	23.02%	7.60%	8.44%
9	7.13%	13.51%	14.06%	-6.38%	-6.93%
10	9.09%	14.84%	15.41%	-5.75%	-6.32%
11	19.10%	14.91%	15.48%	4.19%	3.62%
12	33.62%	34.63%	35.03%	-1.01%	-1.41%
13	37.90%	38.55%	37.85%	-0.65%	0.05%
14	41.24%	24.93%	23.81%	16.31%	17.43%

Table 28. Actual vs. Predicted %Delta for the Fifth Testing Set

Fifth 20%					
	Actual %Delta	Predicted Model 1 %Delta	Predicted Model 2 %Delta	Deviation Model 1	Deviation Model 2
1	35.28%	25.49%	25.88%	9.79%	9.40%
2	11.52%	17.25%	15.97%	-5.73%	-4.45%
3	25.78%	26.36%	25.34%	-0.58%	0.44%
4	22.58%	20.36%	20.34%	2.22%	2.24%
5	38.35%	39.47%	36.29%	-1.12%	2.06%
6	5.75%	15.05%	15.84%	-9.30%	-10.09%
7	41.10%	24.02%	23.34%	17.08%	17.76%
8	34.31%	21.52%	22.07%	12.79%	12.24%
9	15.15%	8.09%	7.09%	7.06%	8.06%
10	6.25%	17.47%	18.18%	-11.22%	-11.93%
11	23.88%	27.98%	27.17%	-4.10%	-3.29%
12	26.12%	28.99%	28.73%	-2.87%	-2.61%
13	41.58%	29.49%	26.68%	12.09%	14.90%

6.2.1. Prediction MSE

Table 29 shows the different cross-validation metrics for each of the two models developed in this study: Model 1 and Model 2. One of the metrics in Table 29 is the Prediction MSE, which was also used in chapter 5. The lower the Prediction MSE, the higher the predictive ability of the model. It should be noted that in chapter 5, we were cross-validating different methods. In this chapter however, we are cross-validating different models. The Prediction MSE is based upon splitting the data into a number of folds and predicting %Delta for unseen data (i.e. the testing sets) using the two regression models, Model 1 and Model 2. More specifically, the Prediction MSE is computed by taking the weighted average of the Mean Squared Error of the 5 testing sets of the cross-validation process, which is equivalent to using the two steps mentioned in section 5.1.2 in chapter 5. As shown in Table 29, Model 1 has a lower Prediction MSE (0.00677) as compared to Model 2 (0.00687). Therefore, Model 1 slightly performs better than Model 2.

Table 29. Cross-Validation Metrics for the Two Cumulative Impact Models

	Model 1	**Model 2**
Sample Size	68	68
Prediction MSE	0.00677	0.00687
PRESS	0.468	0.473
Mean Absolute Percent Error (MAPE)	32.67%	34.01%
%Delta Deviation Range	±20.9%	±18.6%
Correlation Accuracy	72.49%	72.12%

6.2.2. Prediction Sum of Squares (PRESS) Criterion

The PRESS criterion is a form of data splitting used to measure the predictive accuracy of regression models. PRESS resembles to SSE in that it accounts for the sum of the errors. However, PRESS considers each observation as a new observation. PRESS is calculated by deleting an ith case in the dataset, estimating the regression model using the remaining cases, and then predicting %Delta for the ith case using the estimated regression model (Kutner et al. 2004). This procedure is repeated for all n observations, and the sum of the squared prediction errors is computed. In the context of this research, PRESS can be also thought of as a 68-fold cross-validation. For any given regression model, PRESS should be always greater than SSE. However, a close agreement between PRESS and SSE suggests that the model meets its expected predictive ability, which is denoted by the original MSE of the model. Table 29 shows that PRESS for Model 1 and Model 2 are equal to 0.468 and 0.473, respectively. In Chapter 5, SSE for Model 1 and Model 2 were found to be 0.380 and 0.397, respectively. The PRESS values of both models did not vary much from their corresponding SSE. This suggests that both Model 1 and Model 2 are able to adequately predict future observations.

6.2.3. Mean Absolute Percent Error (MAPE)

Another measure of the prediction accuracy of regression models is the Mean Absolute Percent Error (MAPE). The MAPE is computed by using the results of the 5-fold cross-validation process by first getting the absolute percent error of each data point. The MAPE can then be obtained by calculating the average value of the absolute percent errors of the 68 data points. It should be noted that through cross-validation, each data point is considered as a new observation. Therefore, MAPE gives an indication of the future predictive ability of regression models. MAPE is expressed in percentages; it is simpler and easier to understand as compared to the Prediction MSE. Table 29 shows that Model 1 and Model 2 have MAPE of 32.67% and 34.01%, respectively. Recalling that predicting the cumulative impact of changes is a very difficult task, percent errors around 30% are considered adequate for the type of data analyzed.

6.2.4. Deviation Range

The deviation in %Delta for each data point through the cross-validation process can be obtained by computing the difference between the actual %Delta and the predicted %Delta from the testing sets. For Model 1, all sixty-eight projects were predicted within ±20.9% from the actual %Delta, sixty-two projects were predicted within ±15% from the actual %delta, fifty-four projects were predicted within ±10% from the actual %delta, and thirty projects were predicted within ±5% from the actual %delta. For Model 2, all sixty-eight projects were predicted within ±18.6% from the actual %Delta, sixty-three projects were predicted within ±15% from the actual %delta, fifty-two projects were predicted within ±10% from the actual %delta, and thirty-two projects were predicted within ±5% from the actual %delta.

6.2.5. Correlation Accuracy

The relationship between the %Delta values that were predicted through the cross-validation process and the actual %Delta values can indicate the appropriateness of regression models. When the predicted values from the testing sets move in the same direction as the actual values, the regression models are considered appropriate for future use. In other words, when the predicted values have a strong positive trend with the actual values, the model is considered to be accurate. Table 29 shows that the correlation accuracy between the predicted %Deltas from the testing sets and actual %Deltas is high for both Model 1 and Model 2. The correlation accuracy for Model 1 is 72.49%, while the correlation accuracy for Model 2 is 72.12%. Such strong correlations verify the reliability and the reasonableness Model 1 and Model 2. It should be noted that through cross-validation, the correlation accuracy was calculated when each data point was considered as a new observation. Therefore, the correlation accuracy is one of the cross-validation metrics that indicate how well Model 1 and Model 2 can predict new observations.

6.2.6. CII Original Model

Although it is unfair to compare models generated from different sample sizes, it is important to mention the performance of the original CII model in relation to its own sample size (42 projects). The CII original model had a "Mean Absolute Percent Error (MAPE)" of 37.73%. Regarding the "%Delta Deviation Range", all forty-two projects were predicted within $\pm$24% from the actual %delta, thirty-nine projects were predicted within $\pm$15% from the actual %delta, thirty-three projects were predicted within $\pm$10% from the actual %delta, and eighteen projects were predicted within $\pm$5% of the actual %delta (Hanna 2001).

6.3. Predicting New Observations

Although cross-validation provides an adequate measurement of the future prediction accuracy of regression models, collecting and predicting new data is an ideal method for model validation. Accordingly, sixteen new projects were collected in order to verify the performance of the following three cumulative impact models: a) Model 1 (developed in this current study); b) Model 2 (developed in this current study); c) The CII original Model (Hanna 2001).
Table 31 shows the predicted %Delta values for the sixteen new projects for Model 1, Model 2, and the CII original Model, respectively. All three models provided decent performances, with almost the same average percent errors. As shown in Table 32, all three models had a MAPE around 15%. This means that on average, the predictions of the three models had a percent error of 15% of the actual %Delta values. Table 30 shows the values of all %Delta predictors for the

sixteen new projects. It should be noted that each of the three models uses a different set of predictors, as shown below.

Model 1: *%Delta = 0.33466 – 0.14550 %OwnerInitiatedCO – 0.08564 Productivity + 0.02182 Turnover – 0.08426 PM%TimeOnProject + 0.05611 Overmanning + 0.02426 Absenteeism*

Model 2: *%Delta = 0.359326 -0.144190 %OwnerInitiatedCO -0.088081 Productivity+ 0.030369 Turnover -0.082290 PM%TimeOnProject + 0.052126 Overmanning*

CII original Model: *%Delta= 0.36866 +0.11957 PercentChange -0.08065 PM%TimeOnProject -0.16723 %OwnerInitCO -0.09147 Productivity -0.05213 Overmanning +0.022345 ProcessingTime*

Table 30. Values of each variable for the new sixteen projects

No.	*%Change*	*PM%TimeOn Pr.*	*Productivity*	*Processing*	*%OwnerInitCO*	*Overmanning*	*Absentee*	*Turnover*
1	0.428	0.15	0	4	0.675	1	1	1
2	0.406	1	1	5	0.385	1	2	3
3	0.127	0.5	0	5	0.308	1	1	2
4	0.035	0.2	1	3	0.721	1	1	1
5	0.117	0.25	1	4	0.389	1	4	4
6	0.777	1	0	5	0	1	1	1
7	0.121	0.85	0	5	0	1	1	1
8	0.153	0.25	1	5	0	1	1	2
9	0.100	0.35	1	5	0.929	1	1	2
10	0.202	0.9	1	5	1	1	1	2
11	0.166	0.15	0	5	0	1	1	3
12	0.803	1	1	3	0.935	1	1	4
13	0.900	1	1	5	0.304	1	1	1
14	0.081	0.5	0	5	0.183	1	2	4
15	0.113	0.4	0	5	0	1	1	1
16	0.463	0.25	1	2	0	1	1	1

Table 31. Predicted %Deltas for the New 16 Observations

Project No.	Actual %Delta	Predicted %Delta (Model 1)	Predicted %Delta (Model 2)	Predicted %Delta (CII Original Model)
1	26.08%	32.59%	33.20%	33.20%
2	31.40%	27.88%	27.66%	24.03%
3	32.81%	37.17%	38.66%	35.17%
4	25.25%	22.94%	23.33%	15.96%
5	31.33%	41.17%	36.81%	24.32%
6	40.74%	35.26%	35.95%	44.06%
7	38.46%	36.52%	37.19%	37.42%
8	39.53%	35.20%	36.35%	33.49%
9	21.43%	20.83%	22.12%	16.50%
10	13.94%	15.17%	16.59%	12.11%
11	48.81%	46.79%	49.02%	43.61%
12	21.33%	19.64%	22.77%	15.11%
13	32.88%	22.27%	22.76%	31.28%
14	35.09%	45.77%	46.53%	36.70%
15	45.67%	40.31%	40.89%	40.95%
16	43.23%	33.01%	33.32%	30.50%

Table 32. Percent Errors for the New Predictions

Project No.	Percent Error Model 1	Percent Error Model 2	Percent Error CII Original Model
1	24.95%	27.31%	27.31%
2	11.22%	11.90%	23.47%
3	13.28%	17.82%	7.18%
4	9.14%	7.61%	36.80%
5	31.43%	17.52%	22.36%
6	13.46%	11.75%	8.15%
7	5.03%	3.30%	2.70%
8	10.97%	8.04%	15.28%
9	2.82%	3.24%	22.99%
10	8.79%	18.95%	13.14%
11	4.14%	0.44%	10.65%
12	7.95%	6.75%	29.15%
13	32.29%	30.80%	4.86%
14	30.43%	32.59%	4.58%
15	11.72%	10.46%	10.32%
16	23.64%	22.94%	29.45%
MAPE	**15.07%**	**14.46%**	**16.77%**

6.4. Summary

This chapter has thoroughly verified the predictive ability of the cumulative impact models using various statistical techniques. There are no signs of overfitting in Model 1 or Model 2. It was found that both of them are able to explain a generalized trend, rather than only explaining the examined sample. Different validation metrics as well as new observations showed that the newly developed models (Model 1 and Model 2) have high out-of-sample prediction accuracies.

Chapter 7. Conclusions and Recommendations

This research aimed at providing the mechanical and electrical construction industries with means for quantifying the cumulative impact of change orders on labor productivity. It is expected that the findings within this research would assist owners and contractors in improving the performance of their projects. The results of this research provide project stakeholders with alarming productivity loss indications during project execution. The models developed in this study (Model 1 and Model 2) can be proactively used during construction to assist contractors in understanding the potential impact of different project circumstances and factors on the final project outcome. In addition, owners and contractors can apply the quantification models after project close-out. The models aid in resolving legal issues and disputes related to productivity loss associated with the cumulative impact of changes.

7.1. Final Notes regarding the New Models (Model 1 and Model 2)

Among the conventional questions that anyone would ask after reading this thesis report are: "Which model should we use?", "which model is more reliable than the other?"

Both Model 1 and Model 2 are based on a fairly large sample size of 68 impacted projects. According to the comprehensive statistical analysis of this study, both Model 1 and Model 2 are reasonable and have very similar performances. However, Model 2 is more parsimonious than Model 1. Model 2 was able to explain the variability in %Delta with less predictors, which is quite desirable. A smart approach in linear regression is to develop less complex models with as few predictors as possible, while maintaining a satisfactory model performance. Model 1 depends on both "Turnover" and "Absenteeism" in predicting %Delta, while Model 2 only relies on "Turnover." Model 2 could be much closer to the reality than Model 1. Recalling section 5.1 in chapter 5, Model 2 was detected by using the stepwise-BIC criterion, which is a method that primarily looks for the true model. Model 1, however, was detected by using the stepwise-AIC criterion, which is a method that primarily looks for the model with the highest prediction accuracy, regardless whether it is close to the reality.

7.2. Improvements Achieved in this Study

Although the CII original model provides good predictions, the model was based on a smaller sample size as compared to this current study. Moreover, an elemental predictor (Overmanning) in the CII research had a counterintuitive regression coefficient sign that opposes priori knowledge and industry experience. This was interpreted that it was due to multicollinearity. Multicollinearity has serious consequences on regression models in that it affects the true values of the regression coefficients and inflates their variances. Multicollinearity is problematic as it saps the statistical power of the analysis. Multicollinearity results in models where regression coefficient signs are unstable and very sensitive to minor changes in the models. Multicollinearity also results in a problem where the sequential sum of squares (i.e. the amount of explanation in %Delta) associated with any variable depend on which variables are included in the model and which variables are omitted.

In this current study, the new models (Model 1 and Model 2) do not suffer from multicollinearity or their negative effects. All regression coefficient signs agree with priori knowledge. Besides the thorough statistical analysis of this research, an objective data-driven approach was adopted in order to define (Overmanning) based on the jobsite manpower ratios, by computing the ratio of actual peak manpower to actual average manpower.

7.3. Guideline for Model Usage

The models developed in this study can only be applied to projects impacted by change orders. Not every construction project with change orders is impacted; some projects can incorporate large amount of changes without encountering productivity loss. Consequently, before using the cumulative impact models of this study, several project conditions should be first met in order to ascertain whether a project is impacted by changes.

To quantitatively determine whether a project is impacted by changes, the impact logistic model developed in the CII original study (Hanna 2001) should be used (See Appendix B). When the probability of a project being impacted by changes is higher than 0.5, the project is said to have suffered from productivity loss due to change orders.

It should be noted that existing project data is the primary and the best source for quantifying the cumulative impact of change orders. A desirable approach is that owners and contractors track potential cumulative effects of changes throughout the course of the project and agree upon the amount of productivity loss resulting from changes. This would require unique circumstances that rarely occur in Design-Bid-Build projects, such as collaboration, mutual trust, and efficient and transparent communications between owners and contractors. If negotiations between owners and contractors fail, the models developed in this study could be used as substitute means for quantifying the incurred productivity loss. It should be noted that the models developed in this study should be used to support and verify the actual productivity loss calculations conducted by contractors.

7.4. Building a Linear Regression Model

Selecting a linear regression model and computing R^2 and R^2-adjusted are just part of the linear regression story. Building a regression model requires severe and thorough analysis. In this current study, regression models were carefully selected and diagnosed through different statistical methods and measures. Table 33 summarizes some of the noticeable statistical procedures that were adopted in this research throughout the model building process. It is recommended that analysts consider those procedures when building linear regression models.

Table 33. Elemental Statistical Checks in Linear Regression

Statistical Check	**Method/Measure used in this Research**
Constancy of Variance of the Error Terms	-Brown-Forsythe Test
Normality of the Error Terms	-Correlation Test for Normality
Multicollinearity	-Variance Inflation Factors -Condition Indices
Checking Outlying Y Observations	-Studentized residuals -Studentized deleted residuals
Checking Outlying X Observations	-The diagonal elements of the Hat Matrix (The leverage values)
Checking Influential Points	-DFFITS -DFBETAS -Cook's Distance

7.5. Recommendations to Owners and Contractors

Construction projects are affected by many factors and issues that determine their success. Therefore, different project parties should understand and control those factors in order to eliminate the likelihood of decreased labor productivity. This section will highlight practices for Owners and Contractors that are related to the productivity factors that were discussed throughout this thesis report.

7.5.1. Recommendations to Owners and Owner Representatives:

Owners can have a considerable control over many vital project aspects. Following are important recommendations for Owners and Owner representatives:

- **Project scope:** Owners should adequately define the scope of their projects through early constructability reviews in order to decrease the amount of change during construction. In this study, it was found that 39% of change orders for mechanical and electrical projects are due to additions. Moreover, it was found that increased amount of change is more likely to cause labor productivity loss. Impacted projects are characterized by higher amount of changes as compared to unimpacted projects.
- **Role of Architects and Engineers A/E's:** Owner representatives as well as Architects and Engineers should coordinate and solve design issues as early as possible in order to decrease the amount of design problems and insufficiencies that surface during construction. Moreover, adequate A/E's support and coordination between trades during construction helps preventing productivity loss. Also, A/E's should ensure an expeditious change order approval process and faster responses to Requests for Information (RFI's) throughout the course of the project. The longer the contractor waits for the owner's approvals, the higher the probability that a project will suffer from decreased labor productivity.
- **Time extensions:** It is recommended that owners compensate contractors for the amounts of time and cost that are sufficient enough to complete the change order work. Instead of forcing contractors to unfairly bear extra costs or use scheduling acceleration techniques (Overmanning, shiftwork, and overtime) to overcome the impact of changes, owners should owners should grant time extensions or compensate the contractors for the lost productivity that result from applying such schedule compression techniques.

7.5.2. Recommendations to Contractors:

Contractors' management decisions and work practices highly affect a project's outcome. In order to deliver the project with high quality and within its allotted cost and timeframe, the following practices are recommended to contractors:

- **Management Tools:** Manpower loading charts and S-curves are essential management tools that should be used in any project. Tracking the progress of projects and their manpower requirements trigger alarming signs to contractors and help them maintain a satisfactory project performance by taking corrective actions as early as possible. The manpower ratios calculated in this study confirm that when the ratio of estimated peak manpower to actual peak manpower decreases, projects are more likely to suffer from decreased labor productivity. Contractors should track their productivity by continuously measuring the percent of work completed and the actual installed quantities. The earned work hours should be also computed in order to monitor the progress of projects, and forecast the completion date and compare it to the estimated completion date. Contractors should also prepare and update a time schedule and integrate it with the manpower loading analysis and S-curves.
- **Cumulative Impact:** The expenses and work-hours attributable to the indirect impact of change orders on distant work should be reported by contractors in order to formally seek compensation from owners. Contractors can also incorporate allowances in their contracts in order to avoid waiving their rights for additional compensation.

7.6. Conclusions

This study quantified the cumulative impact of change orders on labor productivity for mechanical and electrical contractors. The quantification models of this research serve as verification means for ascertaining the additional work-hours attributable to productivity loss due to changes. Nevertheless, contractors should first rely on existing and actual project data when claiming additional compensation for the cumulative impact of change orders. The models of this study could be used to support the contractors' calculations and assertions during dispute resolutions. The results of this study also assist industry practitioners in understanding the underlying effects that change orders have on labor productivity. Finally, this research highlighted some of the major issues that analysts face when building linear regression models.

Appendix A. Statistical Notes

This appendix provides a detailed description regarding all the statistical terms, tests, and measures that were used throughout this thesis report. The statistical notes begin by describing basic terms and concepts regarding linear regression analysis before illustrating the different tests and measures that were used.

- **ith case:** Refers to a single data point or observation in the sample.
- **n:** Refers to the sample size or the number of observations
- **k:** Refers to the number of predictors or independent variables in a regression model
- **p:** Refers to the number of parameters in a regression model, where $p = k + 1$
- **Observed Value Y_i:** The actual value of %Delta for an ith case in the sample.
- **Fitted Value $\hat{Y}_i$:** The value of %Delta for an ith case obtained from an estimated regression model.
- **Residual e_i:** The deviation of the observed value Y_i from the fitted value $\hat{Y}_i$ on the "estimated regression line" for an ith case, where

$$e_i = Y_i - \hat{Y}_i$$

- **Error ε_i:** The error term of the model is different than the residual. The residual is known, while the error term is unknown. The error is the vertical deviation of the observed value Y_i from the "true regression line." The true regression line is unknown.
- **Total Sum of Squares (SSTO):** The sum of the squared deviations of each observed value Y_i from the overall mean $\bar{Y}_i$. When SSTO = 0, this means that all observed values Y_i are the same. It is a measure of the uncertainty in predicting %Delta when none of the predictors are used (Kutner et al. 2004).

$$SSTO = \sum_{i=1}^{n} (Y_i - \bar{Y}_i)^2$$

- **Error Sum of Squares (SSE):** The sum of the squared deviations of observed values $\boldsymbol{Y_i}$ from the fitted values $\boldsymbol{\hat{Y}_i}$. If SSE = 0, this means that all observations fall on the fitted line. The larger the variation of the observed values around the regression line, the larger SSE will be (Kutner et al. 2004).

$$SSE = \sum_{i=1}^{n} (Y_i - \hat{Y}_i)^2 = \sum_{i=1}^{n} e_i{}^2$$

- **Regression Sum of Squares (SSR):** The sum of the squared deviations of fitted values $\boldsymbol{\hat{Y}_i}$ from the the overall mean $\boldsymbol{\bar{Y}_i}$. It is a measure of the variability of %Delta associated with the regression model (Kutner et al. 2004).

$$SSR = \sum_{i=1}^{n} (\hat{Y}_i - \bar{Y}_i)^2$$

- **Variance of the error terms $\boldsymbol{\sigma^2}$:** The variance σ^2 of the error terms ε_i is unknown. However, it can be estimated using the Error Mean Square (MSE).
- **Error Mean Square (MSE):** It is an unbiased estimator of σ^2. It refers to SSE divided by its associated degrees of freedom $n - p$.

$$MSE = \frac{SSE}{n - p}$$

- **Regression Mean Square (MSR):** Refers to SSR divided by its associated degrees of freedom $p - 1$.

$$MSR = \frac{SSR}{p - 1}$$

- **Coefficient of Multiple Determination R^2:** Refers to the proportionate reduction of the total variation in %Delta associated with the use of k predictors. R^2 ranges from 0 to 1. The larger R^2, the more the total variation in %Delta is decreased by using k predictors (Kutner et al. 2004).

$$R^2 = \frac{SSR}{SSTO} = 1 - \frac{SSE}{SSTO}$$

- **Adjusted Coefficient of Multiple Determination (R^2-Adjusted):** Due to the fact that R^2 can never decrease as p increases, R^2-Adjusted is as an alternative measure of R^2 that accounts for that issue by considering the number of parameters of the regression model. Unlike R^2, R^2-Adjusted can decrease as p increases (Kutner et al. 2004).

$$R^2 - Adjusted = 1 - \left(\frac{n-1}{n-p}\right)\left(\frac{SSE}{SSTO}\right)$$

- **Mallows' Cp Criterion:** It is a criterion used to find the best subset regression model among a set of possible models having different number of parameters. The Cp is often used as a model selection criterion in order to address the issue of overfitting. A desirable Cp should be small in value, and close to the number of parameters p. The Mallows' Cp Criterion takes into account the total error in each fitted value for each subset regression model. The *"total error"* for the ith fitted value $\hat{Y}_i$ is made up of two components: a *"bias"* component for the ith fitted value $\hat{Y}_i$ and a *"random error"* component for the ith fitted value $\hat{Y}_i$ (Kutner et al. 2004).

$$Total\ Error = Bias + Random\ Error$$

In order to understand the "total error in each fitted value", the "bias component in each fitted value", and the "random error in each fitted value", some important terms have to be first defined:

$\boldsymbol{\mu_i}$: The true unknown value of the mean response for the ith case

$\boldsymbol{E\{\hat{Y}_i\}}$: The expected fitted value for the ith case, which is obtained by fitting a regression model to all possible samples.

- The bias component (or the model error component) is the difference between the expected fitted value for the ith case $\boldsymbol{E\{\hat{Y}_i\}}$ and the true mean response $\boldsymbol{\mu_i}$.
- The random error component is the difference between the fitted value $\boldsymbol{\hat{Y}_i}$ and the expected fitted value for the ith case $\boldsymbol{E\{\hat{Y}_i\}}$
- The total error is the sum of the bias component and the random error component for the ith case
- The mean squared error for $\boldsymbol{\hat{Y}_i}$ is the expected value of the square of the total error.
- The total mean squared is the sum of the mean squared errors of all n observations or fitted values
- The Mallows' criterion is measured by dividing the total mean squared error for the fitted values by the true error variance σ^2

As we don't know the true error variance σ^2, it is replaced by its unbiased estimator MSE. According to (Kutner et al. 2004), the Mallows' Cp can be calculated using the following equation:

$$Cp = \frac{SSE\ for\ the\ subset\ regression\ model}{MSE\ for\ the\ model\ with\ all\ possible\ predictors} - (n - 2p)$$

where p is the number of parameters for the best subset regression model.

- **Akaike's Information Criterion (AIC):** A method for providing penalties for adding more variables into the regression model. A desirable model has the smallest AIC.

$$\boldsymbol{AIC = n\, ln(SSE) - n\, ln(n) + 2p}$$

- **Bayesian Information Criterion (BIC):** A method for providing penalties for adding more variables into the regression model. A desirable model has the smallest BIC. The BIC criterion provides higher penalty as compared to the AIC criterion when the sample size n is greater or equal to 8. Consequently, the BIC criterion selects more parsimonious models (i.e. models with less predictors) as compared to the AIC criterion.

$$\boldsymbol{BIC = n\, ln(SSE) - n\, ln(n) + [ln(n)]p}$$

- **Prediction Sum of Squares (PRESS) criterion:** It is a validation method for measuring the precision of regression models. PRESS resembles to SSE in that it accounts for the sum of the errors. However, PRESS considers each observation as a new observation. PRESS is calculated by deleting an ith case in the dataset, estimating the regression model using the remaining cases, and then predicting %Delta for the ith case using the estimated regression model (Kutner et al. 2004). This procedure is repeated for all n observations, and the sum of the squared prediction errors is calculated as follows:

$$\boldsymbol{PRESS = \sum_{i=1}^{n} (Y_i - \hat{Y}_{i(i)})^2}$$

where $\boldsymbol{\hat{Y}_{i(i)}}$ is the predicted %Delta for the ith case, after omitting the same ith case, using the estimated regression model from the remaining $n-1$ cases.

- **Outlying X observations:** Refers to outlier cases in regard to the predictors
- **Outlying Y observations:** Refers to outlier cases in regard to the response variable
- **Fitted Linear Regression Model in the Matrix Form:**

$$\boldsymbol{\hat{Y}_{n\times 1} = X_{n\times p}\, b}$$

where:

$$\hat{\boldsymbol{Y}}_{\boldsymbol{n}\times\boldsymbol{1}} = \boldsymbol{vector\ of\ the\ fitted\ values} = \begin{bmatrix} \hat{\boldsymbol{Y}}_{\boldsymbol{1}} \\ \hat{\boldsymbol{Y}}_{\boldsymbol{2}} \\ \vdots \\ \vdots \\ \hat{\boldsymbol{Y}}_{\boldsymbol{n}} \end{bmatrix}$$

$$\boldsymbol{X}_{\boldsymbol{n}\times\boldsymbol{p}} = \boldsymbol{Design\ matrix} = \begin{bmatrix} \boldsymbol{1} & \boldsymbol{X}_{\boldsymbol{11}} & \cdots & \boldsymbol{X}_{\boldsymbol{1,p-1}} \\ \boldsymbol{1} & \boldsymbol{X}_{\boldsymbol{21}} & \cdots & \boldsymbol{X}_{\boldsymbol{2,p-1}} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{1} & \boldsymbol{X}_{\boldsymbol{n1}} & \cdots & \boldsymbol{X}_{\boldsymbol{n,p-1}} \end{bmatrix}$$

$$\boldsymbol{b} = \boldsymbol{vector\ of\ the\ estimated\ regression\ coefficients} = \begin{bmatrix} \boldsymbol{b}_{\boldsymbol{1}} \\ \boldsymbol{b}_{\boldsymbol{2}} \\ \vdots \\ \vdots \\ \boldsymbol{b}_{\boldsymbol{p-1}} \end{bmatrix}$$

<u>Notes:</u>

- The Design matrix includes the values of all $p-1$ predictors for each of the n observations. It also includes a column of 1 to account for the intercept of the regression model (Kutner et al. 2004).
- $\hat{\boldsymbol{Y}}_{\boldsymbol{1}}$ is the fitted %Delta value for the first case/observation.
- $\boldsymbol{X}_{\boldsymbol{21}}$ is the value of the first predictor for the second case/observation.
- $\boldsymbol{b}_{\boldsymbol{1}}$ is the estimated regression coefficient of the first predictor.
- There are k or $p-1$ predictors in a given fitted regression model.

- **Hat Matrix $\boldsymbol{H}$:** A fitted %Delta value (or Y hat) can be expressed as a linear combination of all actual %Delta values through the following equation (Kutner et al. 2004):

$$\hat{\boldsymbol{Y}}_{\boldsymbol{n}\times\boldsymbol{1}} = \boldsymbol{HY}$$

where $\boldsymbol{Y}$ is the vector of actual %Delta values and $\boldsymbol{H}$ is a square matrix called the "Hat Matrix", computed by using the Design Matrix $\boldsymbol{X}$ as follows:

$$\boldsymbol{H}_{\boldsymbol{n}\times\boldsymbol{n}} = \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-\boldsymbol{1}}\boldsymbol{X}'$$

<u>*Notes:*</u>

- The diagonal elements of the Hat matrix range from 0 to 1.
- The diagonal elements of the Hat matrix are called the "leverage values." They denote the weights of the actual %Delta values in determining the fitted value for an ith case (Kutner et al. 2004).
- There are n leverage values in the Hat matrix, corresponding to each of the n cases/observations
- The leverage value of the ith case is denoted by $\boldsymbol{h_{ii}}$ (i.e. The diagonal element of the ith row and ith column of the Hat matrix).
- The higher the leverage value for an ith case, the greater the distance between the X values for the ith case and the means of all X values for all n observations (Kutner et al. 2004).
- The leverage values in the hat matrix play an important role in detecting outlying X observations
- The sum of the diagonal elements of the Hat matrix equals to the number of parameters p

- **Residuals expressed in the Matrix Form:**

$$\boldsymbol{e}_{\boldsymbol{n}\times\boldsymbol{1}} = \boldsymbol{Y} - \widehat{\boldsymbol{Y}} = \boldsymbol{Y} - \boldsymbol{X}\boldsymbol{b} = (\boldsymbol{I} - \boldsymbol{H})\boldsymbol{Y}$$

Where $\boldsymbol{I}$ is the identity square matrix with size $n\times n$.

- **Hypothesis Test:** In statistical analysis, a hypothesis test consists of two components: a null hypothesis and an alternative hypothesis. A null hypothesis is an assumption that an analyst thinks that it should be rejected or nullified. On the contrary, an alternative hypothesis is the assumption that an analyst thinks it should hold.
- **Test statistic:** It is a standardized value that determines whether the null hypothesis should be rejected.
- **P-value:** It is a measure of the statistical significance in a hypothesis test. Assuming that the null hypothesis is true, the P-value denotes the probability of obtaining a result equal to or more extreme than what was actually observed. The smaller the P-value, the larger the statistical significance in a hypothesis test.
- **α level of risk (or level of significance):** The probability of rejecting the null hypothesis given that it is true. Analysts usually use an α level of risk of 0.05. The null hypothesis is rejected when the P-value is smaller that the chosen level of risk.
- **Normal Distribution:** A probability density function that represents the distribution of a random variable as a symmetrical bell-shaped graph. The center of the graph is the mean of the distribution, and the height of the graph is determined by the standard deviation of the distribution. The larger the standard deviation, the wider and shorter is the bell-shaped graph.
- **Standard Normal Distribution:** The simplest form of a normal distribution. It has a mean of zero, and a standard deviation of 1.
- **Student's t-distribution:** It is a distribution of a random variable when the sample size is small and population standard deviation is unknown. The Student's t −distribution approaches a normal distribution when the sample size increases. The Student's t-

distribution is used to conduct two-sample t-tests in order to examine whether two population means are equal.

- **F-distribution:** A right-skewed probability density function distribution commonly used in the Analysis of Variance. It is a function of the ratio of two independent random variables, each of which has a chi-square distribution and is divided by its number of degrees of freedom. The F-distribution is used to conduct the F-tests. The F-tests are used to compute test statistics in order to compare between different statistical models and check whether a given null hypothesis holds. The F-tests are also used within the Analysis of Variance (ANOVA) table.
- **Chi-square χ^2-distribution:** The Chi-square distribution is the distribution of the sum of squares of independent standard normal random variables. The degrees of freedom of the distribution equals to the number of standard normal random variables being summed.
- **Analysis of Variance (ANOVA):** The ANOVA table discussed in chapter 5 determines the impact that the predictors have on %Delta. It computes the sequential sum of squares for each predictor. The sequential sum of squares for a given predictor is the additional variation that each predictor explains, given that other predictors are already included in the model. The sequential sum of squares can be also thought of as:
    - The reduction in SSE when one or more predictors are added to the model.
    - The increase in SSR when one or more predictors are added to the model.
- **Brown-Forsythe Test:** A test for checking the assumption of constancy of variance of the error terms, even when the error terms are not normally distributed. The test can be separately conducted to each predictor in a regression model. The data is divided into two equally sized groups according to the levels of the subject predictor, where the first group

has relatively low values for the predictor, and the second group has relatively high values for the predictor. For each of the two groups, the test computes the absolute deviations of the residuals around the median of the group. The test then uses the two-sample t test to determine whether the mean of the absolute deviations is the same for the two groups. By controlling the α level of risk at a preferred level and setting the degrees of freedom to $n - 2$, the t distribution can be used to conduct a two-sample t test and obtain a two-sided P-value. If the P-value is higher than the chosen α level of risk, the variance of the error terms is considered constant across the two groups (Kutner et al. 2004).

- **Correlation Test for Normality:** A test conducted to check the assumption of the normality of the error terms. The coefficient of correlation between the residuals and their expected values under normality is compared to a certain threshold. If the coefficient of correlation is higher than the threshold, the error terms can be considered normally distributed. The threshold can be obtained from the values prepared by (Looney and Gulledge), which depend on the sample size and the α level of risk.

  After sorting the residuals from lowest value to highest value, the expected value under normality of the jth smallest residual can be obtained using the following equation (Kutner et al. 2004):

$$(\sqrt{MSE})\left[z\left(\frac{j - 0.375}{n + 0.25}\right)\right]$$

  where:

  j: the rank of the jth smallest residual (after sorting the residuals from lowest value to highest value).

  n: sample size

z: value corresponding to the percentile of the jth smallest residual in the standard normal distribution

Note: Two assumptions are made in order to get the expected values of the residuals under normality: a) the expected value of the error terms is zero; b) the standard deviation of the error terms is $\sqrt{MSE}$ (Kutner et al. 2004).

- **Expected value and Variance of the errors:** The errors follow a normal distribution with mean zero and variance σ^2. The errors are independent from one another.
- **Expected value and Variance of the residuals:** The expected value of the residuals is zero. The standard deviation of the residual $\boldsymbol{e_i}$ equals to $\sqrt{MSE(1-h_{ii})}$. The variance of the residuals is not constant, and depends on $\boldsymbol{h_{ii}}$. The residuals are not independent terms (Kutner et al. 2004).
- **Studentized Residuals $\boldsymbol{r_i}$:** The studentized residuals $\boldsymbol{r_i}$ are refinements to the ordinary residuals $\boldsymbol{e_i}$**.** They account for the issue that the ordinary residuals have different variances ($\sigma_i{}^2$). The studendized residual of an ith case is computed by dividing the ordinary residual of the ith case $\boldsymbol{e_i}$ by its standard deviation (Kutner et al. 2004).

$$\boldsymbol{r_i} = \frac{\boldsymbol{e_i}}{\boldsymbol{s\{e_i\}}}$$

where:

$\boldsymbol{s\{e_i\}}$: The standard deviation of the ith residual, estimated by $\sqrt{MSE(1-h_{ii})}$

$\boldsymbol{h_{ii}}$: Leverage value of the ith case, obtained from the diagonal elements of the Hat matrix.

- **Studentized Deleted Residuals $\boldsymbol{t_i}$:** A second refinement to ordinary residuals is to compute the ith residual when the regression model is fitted on all observations except the ith one (i.e. the ith one is deleted). This procedure is more likely to disclose outlying

Y observations, as discussed in Chapter 5. Fortunately, the studentized deleted residual t_i for the ith case can be computed without having to fit new regression functions each time an observation is deleted using the following equation (Kutner et al. 2004):

$$t_i = e_i \left[\frac{n-p-1}{SSE(1-h_{ii}) - e_i}\right]^{1/2}$$

- **DFFITS:** A measure of the influence of an ith case on a single fitted value. It is based on the difference between the fitted value for the ith case when all 68 cases are used in fitting the regression equation and the fitted value when the ith case is excluded from the fitting process. DFFITS can be computed for each observation without having to fit new regression functions each time an observation is deleted, using the following equation (Kutner et al. 2004):

$$(DFFITS)_i = e_i \left[\frac{n-p-1}{SSE(1-h_{ii}) - e_i}\right]^{1/2} \left(\frac{h_{ii}}{1-h_{ii}}\right)^{1/2} = t_i \left(\frac{h_{ii}}{1-h_{ii}}\right)^{1/2}$$

- **Cook's Distance D_i:** Unlike DFFITS, Cook's distance is a measure of the influence of an ith case on all fitted values. It also involves the procedure of omitting observations and examining their influence. Cook's distance can be easily calculated without fitting a new regression equation each time a case is omitted through the following equation (Kutner et al. 2004):

$$D_i = \frac{e_i^2}{p\ MSE}\left[\frac{h_{ii}}{(1-h_{ii})^2}\right]$$

- **DFBETAS:** It is an indicator of the influence of an ith case on each regression coefficient in the regression function. It is based on examining the difference between the estimated regression coefficient based on all n cases and the estimated regression coefficient obtained after deleting the ith case (Kutner et al. 2004).

- **Variance Inflation Factor (VIF):** VIF is a multicollinearity diagnostic. It is a measure of how much the variances of the regression coefficients are inflated as compared to when the variables are uncorrelated. By regressing a given predictor (X_k) on all other predictors, the coefficient of multiple determination R_k^2 can be obtained and VIF_k can be calculated through the following equation (Kutner et al. 2004):

$$VIF_k = (1 - R_k^2)^{-1}$$

  A rule of thumb is when a VIF is greater than 10, multicollinearity is said to affect the estimated regression coefficients.

- **Condition Index $\emptyset_j$:** The condition index is another multicollinearity diagnostic. It is based on computing the eigenvalues λ of the matrix $\boldsymbol{X'X}$. When there is strong linear dependence in the column vectors of the design matrix, at least one of the eigenvalues of the matrix $\boldsymbol{X'X}$ will be close to zero. The condition indices are computed using the eigenvalues λ of the matrix $\boldsymbol{X'X}$. A rule of thumb is when the Condition Index is in the range of 30 to 100, there is evidence of moderate to severe multicollinearity. Following is the equation used to compute the Condition Index $\emptyset_j$:

$$\emptyset_j = \sqrt{\frac{\lambda_{max}}{\lambda_j}} \text{ for } j = (1,2, \dots, p)$$

  where:

  λ_j: The jth eigenvalue

  λ_{max}: The maximum eigenvalue

Appendix B. Screening New Projects

As stated in chapter 6, sixteen new projects were collected in order to validate the performance of the cumulative impact models. Although the new projects were primarily impacted by changes, the CII logistic regression impact model (Hanna 2001) was used as a screening criterion to illustrate the prerequisite step that should be met before using the %Delta predictive models. The CII logistic regression impact model (Hanna 2001) is able to accurately determine the probability of a project being impacted by changes. A project with a probability higher than 50% is considered to be impacted by changes, and is qualified to proceed to a further step. The further step consists of predicting %Delta using the cumulative impact linear regression models. Table 34 shows the results of the quantitative screening procedure applied to the sixteen new projects. As was expected, all sixteen projects were found to be impacted by change orders. Below is the logistic impact model adopted from (Hanna 2001):

- *Sum of Factors = - 6.997 - 1.0939 MorE + 3.889 PerChange*MorE - 1.0371Estimated Peak/Actual Peak+ 0.6342 ProcessingTime+2.6433 Overmanning+1.1933 Overtime + 1.2048 Peak/Average Manpower + 0.017154 Percent Change Orders related to Design Issues*100*
- *Probability (Impact)* $= \frac{e^{Sum\ of\ Factors}}{1+ e^{Sum\ of\ Factors}}$

Table 34. Screening the new projects

%Change	MorE	EA_P	Process	P/A	Overman -ning	PerDesign	Overti -me	Sum of Factors	Probability (Impact)
0.428	0	0.33	4	3.88	1	0.5	1	4.566481	98.97%
0.406	0	0.53	5	2.02	1	0.4	1	2.580793	92.96%
0.127	0	0.48	5	2.4	1	0.95	0	2.840642	94.48%
0.035	0	0.46	3	3.15	1	0.4	1	2.746414	93.97%
0.117	0	0.87	4	1.69	1	0.5	1	1.367935	79.70%
0.777	0	0.43	5	2.96	1	0.2	1	3.473935	96.99%
0.121	0	0.4	5	3.02	1	0.75	1	4.520806	98.92%
0.153	0	0.57	5	2.31	1	0.5	1	3.060241	95.52%
0.100	0	0.32	5	5.14	1	1	1	7.5868	99.95%
0.202	0	0.27	5	4.71	1	0.4	1	6.091351	99.77%
0.166	0	0.42	5	2.42	1	0.8	0	2.669654	93.52%
0.803	0	0.4	3	2.15	1	1	1	2.63308	93.30%
0.900	0	0.22	5	3.06	1	0.15	1	3.726436	97.65%
0.081	0	0.68	5	2.04	1	0.3	1	2.277784	90.70%
0.113	0	0.39	5	3.17	1	0.65	1	4.540357	98.94%
0.463	0	0.36	2	2.73	1	0.55	1	1.967218	87.73%

Appendix C. Bibliography

AbouRizk, H., Lee, S., Gellatly, I., & Fayek, A. (2010). Understanding Withdrawal Behavior in the Construction Industry. In Construction Research Congress 2010 (pp. 809-816).

Ahn, S., Lee, S., & Steel, R. P. (2013). Construction workers' perceptions and attitudes toward social norms as predictors of their absence behavior. Journal of Construction Engineering and Management, 140(5), 04013069.

Allen, W. (1979). "Developing the project plan." Notes, Engineering Institute of Canada Annual Congress Workshop, Toronto, Ont., pp. 3-9.

Allmon, E., Hass, C. T., Borcherding, J. D., and Goodrum, P. M. (2000). "U.S. construction labor productivity trends, 1970–1998." J. Constr. Eng. Manage., 10.1061/(ASCE)0733-9364(2000)126:2(97), 97–104.

American Association of Cost Engineers (AACE). (2013). AACE international recommended practice 10S-90, cost engineering terminology, AACE International, Morgantown, WV.

Belsley, D. A. (1991). Conditioning diagnostics. John Wiley & Sons, Inc.

Business Roundtable (1982). "Absenteeism and Turnover." Report C-6. BRT, June 1982.

Business Roundtable (1988). "Construction Industry Cost Effectiveness (CICE): The next five years and beyond", Business Round Table, New York.

Chan, A. P. C., & Chan, A. P. L. (2004). Key performance indicators for measuring construction success. Benchmarking: An International Journal, 11(2), 203-221. http://dx.doi.org/10.1108/14635770410532624

Coburn, E. (1997). “Shiftworker fatigue: The $77 billion problem.” Cost Eng., AACE, 39(4), 26–28.

Finke, M. R. (1997). Claims for construction productivity losses. Public Contract Law Journal, 311-338.

Hanna, A. S., Russell, J. S., Gotzion, T. W., and Nordheim, E. V. (1999a). “Impact of change orders on labor efficiency for mechanical construc- tion.” J. Constr. Eng. Manage., 10.1061/(ASCE)0733-9364(1999) 125:3(176), 176–184.

Hanna, A. S., Russell, J. S., Gotzion, T. W., and Vandenberg, P. J. (1999b). “The impact of change orders on mechanical labor efficiency.” Constr. Manage. Econ., 17(6), 721–731.

Hanna, A. S. (2001). “Quantifying the cumulative impact of change orders for electrical and mechanical contractors.” Research Report 158– 11, Construction Industry Institute, Univ. of Texas at Austin, Austin, Tex.

Hanna, A. S., Menches, C. L., Sullivan, K. T., & Sargent, J. R. (2005). “Factors affecting absenteeism in electrical construction.” Journal of construction engineering and management, 131(11), 1212-1218.

Hanna, A. S., Taylor, C. S., and Sullivan, K. T. (2005). “Impact of extended overtime on construction labor productivity.” J. Constr. Eng. Manage., 10.1061/(ASCE)0733-

9364(2005)131:6(734), 734–739.

Hanna, A. S., Chang, C., Lackney, J. A., & Sullivan, K. T. (2005). Overmanning impact on construction labor productivity. In Construction Research Congress 2005: Broadening Perspectives (pp. 1-10).

Hanna, A., Chang, C., Sullivan, K., and Lackney, J. (2008). "Impact of shift work on labor productivity for labor intensive contractor." J. Constr. Eng. Manage., 10.1061/(ASCE)0733-9364(2008)134:3(197), 197–204.

Haring, D. B. (1981). "Productivity-thorough shift work." District 2, National Electrical Contractors Association.

Hester, W. T., Chang, T. C., & Kuprenas, J. A. (1991). Construction changes and change orders: their magnitude and impact. Construction Industry Institute.

Hinze, Jimmie. (1985) "Absenteeism in Construction Industry." Journal of Management in Engineering, Oct. 1985, v1, no.4, pp188-200.

Howerton, J. (1969). "Do you know the hidden costs of overtime?" Qualified Contractor, (Mar.).

Ibbs, W. (2005). "Impact of change's timing on labor productivity." J. Constr. Eng. Manage., 10.1061/(ASCE)0733-9364(2005)131:11 (1219), 1219–1223.

Ibbs, W., and Liu, M. (2005). "Improved measured mile analysis technique." J. Constr. Eng. Manage., 10.1061/(ASCE)0733-9364(2005) 131:12(1249), 1249–1256.

Ibbs, W. (2008). The Cumulative Impact of Change on Construction Labour Productivity.

Jarkas, A. (2010). “Critical investigation into the applicability of the learn- ing curve theory to rebar fixing labor productivity.” J. Constr. Eng. Manage., 10.1061/(ASCE)CO.1943-7862.0000236, 1279–1288.

Jones, R. M. (2001). Lost productivity: Claims for the cumulative impact of multiple change orders. Public Contract Law Journal, 1-46.

Jones, R. M., & Driscoll, T. J. (2002). Cumulative Impact Claims.

Kutner, M. H., Nachtsheim, C., & Neter, J. (2004). Applied linear regression models. McGraw-Hill/Irwin.

Lee, M. J., Hanna, A. S., and Loh, W. Y. (2004). “Decision tree approach to classify and quantify cumulative impact of change orders on productiv- ity.” J. Comput. Civ. Eng., 10.1061/(ASCE)0887-3801(2004)18:2 (132), 132–144.

Leonard, C. A. (1988). The effects of change orders on productivity (Doctoral dissertation, Concordia University).

Looney, S. W., & Gulledge Jr, T. R. (1985). Use of the correlation coefficient with normal probability plots. The American Statistician, 39(1), 75-79.

Lu, W., Shen, L., & Yam, M. C. (2008). Critical success factors for competitiveness of contractors: China study. Journal of construction engineering and management, 134(12), 972-982.

Means, R. S. (2010). Means Illustrated Construction Dictionary.

Moselhi, O., Charles, L., and Fazio, P. (1991a). "Impact of change orders on construction productivity." Can. J. Civ. Eng., Vol. 18,pp 484–492.

Overtime and productivity in electrical construction. (1969). National Electrical Con- tractors Association, Washington, D.C.

Penkala, D. (1997). "Improving productivity and profitability around the clock." Natl. Prod. Rev., 16(3), 29–35.

Pittman Construction Co., GSBCA Nos. 4897, 4923,81-1 BCA 14,847

Revay, S. O. (2002). Coping with extras. The Revay Report, 21(2), 1-4.

Schwartzkopf, W., McNamara, J. J., and Hoffar, Julian F. (1992). Calculating construction damages. Aspen Publishers,Inc.

Schwartzkopf, W., McNamara, J. J., and Hoffar, Julian F. (1997). Calculating construction damages. Aspen Publishers,Inc.

Salehi Sichani, M., Lee, S., & Robinson Fayek, A. (2011). Understanding construction workforce absenteeism in industrial construction. Canadian Journal of Civil Engineering, 38(8), 849-858.

Sargent, J., Sullivan, K. T., & Hanna, A. S. (2003, March). "Absenteeism and turnover impact on labor productivity for electrical contractors." In ASCE Construction Research Congress. Honolulu, Hawaii. March.

Statista (2016). "Construction Industry - Statistics & Facts."

Sullivan, Kenneth (2004). "Quantification of the Cumulative Impact of Change Orders on Sheet Metal Labor Productivity." (Doctoral dissertation, University of Wisconsin-Madison)

Sweet, Justin. (1994). Legal Aspects of Architecture, Engineering and the Construction Process, Fifth Edition, West Publishing Company, St. Paul, MN.

Teicholz, P. (2000). "Productivity trends in the construction industry." AISC Annual Convention, AISC, Chicago.

"The effects of scheduled overtime and shift schedule on construction craft productivity." (1988). Source Document 43, Construction Industry Institute, Austin, Tex. The four-ten hour/day work week. (1979). Daniel International Construction Company, Inc., Greenville, S.C.

Thomas, H. R. (1992). Effects of scheduled overtime on labor productivity. Journal of Construction Engineering and Management, 118(1), 60-76.

Thomas, H. R., & Raynar, K. A. (1997). Scheduled overtime and labor productivity: quantitative analysis. Journal of Construction Engineering and Management, 123(2), 181-188.

Thomas, H. Randolph and Napolitan, Carmen L. (1995). "Quantitative Effects of Construction Changes on Labor Productivity." Journal of Construction Engineering and Management., Vol. 121, No. 3, pp. 290-296.

U.S. Army Corps of Engineers. (1979). "Modification impact evaluation guide" EP 415-1-3, Dept. of the Army, Washington, D.C.

Wideman, Max. (1994) "A Pragmatic Approach to Using Resource Loading, Production, and Learning Curves on Construction Projects" Canadian Journal of Civil Engineering, Vol. 21

Yi, W., & Chan, A. P. (2013). Critical review of labor productivity research in construction journals. Journal of management in engineering, 30(2), 214-225.