Modeling and Benchmarking Performance for the Integrated Project Delivery (IPD) System

By

Mounir El Asmar

A dissertation submitted in partial fulfillment of
the requirements for the degree of

Doctor of Philosophy
(Civil and Environmental Engineering)

at the
UNIVERSITY OF WISCONSIN – MADISON
2012

Date of final oral examination: July 30, 2012

The dissertation is approved by the following members of the Final Oral Committee:

Awad S. Hanna, Professor, Civil and Environmental Engineering
Jeffrey S. Russell, Professor, Civil and Environmental Engineering
John S. Nelson, Adjunct Professor, Civil and Environmental Engineering
Carol C. Menassa, Assistant Professor, Civil and Environmental Engineering
Wei-Yin Loh, Professor, Statistics
Wafik Lotfallah, Visiting Professor, Mathematics

I dedicate this dissertation to my loving family, to whom I am forever thankful: Maurice, Noelle, José, Sandra, and Lucien. I am grateful for the invaluable education they provided me, and I am blessed with their endless love, unwavering support, and continuing inspiration.

Acknowledgements

My doctoral journey has been an unforgettable experience that will forever shape who I am as a person. I could not have completed this journey without the guidance and support of many individuals, including my advisor, committee members, colleagues, friends, and industry professionals.

Professor Awad Hanna provided steadfast support and invaluable advice throughout the past several years. He instilled in me hard work and determination, and taught me how to become a confident and comprehensive researcher. I am truly thankful to him for inspiring me to pursue this career.

I wish to thank each Ph.D. committee member who offered invaluable recommendations and guidance. Their constructive feedback facilitated the completion of this work. Specifically, I would like to thank Professor Jeffrey Russell for sharing his distinctive outlook regarding the broader framework of my research and providing key advice to maximize the contributions of this work to both academia and industry. Through his statistical expertise, Professor Wei-Yin Loh provided instrumental assistance from a non-engineering perspective, which was extremely helpful in the development and interpretation of results. Professor John Nelson’s unique systems thinking abilities and extensive industry experience assisted in framing the motivation of this work, while also contributing to key aspects, such as formalizing the *3Ts* framework. In addition to pushing me to explore each facet further, Professor Carol Menassa consistently provided insightful advice and thorough reviews that were essential to finalizing this work. I also would like to thank Professor

Wafik Lotfallah for his precious assistance with the development of the project quarterback rating performance model.

Through countless hours of reviewing, Mr. Gary Whited enhanced the quality and flow of the document as a whole. Mr. John Stevenson from the University of Wisconsin Survey Center provided valuable feedback on the questionnaire to maximize the usability of collected data. I also wish to acknowledge the support and funding provided by the Construction and Materials Support Center (CMSC) at the University of Wisconsin-Madison in the preparation of this study.

Several University of Wisconsin–Madison graduate students and friends also provided support in various ways, including reviewing this work and volunteering insightful feedback during research seminars. In particular, I would like to thank Mr. Elie Azar and Mr. Matthew Iwanski. A very special thank you is extended to my best friend Ms. Martina Gast. Her selfless love and support has remained steady, and I am deeply grateful for her persistent recommendations and advice from a non-construction perspective.

I am equally grateful for each one of the industry collaborators who took the time and effort to participate in the development stage, and engage in interviews and surveys. Throughout the course of this research, I met several accomplished industry leaders with who I hope to have the opportunity to collaborate again. The supportive industry collaborators include the Boldt Company, DPR Construction, Mortenson Construction, J.H. Findorff & Son Inc., Tocci Building Companies, Alberici Constructors, BMW Constructors Inc., the

Erdman Company, J. P. Cullen and Sons Inc., and CG Schmidt Inc. Specific individuals I would like to thank include:

- Mr. Jeffrey Niesen
- Mr. Gus Schultz
- Mr. Pete Selmo
- Mr. Blaine Tuchscherer
- Mr. David Thomack
- Mr. Jay Harris
- Mr. Josh Bronitsky
- Mr. Brian Boecker
- Mr. Daniel Loosbrock
- Mr. Jim Yehle
- Mr. John Tocci
- Mr. Timothy M. Gunn
- Mr. Jim Schumacher
- Mr. Jeff Kenley
- Mr. Mikelis Abuls
- Mr. Mark Rounds
- Mr. John K. Huggett
- Mr. Trent Jezwinski
- Ms. Shelly Peterson
- Mr. James McConachie
- Mr. Daniel Obara
- Mr. Jeffrey Madden
- Mr. Breck Dane
- Mr. Craig Wacker
- Mr. Eric Plautz
- Ms. Laura Handler
- Mr. Dave Koester
- Mr. Bob Klaas
- Mr. David Albrecht
- Mr. Pat Loughrin
- Mr. Michael Radtke
- Mr. Tim London
- Mr. Brian Lubnow
- Mr. Eric Swanlund
- Ms. Elke O'Neill
- Mr. Robert Nartonis
- Mr. Mark Tomsyck
- Mr. Rob Weise
- Mr. Jason Mattila
- Ms. Amy T. West
- Mr. John Hetland
- Mr. Davis Baran
- Mr. Dan Dherrwaechter

All of these committed individuals constantly strive to enhance the construction industry. Without their contributions of data, this study would not have been possible.

Finally, I would like to extend my gratitude to all the individuals not specifically listed who also contributed to my research and the completion of this study. I feel very fortunate to be surrounded by such a distinguished group of people. It is my hope that this dissertation demonstrates my gratefulness to all those involved.

Abstract

Integrated Project Delivery (IPD) is an emerging construction project delivery system that collaboratively involves key participants very early in the project timeline, often before the design is started. It is distinguished by a multiparty contractual agreement that typically allows risks and rewards to be shared among project stakeholders. As IPD is becoming increasingly popular, various organizations are expressing interest in its benefits to the AEC industry. However, no research studies have shown statistically significant performance differences between IPD and more established delivery systems.

This study fills that missing gap by evaluating the performance of IPD projects compared to projects delivered using the more traditional design-bid-build, design-build, and construction management at-risk systems, and showing statistically significant improvements for IPD. Project delivery performance literature was analyzed to identify key variables, which were classified in three domains that reflect the social, organizational, and functional aspects of project delivery. A data-collection survey was developed and used in detailed interviews to gather quantitative delivery and performance data from the industry. Univariate and multivariate data analyses were performed, leading to the development of benchmarks and models of IPD project performance. An overall project performance model, the *Project Quarterback Rating*, also was developed and used to evaluate IPD performance.

Results indicate that IPD achieves statistically significant improvements in 14 metrics across six performance areas: building quality, project schedule, project changes, communication among stakeholders, recycling and financial performance. One major interpretation of the results is that IPD provides higher quality facilities at no significant cost premium. The findings provide a comprehensive understanding of IPD performance by specifically identifying metrics that are enhanced by this emerging delivery system.

Furthermore, the results highlight individual delivery characteristics that the industry can use to improve its performance. These characteristics include contractual incentives, use of Building Information Modeling processes, avoiding lump sum compensation, and perhaps most interestingly, social factors like the unconventional participation of stakeholders along the project timeline. Although some of these delivery characteristics are typically associated with IPD, the AEC industry can apply most of these characteristics with other delivery systems to improve performance in both private and public sector projects.

Table of Contents

List of Tables

List of Figures

Chapter 1. Introduction

Integrated Project Delivery (IPD) is the subject of great interest in the construction industry today. The Construction Industry Institute, the American Institute of Architects, the Construction Users Roundtable, and other organizations, have weighed in on the topic, as evidenced by the several reports and publications dedicated to IPD or closely related subjects (e.g. CII 2011; AIA 2011; CURT 2007). Additionally, construction magazines, such as *Engineering News Record* (ENR) and *Tradeline*, have featured projects that have used IPD (ENR 2011; Tradeline 2007; Tradeline 2009). *Figure 1* shows the cover of the September 12, 2011 issue of *ENR*, along with a copy of the multiparty contract signed by eleven stakeholders of a California hospital project. Articles in several journals, including the *Journal of Construction Engineering and Management*, the *Construction Lawyer*, and the *Lean Construction Journal* (LCJ) have commented on experiences and potential benefits of IPD, such as the reduction of project costs and increased cooperation in the construction process (Matthews and Howell 2005). In 2011, *LCJ* dedicated an entire issue to IPD, discussing integrated delivery and its implementation, suggesting positive outcomes from integration, and illustrating barriers to this transition. This study examines the claims of superior performance by statistically studying the performance of IPD.

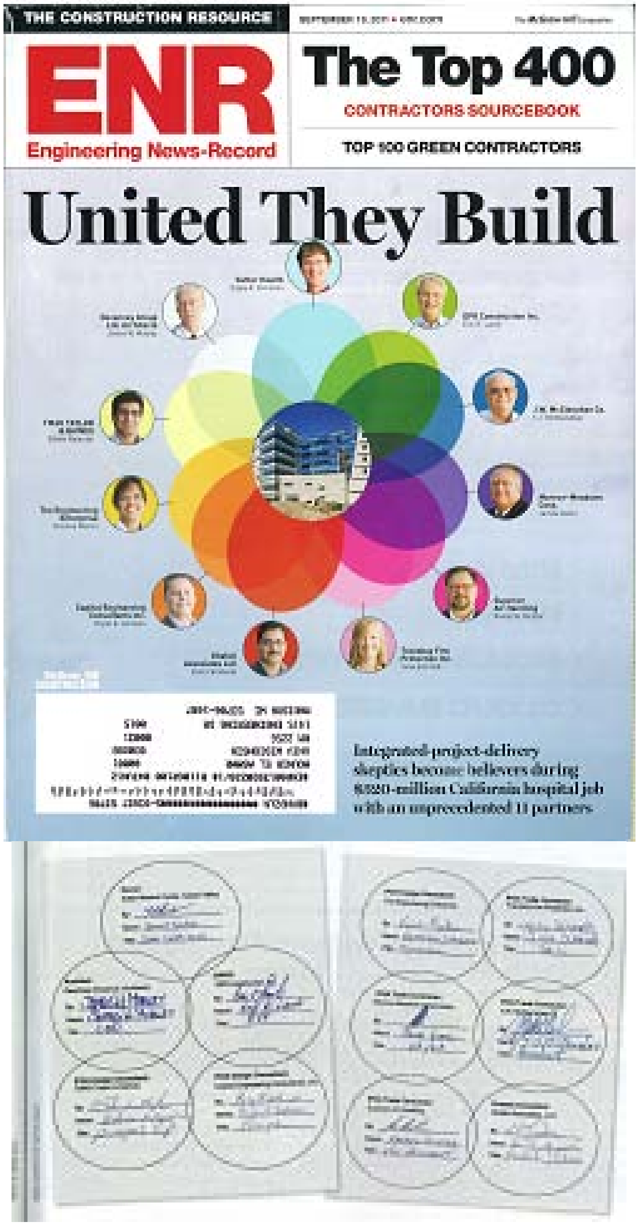


FIGURE 1: ENR Featuring Eleven-Party Contract

To ensure the topic is introduced appropriately, this dissertation will start with a brief section that defines exactly what IPD is and how it compares to other project delivery systems. From there, the chapter will cover the background and motivation for this study, and then the problem statement and the methodology used will be discussed.

1.1 Definitions of Terms

There are several existing definitions of “project delivery systems”. Cho et al. (2010) have tried to summarize the different definitions under three components: commercial terms, organizational structure, and management system. This thesis will discuss these components in greater detail in *Section 3.1*. However, two elements are consistently found in the majority of delivery systems definitions: (1) relationships of project stakeholders, and (2) their timing of engagement in the project (Hanna 2011). Therefore, this thesis defines a project delivery system as *a system that determines the relationships between the different project stakeholders and their timing of engagement to provide a facility*.

There are several types of project delivery systems being used today. *Figure 2* displays certain differences between the traditional design-bid-build (DBB) system, the more integrated design-build (DB) system, and the emerging IPD system.

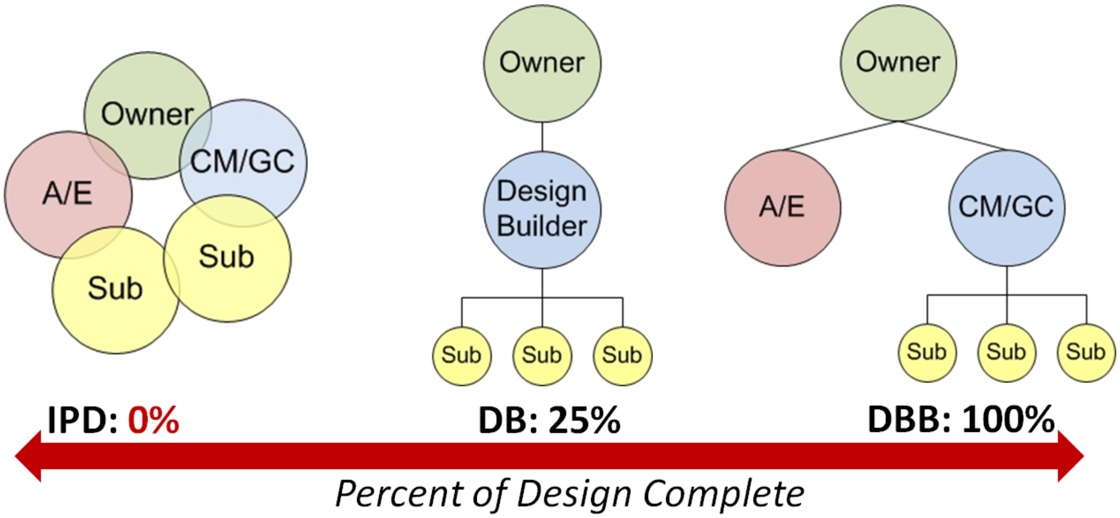


FIGURE 2: Differences between DB, DBB, and IPD

The two key focus areas are in accordance with the definition stated above with respect to the relationships between project stakeholders and their timing of engagement. For example, under DBB, the owner contracts with the designers, and then when their design is 100% complete, the owner would contract separately with a general contractor (GC) to build the facility. In DB, the contractor generally would be involved when the design is around 20% complete (the portion of design complete varies based on the project at hand), and the designer and GC would join forces, therefore providing a single point of responsibility for the owner.

In contrast, IPD is different in these two key aspects: (1) all key project stakeholders sign *one* multiparty contract (2) before the design even starts, i.e. when *0%* of the design is complete. Key stakeholders can include many project parties, such as the owner, GC, architect, consultants, subcontractors, and suppliers. Therefore, this thesis defines IPD as *a delivery system distinguished by a multiparty agreement and the very early involvement of the key participants*. As discussed later in *Chapter 2*, the term *IPD-ish* will be used to describe projects that were originally referred to as IPD, although they do not include all the necessary characteristics of a true IPD project, namely a multiparty contract.

1.2 Background and Motivation

1.2.1 AEC Industry Problems and Changing Factors

Even after the economic downturn, the U.S. annual value of *construction put in place* was $830 billion, as estimated in May 2012 (U.S. Census Bureau 2012). Estimates of the U.S. nominal Gross Domestic Product (GDP) are around $15 trillion (U.S. Department of

Commerce 2012), which implies that the U.S. construction industry accounts for approximately 6% of the country's GDP. Other sources estimate this number to be larger. Even though the construction industry is a significant part of the U.S. economic mix, major inefficiencies in the industry are evident in several recent studies.

Although technology is advancing, a U.S. Bureau of Labor Statistics study demonstrates a consistent 10% decline in productivity in the construction industry between 1964 and 2004 (see *Figure 3*), whereas all other non-farm industries have benefited from a doubling in their productivity over the same period (Teicholz et al. 2001 and 2004).

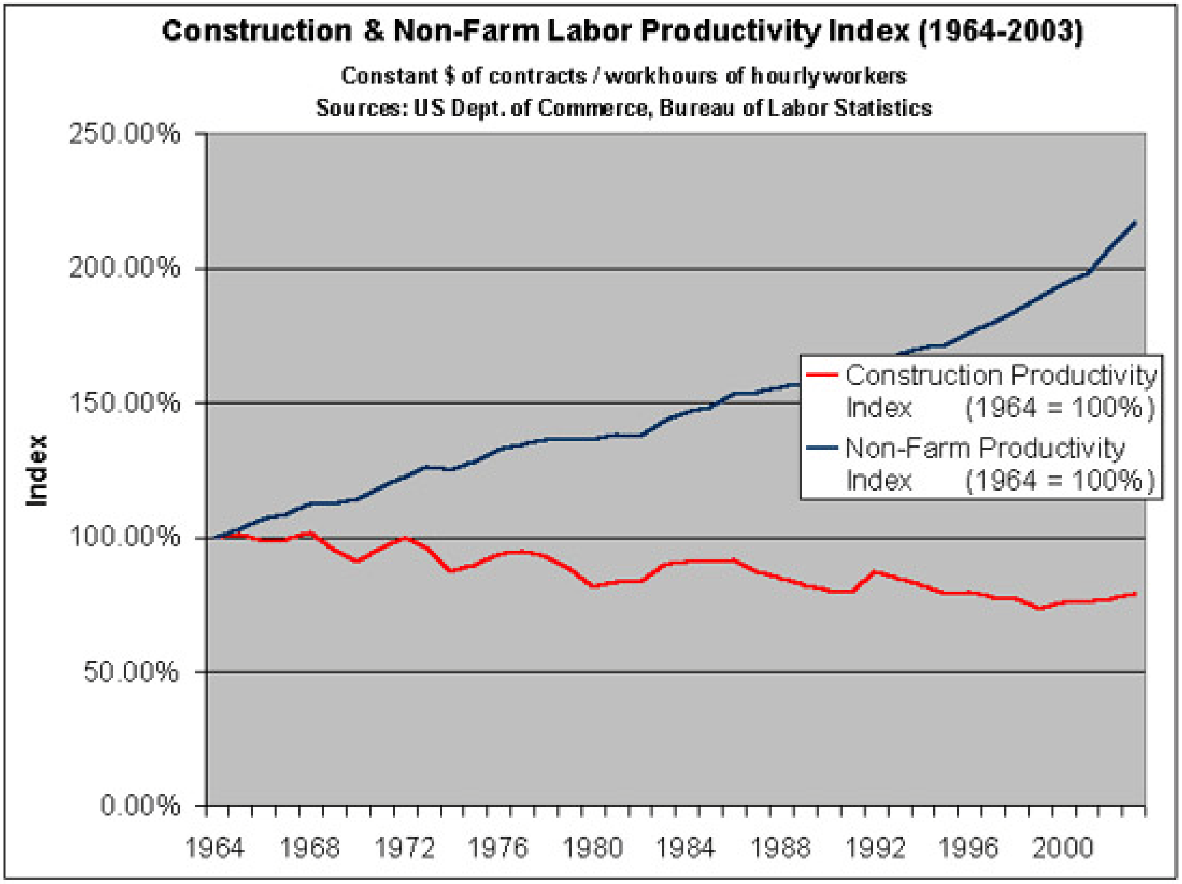


FIGURE 3: Construction Industry Productivity (Teicholz 2004)

This finding, however, should be cited in the proper context so it does not mislead. Labor productivity measured only takes account of the construction phase of projects; therefore, the trend does not take into account the productivity issues that occur in the design phase through the process of rework or organic design decisions that can radically affect construction productivity. Hanna (2010) draws the same conclusions related to productivity declines for the Canadian construction industry.

The Construction Industry Institute (CII) explains that one potential cause of the productivity decline is the existence of waste during project delivery. “Waste” is defined as *any activities that do not add value to the building, as perceived by the customer*, such as waiting time and unnecessary movement or transportation of materials. Liker (2004) defines eight distinct types of waste, including defects, over processing, and unused employee creativity. Horman and Kenley (2005) report an average 49.6% of construction operative time is spent on wasteful activities. Hanna (2010) shows value-added-activities only make up 41% of a construction workday. The remaining 59% of the time can be attributed to wasteful activities, as shown in *Figure 4*.

Although studies show differing waste values, they all lead to the same uncontested conclusion: construction waste is a major problem. To display the magnitude of the problem, if Hanna’s results are multiplied by typical labor cost percentages and the GDP numbers discussed previously, one could argue that around 1.5% of the total U.S. GDP can be labeled as pure waste; waste entirely resulting from construction.

Waste is, indeed, engrained throughout industry practices. In fact, numerous studies list a variety of problems found in the construction industry today. One of the industry-level factors discussed by Hanna (2010) is the project delivery system. For example, waste can be seen in the traditional design-bid-build (DBB) delivery system, which is sequential and linear, lacking collaboration between the several project stakeholders. Collaboration deficiencies ultimately result in inefficient processes, which can lead to even more wasteful practices, such as design and construction rework. Specifically, rework in design and construction has been found to account for 52% of the total cost growth for projects (Love 2002).

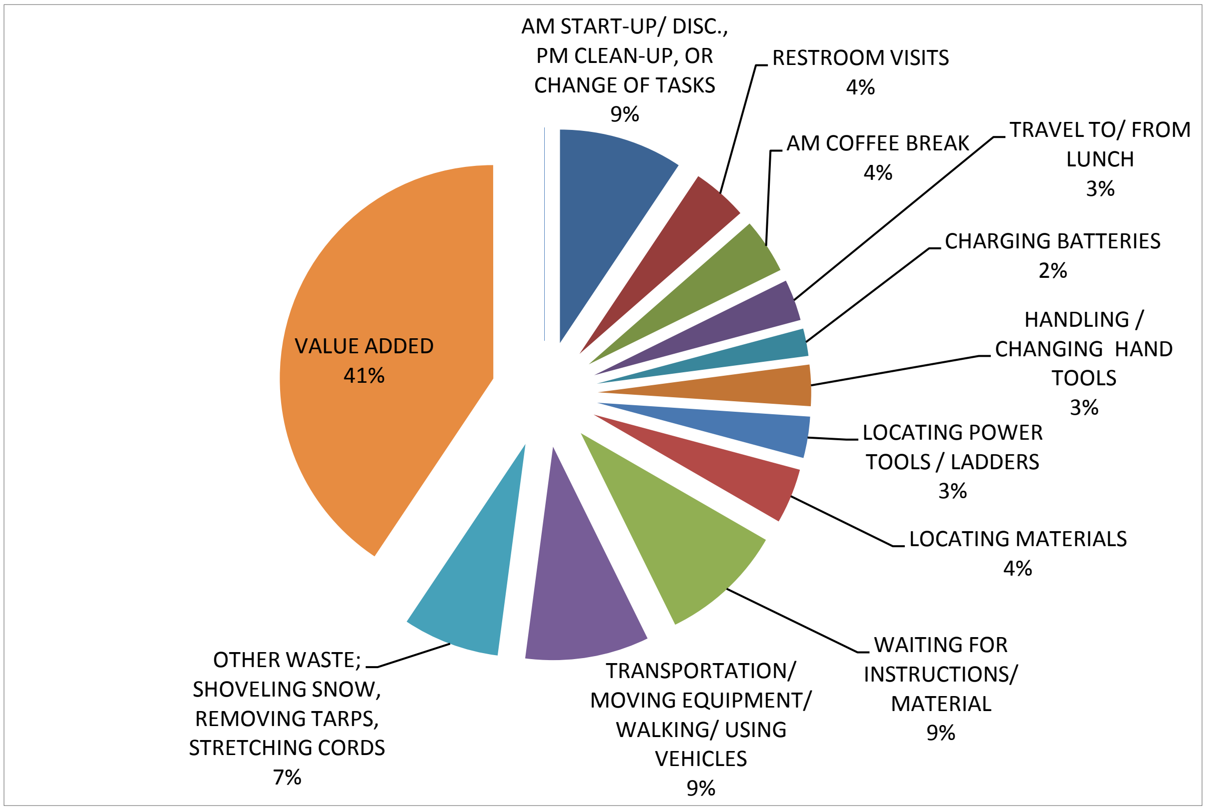


FIGURE 4: Value vs. Waste in Construction (Hanna 2010)

In addition to time wasted and rework, schedule and budget compliance also are systemic problems the industry is facing. For example, the National Cooperative Highway Research Program reports that only 35% of transportation projects that cost above $5 million are delivered on time, and less than 20% of such projects are delivered on or under budget (NCHRP 2007). Moreover, the Construction Management Association of America (CMAA) found 40% to 50% of all construction projects are not delivered on time (Thomsen et al. 2010).

Smith et al. (2011) list additional industry problems, including: safety, productivity, errors and litigation, design quality, an adversarial culture caused by traditional contracts, owners losing money, and constructors bearing too much risk. The issue of improper risk allocation has also been shown by previous studies (Hanna and Swanson 2007). Hanna and Swanson also estimated that misallocation of risk increases project cost by 13% as a result of increasing contingencies. One might add to the above list issues such as lack of coordination between trades, lowered quality of design and construction caused by an increased competition, cost of litigation, latent defects, lack of sufficient industry data to conduct effective studies, and systemic lack of planning and tracking throughout the industry.

In addition to the myriad of problems, there are also some significant changing factors in the construction industry. The makeup of the industry is changing with the number of contractors skyrocketing and the percentage of union memberships decreasing. Technology is also changing, affecting behaviors and new capabilities. For example, Building Information Modeling (BIM) is resulting in early collaboration between designers and contractors, which defies the traditional sequential over-the-wall behaviors. These industry problems and changing factors ultimately led to the development of IPD.

1.2.2 IPD as a Potential Solution

The emerging IPD system is believed by many in the industry to be revolutionizing the way projects are delivered by fostering early involvement and collaboration of project stakeholders through the use of different concepts, such as: shared project leadership, shared risk and reward between all project participants, multiparty contracts, and liability waivers.

The need for more collaboration in general, and for IPD specifically, is best expressed by the 2004 and 2007 reports of the Construction Users Roundtable (CURT). The earlier report encouraged owners to drive the construction industry change “by leading the creation of collaborative, cross-functional teams comprised of design, construction, and facility management professionals.” The second report specifically spells out CURT’s path toward embarking on IPD projects.

Several sources estimate great benefits of utilizing IPD. The United Kingdom’s Office of Government Commerce (UKOGC) estimates savings anywhere from 2% to 10% in the cost of construction for single projects, and up to 30% for strategic partnering where integrated teams work together for more than a single project (UKOGC 2007). Other studies by the American Institute of Architects (e.g. AIA 2010) showcase a handful of successful IPD case studies. Mossman et al. (2010) discuss potential benefits of integrated delivery through case studies. For example, clients obtain more value and reduced energy costs of use; designers see reduced design documentation time and it is easier for them to keep the design within the target cost; and constructors experience less rework and work on more buildable facilities.

Cho and Ballard (2011) statistically compared the performance of IPD projects to non-IPD projects and found no significant performance differences. However, their definition of *performance metrics* by time and cost growths could potentially be considered narrow and not comprehensive enough; IPD potential benefits cannot be restricted to cost and time reductions in construction. For example, customers might reinvest the saved cost back into the project, getting more value out of their facility. This likely situation results in no visible differences in the authors' analysis of cost and time reductions.

Therefore, due to the lack of comprehensive studies targeting IPD performance, there still exists a need to evaluate IPD and understand its true performance based on several important metrics used in construction. The performance of IPD projects needs to be compared to the performance of projects delivered with other more traditional delivery systems, which will serve as a baseline for this study.

1.3 Problem Statement

As stated earlier, there are major problems in the construction industry today, and IPD seems to be a potential solution. A survey of the literature to date shows no comprehensive studies that have statistically compared and quantified the benefits of IPD projects relative to non-IPD projects based on performance metrics. Despite the references to several sources discussing IPD benefits, one can see that the UKOGC report gives mere estimates, and the Mossman and AIA studies only discuss the data they collected from individual case study perspectives.

Aside from a few case studies and anecdotal examples, no significant literature exists to support the claim of superior IPD performance. In fact, the only research study that statistically investigates this claim found no performance differences between IPD and non-IPD projects (Cho and Ballard 2011). The hypothesis that the implementation of IPD would improve project performance is not supported by their statistical analysis. Since no solid statistical inference can be made based on the findings, data collection and analysis are still necessary to investigate the relationship between IPD and project performance.

Unlike other more established delivery systems, it is very challenging to quantify performance metrics for IPD. Most IPD publications that report an enhanced performance are only based on a handful of case studies, because collecting IPD data is not an easy task. The system has only been applied recently, which means very few complete projects can be targeted to collect data. This study focuses on contractors because the have access to most of the quantitative project data; however, obtaining detailed information from contractors about their use of IPD is a major hurdle because it might involve protected trade secrets. A comprehensive and correct measurement of performance also requires heavy involvement and commitment from industry collaborators to provide project information they may or may not have readily available.

As demanding as this task might appear, successfully undertaking such a study will lead to worthy outcomes. The completion of this research will not only show how IPD projects could save waste resulting from construction, but also how IPD could help all project stakeholders and especially owners fulfill their own missions by potentially resulting in more efficient projects and higher-quality facilities.

1.4 Methodology

The methodology for this study encompasses three distinct stages, as illustrated by the flowchart in *Figure 5*. The flowchart serves as a roadmap for the major research stages and steps discussed in this chapter. The following is a discussion of the flowchart, presenting more detail about the research stages and the steps within each stage. A section is devoted to each individual step to explain it thouroughly.

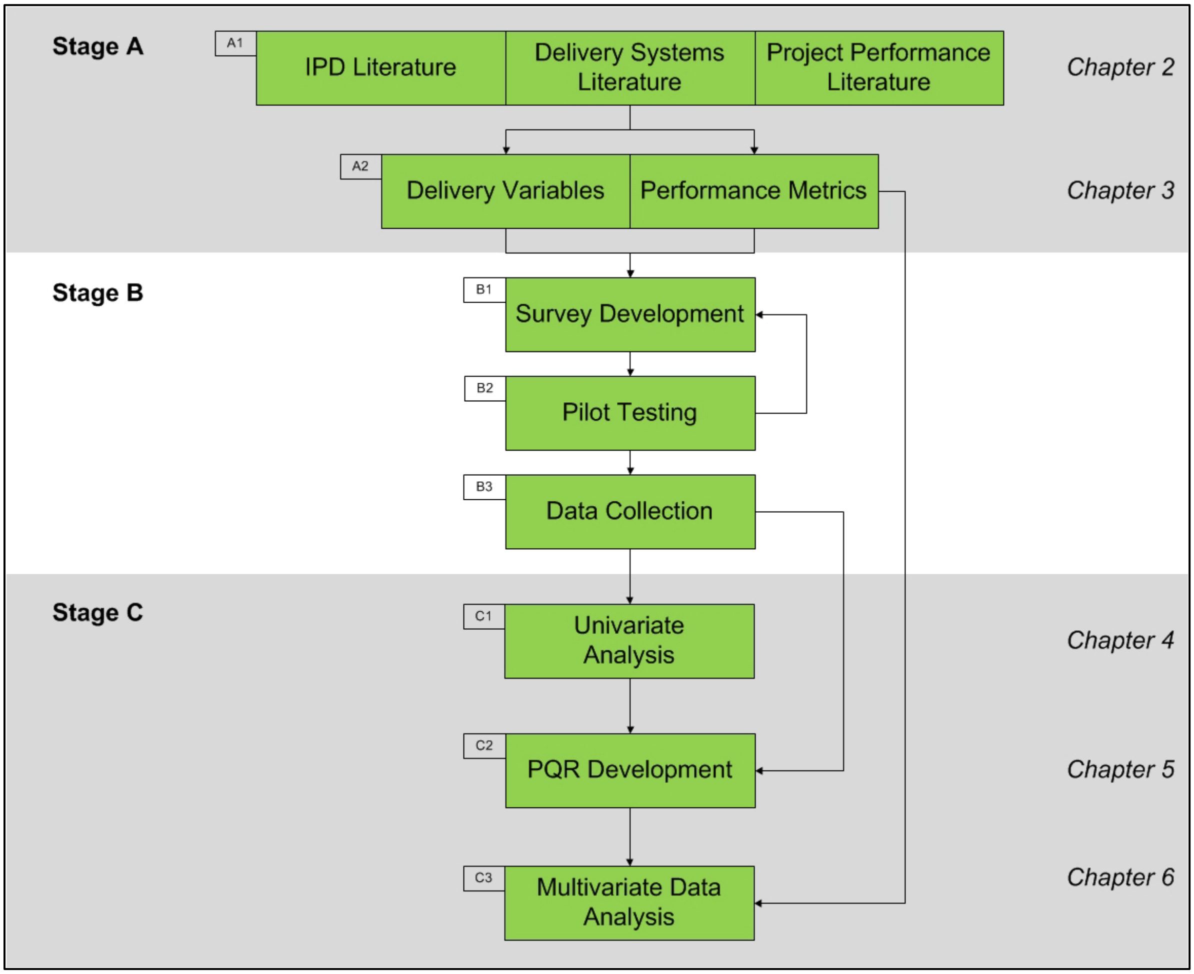


FIGURE 5: Research Methodology

Stage A is an assessment of the literature and industry practices that will lay the ground for the rest of the study. *Stage A* consists of two steps, the first of which is meant to appreciate the current state of knowledge regarding IPD, project delivery comparisons, and project performance metrics. Based on the findings from the literature, the second step is meant to identify key variables that need to be analyzed in order to answer the research questions.

1.4.1 Step A1 – Review of IPD and Delivery Performance Literature

As will be shown in *Chapter 2*, a methodical examination of the construction literature was completed to identify previous studies that focused on the performance of major project delivery systems, as well as studies that looked specifically at IPD. Case studies also were considered for IPD due to the newness of the topic, since IPD performance benchmarks are not as established as for other delivery systems. Unlike the project data collected for this study, the literature review was not restricted to a specific timeframe. The absence of a time restriction allows the inclusion of all relevant publications to allow the background of this study to be as comprehensive as possible.

Several databases were researched for journal articles, conference proceedings, and published books, as well as studies completed for national bodies. An extensive list of publications was reviewed, including American Society of Civil Engineering (ASCE) publications such as the *Journal of Construction Engineering and Management* and the *Journal of Management in Engineering*, and studies conducted by the Construction Industry Institute (CII), the Lean Construction Institute (LCI), and the American Institute of Architects (AIA).

AIA publications provided particular benefits during the very early stages of this research. One AIA report included valuable project data that was used in the early stages of this study to conduct a preliminary analysis that served as a "proof of concept". The preliminary analysis, parts of which are available in *Appendix D*, confirmed the significant value and feasibility of the research and was shared with the industry participants involved in the data collection phase to encourage their participation.

1.4.2 Step A2 – Key Variables

The key variables used in this study can be separated between inputs and outputs. The inputs are independent and control variables and will be discussed first, while the outputs are performance metrics and will be discussed second.

Independent and Control Variables

As will be explained in further detail in the *Section 2.3*, key variables that differentiate projects from one another were identified from the literature. Independent and control variables are characteristics that exist throughout the project delivery phase, *not* an outcome measured at the end of a project. The independent variables, such as the project delivery system, are characteristics of interest to this study and will be tested to determine their effect on the performance metrics. Alternatively, the control variables are factors that have an effect on performance but are mostly unmanageable, such as the project size, type, and complexity. These variables are not of particular interest to the study but should still be accounted for and controlled to understand the variations stemming from the independent variables.

Independent variables that can be most influential on the tested performance metrics include those defined by Sanvido and Konchar (1998), and Korkmaz et al. (2010), and will be discussed in more detail in *Chapter 2*. Other characteristics this thesis considered are the existence of incentive clauses through sharing risk and reward, and the use of technology and information management systems such as Building Information Modeling (BIM). Some additional factors are the use of claim waivers and the involvement of stakeholders in the project management structure. All of the variables mentioned, and others, will be denoted as delivery characteristics. *Chapter 3* will include a full listing of these variables and a discussion of how these delivery characteristics were grouped in three areas: (1) Contractual Terms, (2) Social Tone, and (3) Tools and Techniques.

Project Performance Metrics

Quantitative and qualitative project performance metrics are dependent variables measured after project completion. The initial list of performance metrics used for this study was based on the factors measured in previous studies, and was later complemented with additional factors recommended by this study's industry panel to compare performance of projects and gauge project success.

One example of a pioneering study that was used as background for this study is Sanvido and Konchar's (1998) work on project delivery systems. Sanvido studied several metrics, including: unit cost, construction speed, delivery speed, cost and schedule growth, turnover quality, and systems quality. For the purpose of this study, the list was updated to include metrics identified in more recent studies as will be shown in *Chapters 2* and *3*. The compilation of these metrics will provide a more comprehensive picture of project performance.

Industry collaborators for this study identified certain "waste factors" as candidates to be included in the list of performance metrics to investigate, such as: RFI's, resubmittals, and deficiency issues. These are items that should be minimized to avoid wasteful practices in construction projects. The research team also suggested supplementary metrics, such as the number and cost of claims.

The achieved LEED rating of buildings will not be considered as a performance metric based on the several studies that show no relation between the LEED rating of a building and its energy performance (Menassa et al. 2012). The list of performance metrics was finalized through several interactions with industry professionals and executives. The full list is shown in *Chapter 3*.

Identifying key variables highlights the important factors that need to be investigated and allows the research to move on to the next stage of this study. In fact, identifying the key variables most important to answer the research questions provides guidance about the type of data that needs to be collected. The completion of the first stage serves as a solid basis for the survey development. The survey is the backbone of the second stage of this research – Data Collection. Three steps are needed for the completion of *Stage B*: survey development; pilot testing; and data collection.

1.4.3 Step B1 – Survey Development

As stated above, the review of literature and consultation with industry partners allowed for the identification of key variables. A data collection questionnaire, or survey, was developed based on these key variables. The survey, acting as a foundation for the data collection efforts of this research, is designed to gather data on quantitative and qualitative delivery characteristics and performance metrics.

The survey was shared with industry participants, specifically the general contractor or construction manager of each targeted project, to act as a roadmap for interviews, and to allow for the gathering of data in a consistent format. Data was collected from recently completed projects delivered with DBB, DB, CM, and IPD.

The survey consists of thirteen pages, is aimed at project managers, and rigorously helps collect data related to the delivery phase of projects. The survey, which can be found in *Appendix C*, comprises five major sections. The first section requests data regarding general project characteristics, such as size, location, and the project delivery system used, as well as general contract aspects, such as the type of compensation, incentives, and the inclusion of specific clauses in the contract. The second section of the survey looks at project performance directly, as well as indirectly through gathering data needed to calculate performance metrics that included safety performance, cost performance, schedule performance, project changes, and labor performance. The third section is an overview of project systems looking at both the complexity and the quality of each major building system. The fourth section gathers data about project team characteristics with an emphasis placed on collaboration. The respondents are asked about various project aspects such as their prior experiences, the procurement of contractors for the project, the project management structure, the stakeholders' involvement in different phases of the project, and the use of BIM and lean construction tools and techniques. In the fifth and last section of the survey, respondents are asked to define factors that they consider when assessing the success of a project as applied to the specific project at hand. An introductory letter and a set of definitions accompany the survey to provide some background information for the interested participants in preparation for detailed interviews during which project data is collected.

Some sections of the first version of the survey were partly based on a previous data collection instrument developed for a CII study on change orders, as well as another CII study conducted by Sanvido (1998) to compare delivery systems, and several other studies found in the literature. Before the survey was used to collect data, it went though three specific review stages. The first stage consists of individual reviews, the second stage consists of collective reviews, and the third stage consists of reviews by a professional survey center:

Individual reviews: After the survey was developed, it was reviewed individually by industry experts (contractors, designers and owners) as well as UW-Madison construction engineering faculty members and graduate students. After refining the questions, another review round was conducted by senior industry executives.

Collective reviews: After the questionnaire was updated based on the valuable feedback from the first round of review, it then was presented to three panels of industry experts, including contractors, designers and owners. Information was shared through web conference calls; software including Windows Live Meeting and GoToMeeting were used to share the survey document between meeting sites. The questionnaire was updated in real time, guided by input from the panel participants including executives, project managers, improvement professionals, and estimating professionals. The goal of the review was to provide answers to the following:

1. Is the terminology clear? Are any definitions or clarifications needed?
2. Is the data available? Is it accessible by the project manager?
3. Are the questions adequate for the research purpose? Are there any questions that need to be added or deleted?

The participants provided constructive comments that contributed to finalizing the questionnaire. Once the questionnaire was updated, the next step was a review with survey professionals.

UW Survey Center reviews: The University of Wisconsin–Madison is home to a nationally renowned survey center. After the individual and collective reviews, the survey was reviewed by the professionals at the survey center, who suggested several changes, ranging from rephrasing questions to formatting the survey in order to make it more appealing to the respondents and ultimately increase the response rate.

1.4.4 Step B2 – Pilot Testing

A pilot study is a preliminary study conducted on a small scale to test and improve the survey before performing the full-scale data collection. Following the thorough survey development stage, the final data collection questionnaire was ready to be pilot tested on projects from contractors who expressed great interest in the study. The pilot test covered a limited number of projects in order to refine the questions and maximize their effectiveness. The survey was used as a thorough outline to conduct three detailed project interviews with the respective project managers. The pilot testing offered insights from the respondents' perspective, and helped finalize the survey before performing the full-scale data collection effort. The resulting comprehensive thirteen-page survey, shown in Appendix C, allowed for an intense data collection effort targeting delivery characteristics and performance metrics for individual construction projects.

1.4.5 Step B3 – Data Collection

Data collection is the step where the developed survey is used as a roadmap for interviews with the contractors' project managers to gather quantitative and qualitative project data. Typically, randomized sampling is the ideal data sampling strategy to statistically support a hypothesis. However, since IPD is a new delivery system that is not widely used in the industry, a random sample of projects would be unlikely to include any IPD projects. It is practically impossible to randomly select companies for the data collection stage because of the limited number of companies using IPD to deliver building projects. Therefore, this study uses purposive sampling (Babbie 2010) and targets a specific set of companies that are known to have recently used IPD, and from that set, a list of interested companies was identified for the data collection phase. These companies were asked to provide data from IPD projects as well as comparable non-IPD projects.

Company executives were first contacted by email or phone in order to identify the correct individual with access to the required information needed to answer the survey questions. After initiating contact with the project managers or other representatives identified by the company executives, the survey was sent to these individuals by email to initiate the data collection efforts.

Face to face interviews were conducted with the project managers to collect data following the questionnaire format. If in-person interviews were not feasible, teleconferences, phone interviews, and mailing surveys were conducted. Interviews, as opposed to other data collection techniques, have a greater potential to result in higher quality data. Furthermore, interviews allow for opportunities to clarify the survey by answering questions on the spot. This method will leave no room for uncertainty in the participants' interpretation of potentially unclear questions or when specific data is unavailable. Also, interviews are likely to increase the response rate, especially if travel and expenditure of personal resources are involved.

In some instances, interviews could not be conducted. In those circumstances, clarifications regarding the survey questions were provided via email and telephone communication. Further follow-up was conducted to ask additional questions when needed.

Upon completion of *Stage B*, the developed questionnaire had been used to gather responses, and the resulting project data had been verified and ready for analysis. After the data collection phase is completed, the focus shifts to the performance metrics for which data was most readily available. The third and last stage of this research builds on the previous two and consists of analyzing the data collected, and developing benchmarks and models for IPD project performance. *Stage C* encompasses three distinct steps: Univariate Analysis; *Project Quarterback Rating* (PQR) Development; and Multivariate Analysis and Models Development.

The statistical analyses in *Step C1* consist of testing whether IPD leads to superior performance. As discussed in *Section 1.5*, these findings will satisfy the first contribution of this research. *Step C2* discusses the *Project Quarterback Rating* (PQR). Drawing inspiration from the National Football League's passer rating for quarterbacks, the PQR aggregates different performance metrics (i.e. cost, schedule, quality) into *one* number to rate project success. The development of the PQR satisfies the second contribution of this research and will be used for the multivariate analysis in *Step C3*. *Step C3* accomplishes the third research contribution by developing overall performance forecasting tools for IPD projects, based on several key project delivery characteristics.

1.4.6 Step C1 – Univariate Analysis

Univariate analyses are utilized to understand the effect of a single independent variable (e.g. project delivery system) on project performance. However, before any analysis is conducted, some collected project data needs to be adjusted in order to accurately compare projects built in different cities in different years. Cost data was adjusted for time and location using the latest Engineering News Record (ENR) and RS Means historical cost indexes.

The first set of analyses tested the individual effect of delivery systems on performance. The collected project data is used to compare the various delivery systems and test whether IPD is more successful than other types of PDS based on each individual performance metric. *Table 1* presents a sample of univariate analyses that were performed for this study. One example of a hypothesis used for this part of the study is: "**IPD projects result in a higher delivery speed than non-IPD projects.**" Similar hypotheses were developed for the remaining performance metrics. For each metric, normality tests are conducted, and then two types of analysis are used for each performance metric, (1) t-tests and (2) Mann-Whitney-Wilcoxon tests. *Chapter 4* will discuss these analyses in greater detail.

The second set of univariate analyses looked at other identified delivery characteristics individually. Individual characteristics, such as the percent of design complete when the contractors get involved in the project, or the extent of Building Information Modeling (BIM) use, were tested against project performance. *Section 6.4* discusses the statistical analyses performed to quantitatively inspect any differences in performance based on these delivery characteristics.

TABLE 1: Sample of Univariate Analyses

Independent Variable	**Dependent Variable**	**Analysis**
IPD vs non-IPD (*categorical*)	Delivery Speed	t-test and MWW test
IPD vs non-IPD (*categorical*)	Lost Time Injuries	t-test and MWW test
IPD vs non-IPD (*categorical*)	Systems Quality	t-test and MWW test

After univariate analyses were performed to study the impact of the project delivery system on each project performance metric, these individual metrics are combined into a more comprehensive rating to allow for an all-inclusive analysis of project performance.

1.4.7 Step C2 – Project Quarterback Rating (PQR)

The *Project Quarterback Rating* (PQR) aggregates key performance metrics, such as cost, schedule, quality, and safety into *one* number to compare the level of success for construction projects on the same basis. The development of the PQR complements the analyses of individual performance metrics by providing an overall project success metric allowing project-to-project comparison using a single comprehensive value for each. Additionally, PQR is used as one supplementary metric to test performance for IPD projects as compared to non-IPD projects. PQR allows for a comprehensive multivariate comparison of performance between IPD and other delivery systems, as discussed in *Chapter 5*.

1.4.8 Step C3 – Multivariate Analysis and Models Development

By concurrently identifying variables that have the most impact on project performance, a multivariate analysis can present a more complete explanation of performance variations for IPD. The multivariate analysis can verify univariate results by ensuring that these results are robust when additional variables also are considered to explain performance variation. A key product of the analysis is to understand which of the IPD delivery characteristics were causing better performance, highlighting important contractual factors, social factors, and tools used on projects to help the architecture/engineering/construction (AEC) industry achieve successful project completion. *Chapter 6* presents a more thorough discussion of the multivariate analyses.

1.5 Research Contributions

This research offers three main contributions to the construction engineering and management literature and to the AEC industry:

1. A demonstration of IPD performance to guide project stakeholders in making informed decisions when choosing a delivery system and related characteristics;

2. The creation of a unified comprehensive project performance metric for a practical comparison of overall performance for any project under any given delivery system; and

3. The development of an explanatory model for IPD projects to understand the effect of specific delivery characteristics on project performance.

1.6 Dissertation Organization

Chapter 1 acted as an introduction, providing the background and motivation for this study, as well as definitions of some key terms and an overview of the research methodology covering three main research stages that encompass 8 distinct steps. Next, *Chapter 2* consists of a review of the literature, structured in three main parts: IPD literature, delivery systems comparisons, and key variables and metrics to analyze. The chapter also covers the research opportunities and the objectives needed to fill gaps in the literature, as well as the scope of this study. *Chapter 3* presents the data collection results and discusses the final variables used in this study as well as their organization and coding. *Chapter 4* discusses the univariate analysis and results through a comparison of IPD projects to non-IPD projects based on individual performance metrics. *Chapter 5* discusses the development of the *Project Quarterback Rating* (PQR), a comprehensive performance metric to gauge overall project success, and its use to compare IPD and non-IPD project performance. *Chapter 6* discusses the multivariate data analysis techniques that were used and presents the results in terms of modeling overall project performance based on project delivery characteristics. Finally *Chapter 7* presents the conclusions and recommendations of this study. The bibliography used for this study can be found in *Appendix A*, and then a list of definitions and abbreviations is presented in *Appendix B*. The data collection survey developed for this study can be found in *Appendix C*, and results of a preliminary analysis on an AIA dataset are shown in *Appendix D*. Then *Appendix E* shows selected outputs of the Minitab software. Finally, *Appendix F* presents the correlations of the performance metrics used in this study and *Appendix G* presents an example to help users apply the PQR formula.

Chapter 2. Literature Review and Research Opportunities

This chapter consists of a comprehensive survey of the literature to understand the current state of knowledge in the construction field. The focus of the literature review will be on three aspects: (1) Integrated Project Delivery (IPD), (2) major project delivery comparisons, and (3) key analysis variables proven to be important when studying delivery performance.

2.1 Integrated Project Delivery (IPD)

There is no single definition for IPD. In fact, numerous definitions can be found throughout published studies and reports. As shown later in this section, some of the definitions continue to evolve even within the same organization. This study develops its own IPD definition in *Section 3.2*, based on the literature reviewed here and complemented by the collected project data, while also remaining consistent with the definition of a project delivery system.

Mathews and Howell (2005) define IPD as "*a relational contracting approach that aligns project objectives with the interests of key participants.*" The authors illustrate IPD through a case study of the Orlando Utilities Commission North Plant. The performance results of this project show a 10% cost saving as compared to the initial Guaranteed Maximum Price (GMP).

Lichtig (2005) developed a contracting model to support IPD for Sutter Health in California. The paper discusses commercial strategies, such as team selection through quality-based evaluations and interviews, early involvement of major subcontractors, and collaborative design. An interesting discussion is included regarding the goal of risk management being to share and reduce overall project risk, rather than just shifting the risk between project participants. Along with shared risk also come shared savings through an incentive program, as well as a dispute resolution process that preserves relationships between the parties. The author also presents the Sutter approach which revolves around Sutter Health's *Five Big Ideas*:

- Collaborate; really collaborate
- Projects as networks of commitments
- Tightly couple learning with action
- Increase relatedness
- Optimize the project not the pieces

The American Institute of Architects (AIA) first defined IPD as "*a project delivery approach that integrates people, systems, business structures and practices into a process that collaboratively harnesses the talents and insights of all participants to optimize project results, increase value to the owner, reduce waste, and maximize efficiency through all phases of design, fabrication, and construction*." The authors go on to add that IPD is distinguished by extensive and early collaboration among project stakeholders (AIA 2007a). The AIA document lists the principles of IPD under nine sections, shown in the first column of *Table 2*. The second and third columns of the table will be discussed at a later point in this section. The 2007 AIA report expressed two key messages. First, although challenged by some other IPD definitions, AIA stated that IPD principles can be applied to a variety of contractual arrangements and are not restricted to multiparty contracts. Second, IPD teams can include members well beyond the basic triad of owner, architect, and contractor. The report also discusses the need to redefine project phases.

TABLE 2: The American Institute of Architects' Principles of IPD – An Evolving Definition

AIA 2007	AIA 2010	AIA 2011
		MARKERS
		Relational Contracts
		Protection from litigation
Early Goal Definition	Jointly Developed Project Target Criteria **(C)**	Aligned project goals (Jointly Developed Project, Target Criteria)
Collaborative Innovation and Decision Making	Collaborative Decision Making **(C)** + Willingness to Collaborate **(B)**	Informed and balanced decision-making (Collaborative Decision Making)
Open Communication	Open Communication **(B)**	Open Communication
		Risks Identified and Accepted Early
		STRATEGIES
	Key Participants Bound Together as Equals **(C)**	Key Participants Bound Together as Equals (Multiparty Agreement)
Intensified Planning		Budget & create team for design intensive work
Early Involvement of Key Participants	Early Involvement of Key Participants **(C)**	Early contribution of expertise (Early Involvement of Key Participants)
Mutual Respect and Trust	Mutual Respect and Trust **(B)**	Pre-existing relationships between parties
Organization and Leadership		Champion/ Facilitator (Leadership by All)
Mutual Benefit and Reward	Shared Financial Risk and Reward Based on Project Outcome **(C)**	Shared Financial Risk and Reward Based on Project Outcome
	Liability Waivers between Key Participants **(C)**	Liability Waivers between Key Participants
	Fiscal Transparency between Key Participants **(C)**	Fiscal Transparency between Key Participants
Appropriate Technology		BIM - virtual rehearsal of construction and ongoing constructability reviews
		Lean Construction processes
		Co-location

Another AIA document highlights the differences between integrated and traditional project delivery methods, as shown in *Figure 6* (AIA 2007b). One major difference is the early involvement of the key stakeholders, which results in an earlier and clearer project definition. A yet different AIA report on collaboration discusses the value proposition of IPD, including a higher efficiency, cost predictability, and responsiveness to changing markets (AIA 2009).

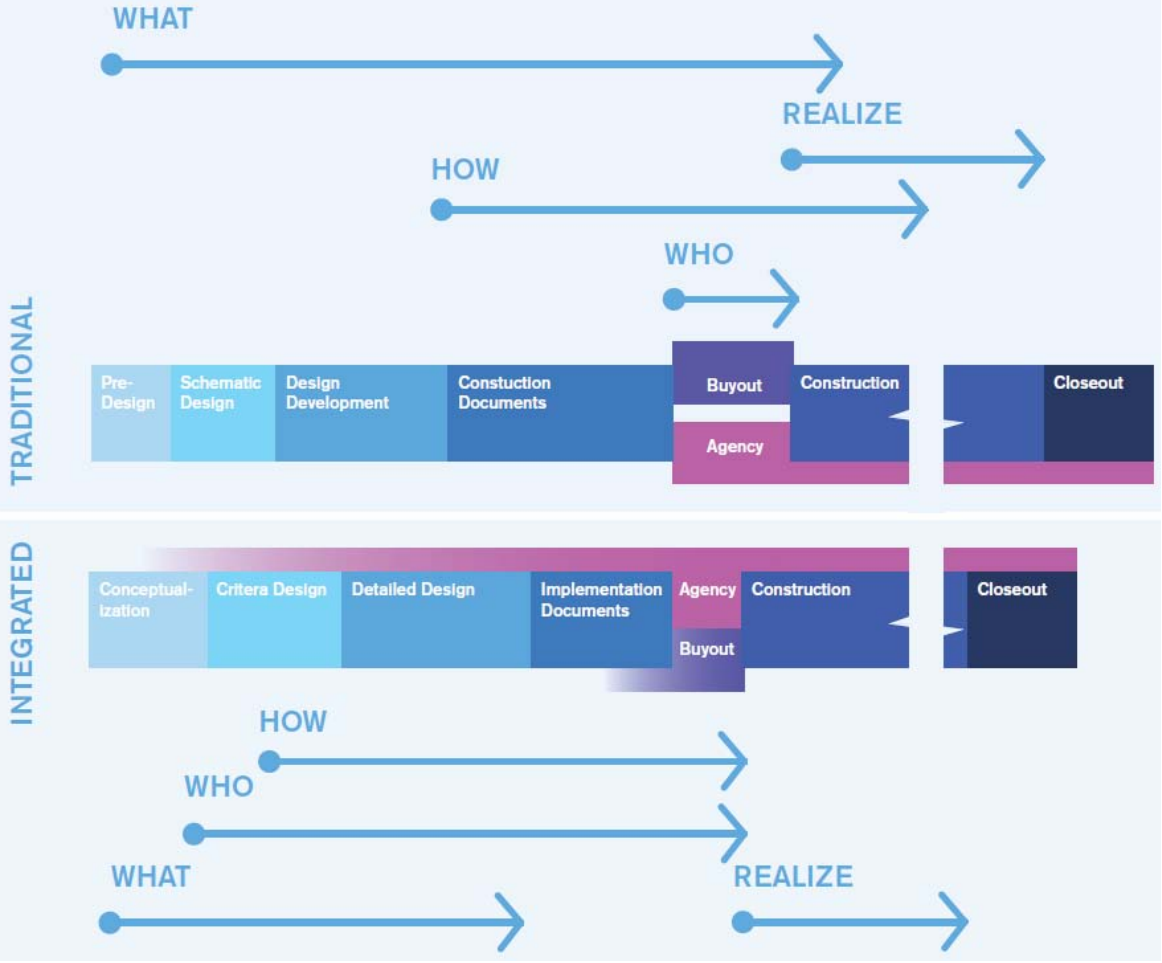


FIGURE 6: Comparing DBB to IPD (AIA 2007)

In 2010, AIA published six IPD case studies with completed project data along with the IPD characteristics used on each project. These case studies were used to complete a preliminary analysis that served as a proof of concept for this study. The analysis of these projects is discussed in more detail in *Appendix D*. The document included a refined definition of IPD: "*a project delivery method distinguished by a contractual agreement between a minimum of the owner, design professional, and builder where risk and reward are shared and stakeholder success is dependent on project success*." This definition is tied to a set of "Contractual Principles" (C) and "Behavioral Principles" (B), shown in the second column of *Table 2*.

More recently in 2011, AIA published another report on a different set of case studies. The report covers five projects that are still in their early delivery phases, focusing on activities that lay the foundation for IPD. The AIA report found very few "pure IPD" projects that would fit under their previous IPD definition. Therefore, for this most recent report, the IPD definition was modified to include "IPD Markers" and "IPD Strategies", a list of which is shown in the third column of *Table 2*. According to this new definition, a project is considered IPD if it follows IPD Markers, while IPD Strategies are optional and can vary between IPD projects.

The AIA literature on IPD can be summarized in two key points. First, as can be seen in *Table 2*, their IPD definition is clearly evolving. Most of their early IPD *Principles* are being carried forward, gaining a clearer definition with time. However, their IPD requirements appear to be loosening to accommodate what is available in the industry today; the most noticeable example is considering multiparty contracts to be a strategy and not a required IPD characteristic. The second point is that their reports discuss case studies on an individual basis, and no statistical comparative analyses were performed to draw strong conclusions regarding IPD performance.

A joint effort of the National Association of State Facilities Administrators (NASFA); Construction Owners Association of America (COAA); APPA: The Association of Higher Education Facilities Officers; Associated General Contractors of America (AGC); and American Institute of Architects (AIA), culminated in a report on IPD that discusses the different levels of collaboration in project delivery (NASFA 2010). As shown in *Figures 7 and 8*, the lower level (*Level 1*) uses IPD as a philosophy with a CMR or DB delivery approach. The middle hybrid level (*Level 2*) employs some IPD characteristics, without a multiparty contract. The higher level (*Level 3*) uses IPD as a delivery system with a multiparty contract. This idea of different collaboration levels served as an inspiration for the definition of IPD in this study, and will be explored further in this study.

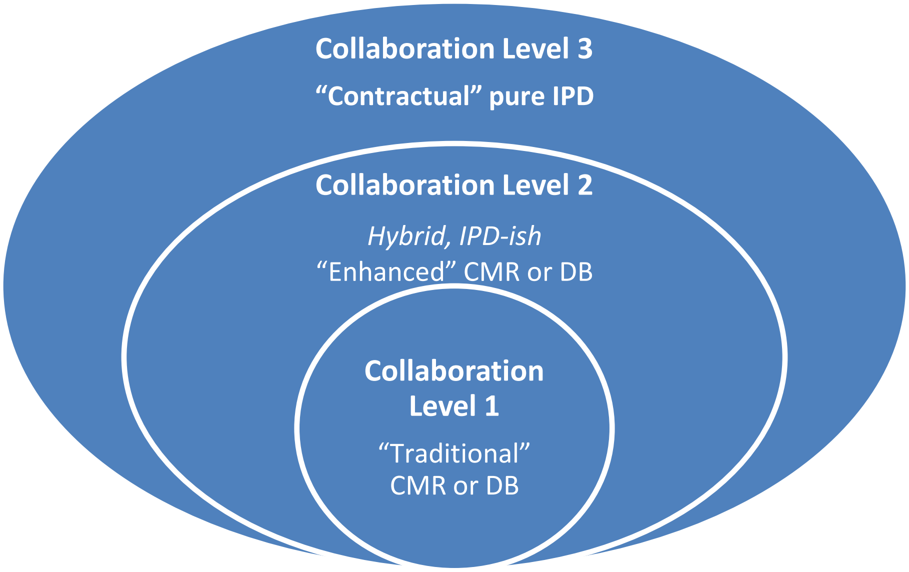
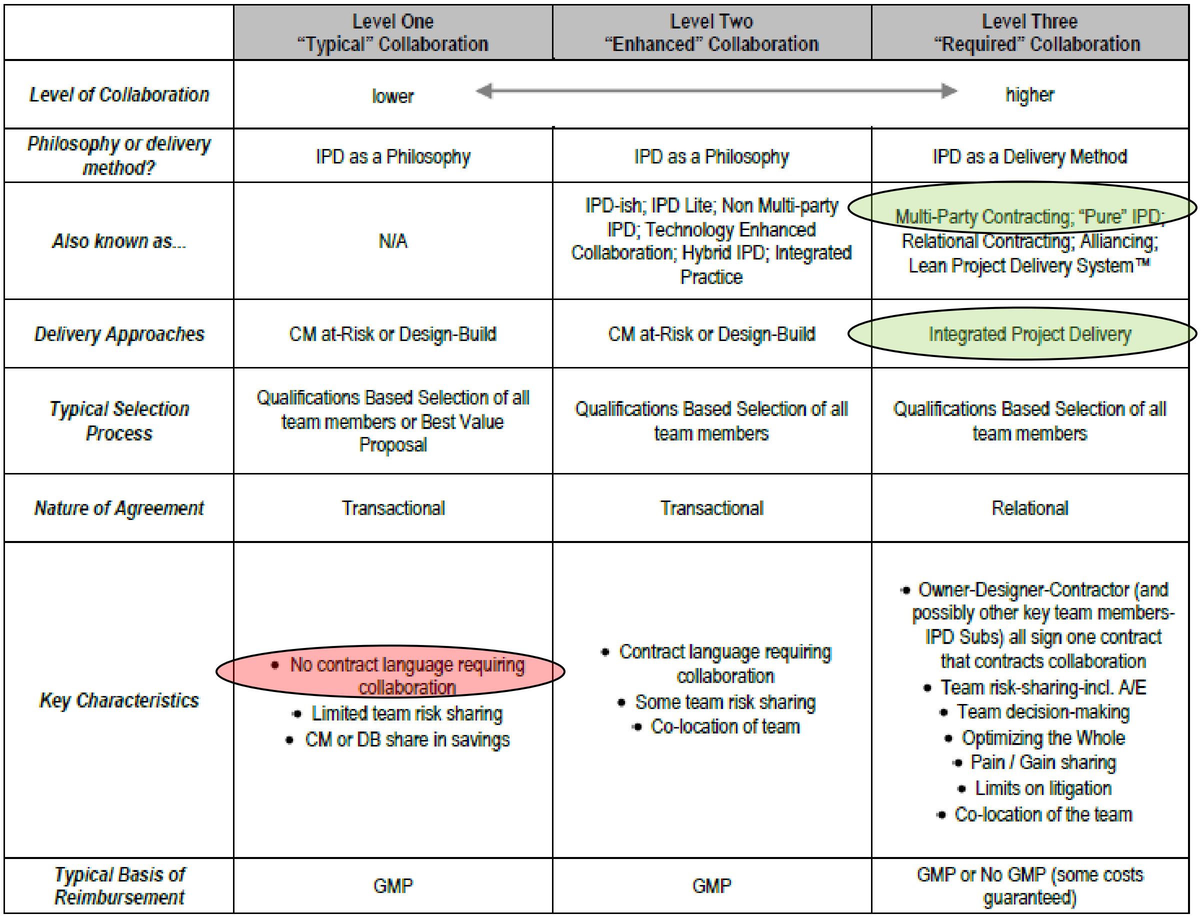


FIGURE 7: IPD Collaboration Levels Based on NASFA

The NASFA IPD definition is somewhat similar to the one adopted by AIA, including most of the same elements although grouped differently, as seen in the Key Characteristics row of *Figure 8*. It emphasizes the grouping structure for the project stakeholders; at the heart of the project is the project management and leadership team (or core group) that handles day-to-day decision making. Projects could have two other team levels; a higher-level executive team for circumstances where consensus is not reached by the core group, and lower-level project implementation teams (or cluster teams) responsible for designing, detailing and constructing specific aspects of the project.



	Level One "Typical" Collaboration	Level Two "Enhanced" Collaboration	Level Three "Required" Collaboration
Level of Collaboration	lower	← →	higher
Philosophy or delivery method?	IPD as a Philosophy	IPD as a Philosophy	IPD as a Delivery Method
Also known as...	N/A	IPD-ish; IPD Lite; Non Multi-party IPD; Technology Enhanced Collaboration; Hybrid IPD; Integrated Practice	Multi-Party Contracting; "Pure" IPD; Relational Contracting; Alliancing; Lean Project Delivery System™
Delivery Approaches	CM at-Risk or Design-Build	CM at-Risk or Design-Build	Integrated Project Delivery
Typical Selection Process	Qualifications Based Selection of all team members or Best Value Proposal	Qualifications Based Selection of all team members	Qualifications Based Selection of all team members
Nature of Agreement	Transactional	Transactional	Relational
Key Characteristics	• No contract language requiring collaboration • Limited team risk sharing • CM or DB share in savings	• Contract language requiring collaboration • Some team risk sharing • Co-location of team	• Owner-Designer-Contractor (and possibly other key team members-IPD Subs) all sign one contract that contracts collaboration • Team risk-sharing-incl. A/E • Team decision-making • Optimizing the Whole • Pain / Gain sharing • Limits on litigation • Co-location of the team
Typical Basis of Reimbursement	GMP	GMP	GMP or No GMP (some costs guaranteed)

FIGURE 8: IPD Collaboration Levels (NASFA 2010)

The report states that adopting IPD drives waste out of the project, reduces or eliminates changes, improves schedules, and results in avoiding conflicts and resolving disputes by the core group. However, no analysis was completed using actual project data. Finally, the report comments on the AIA 2010 case studies, noting that five projects (out of six) used IPD as a *Level 3* delivery method with a multiparty contract. The sixth project, “Walter Cronkite”, although employing highly collaborative IPD principles, was *Level 2* or *IPD-ish* (NASFA 2010). This study adopts the NASFA definition of contractual IPD; the term *IPD-ish* will be used in the remainder of this dissertation to describe projects that were originally referred to as IPD, although they do not include all the necessary characteristics of a true IPD project, namely a multiparty contract.

Mossman et al. (2010) combined two graphs from Eckblad et al. (2007) and Lichtig (2007) to compare traditional and integrated delivery timelines, and their impact on developing a shared understanding of the project by the whole team. This comparison, shown in *Figure 9*, is interesting as it summarizes some of the key differences between IPD and the traditional DBB delivery system. One major difference is the early involvement of project stakeholders in IPD, which results in an earlier common understanding of the project as well as a reduction in clashes and in project schedule. However, this last statement is not scientifically proven and the yellow curve that represents the common understanding of the project is not based on actual project data.

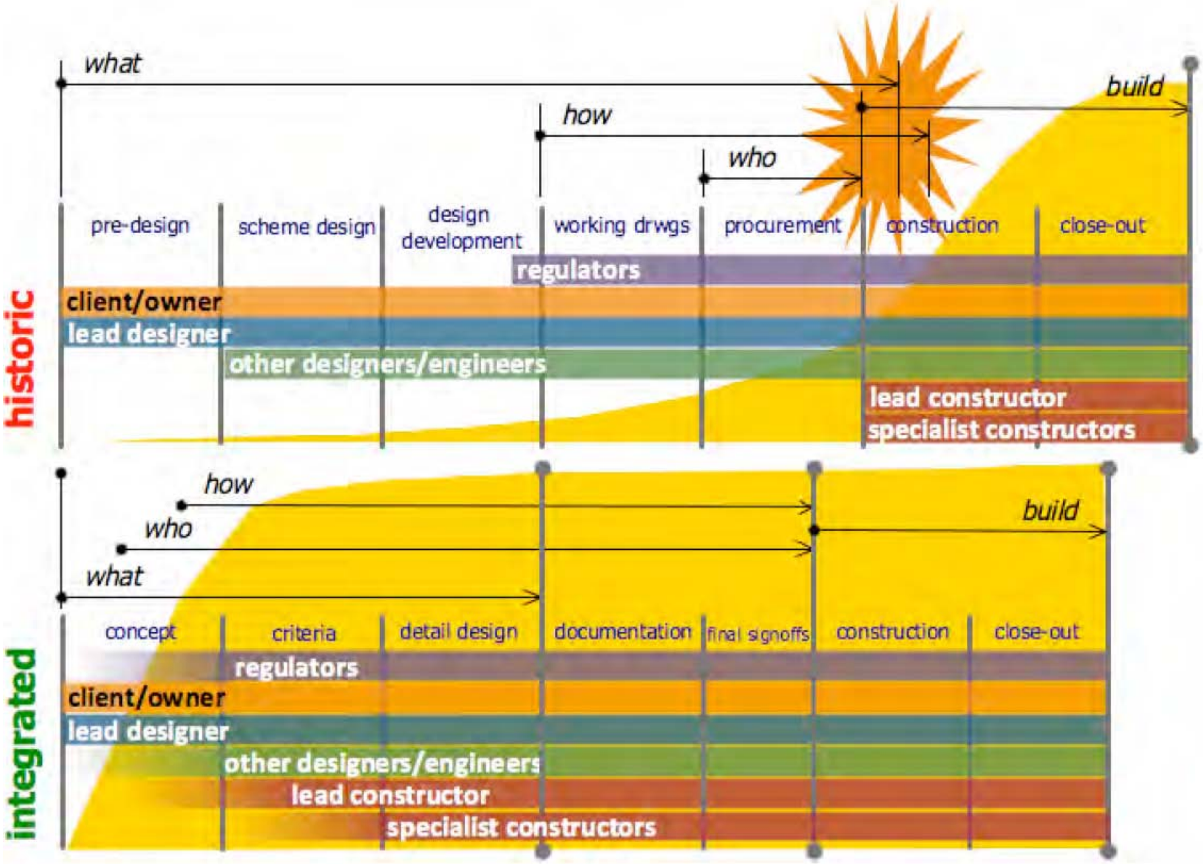


FIGURE 9: Comparing traditional to integrated delivery (Mossman et al. 2010)

Mossman's paper also discusses different delivery tools and techniques, namely set-based design and target value design, illustrated with examples mostly from the industry. For example, a discussion on insurance explains how the general contractor for the London Heathrow Airport's Terminal 5 bought a single project insurance to cover all parties, and worked with the different parties to manage and reduce project risk, as opposed to just shifting the risk between the different parties.

To summarize this first section of the literature review, one can see there are different definitions of IPD, some more stringent than others. However, none of the previously listed definitions are perfect because they are not consistent with the two key elements of a project delivery system definition: relationships and timing of engagement of the key stakeholders. Summarizing the definitions found in the literature, and based on the project data collected, this study develops its own definition of IPD. This definition is in essence very similar to the NASFA (2010) definition of contractual IPD, and it also has a format consistent with the definition of a project delivery system: IPD is *a delivery system distinguished by a multiparty agreement and the very early involvement of the key participants*. The literature also reveals a myriad of benefits stemming from the use of IPD, but the claims are often unsupported or only based on a limited number of case studies.

2.1.1 Differences Between IPD and Lean Project Delivery System (LPDS)

Although IPD and LPDS are sometimes used interchangeably in the construction industry, some researchers draw boundaries between these two systems. The requirements for LPD consist of three basic domains: (1) an operating system, (2) relational commercial terms, and (3) a collaborative project organization. On the other hand, IPD does not necessarily require the use of specific tools, or what is referred to as an operating system. IPD is defined by the early involvement of key participants signing the same multiparty contract, regardless of what “operating system” is being used.

The Construction Industry Institute published a research report presenting a roadmap for lean implementation at the project level (CII 2007). The report describes LPDS as a system where representatives of every stage in the lifecycle of the facility are involved from the very beginning of the project, stressing on the Set-Based Design process and feedback loops throughout the building lifecycle. Ballard (2008) states LPDS is still under development. He discusses how concepts, such as set-based design, evidence-based design, and target costing, allow the project teams to help customers decide what they want, which ultimately results in a superior outcome in terms of value for the customer. His hypothesis is that through the Lean Project Delivery System, one can provide facilities more fit for the customer's purpose at a lower cost. The paper features an LPDS case study of the ThedaCare Shawano Clinic, which was completed 3.5 months ahead of schedule with a cost 17.6% below the market benchmark. The author expresses the need for descriptive research that enables a better understanding of *what works and what does not*, in terms of delivery characteristics.

This gap will be addressed in this study. Although this study focuses on IPD performance (not LDPS), it will also present results for specific delivery characteristics that are proved to increase performance. *Chapter 6* includes a detailed discussion of the results.

2.2 Comparing Project Delivery Systems

There are numerous systems being used to deliver buildings around the world. However, Branca (1987) identified the three delivery systems most commonly employed in the U.S. construction industry: (1) traditional design-bid-build (DBB), (2) construction management at risk (CMR), and (3) design-build (DB). There is an abundance of construction delivery literature comparing the performance of DBB, CMR, DB, and other delivery systems. The studies differ based on specifics, such as the types of projects studied and the performance metrics used.

Pocock (1996) compared the performance of traditional and alternative project delivery approaches using 209 military construction projects. He also measured the degree of team integration, which he demonstrated was directly impacting project performance. The metrics used to compare delivery types were: (1) schedule growth, for which partnered projects were the most successful; (2) cost growth and (3) design deficiencies, both of which were dominated by DB; and (4) modifications, at which combination projects (hybrid use of delivery systems) had an enhanced performance. Traditional DBB projects were shown to perform the worst when comparing schedule growth, modifications, and design deficiencies. Other performance metrics considered for the study included: cost of claims, value engineering savings, safety, and average user satisfaction rating.

Several studies compared the performance of the DB delivery system to the traditional DBB, most notably Molenaar (1995), where the focus was only on the public sector, and Songer and Molenaar (1996) that focused on owner attitudes toward DB and analyzed seven DB selection factors. Oberlender and Zeitoun (1993) also included CMR in the comparison, in addition to DB and DBB, identifying early warning signs of cost and schedule growth. However, the study was not able to correlate the effect of different delivery types on cost and schedule growth.

Bennett et al. (1996) used both univariate and multivariate analyses to compare cost, schedule and quality performance of 332 DB and DBB projects in the UK. They identified variables that influence these performance metrics, and their study showed DB projects result in improvements of delivery speed by 30% and construction speed by 12%, as well as a 13% reduction in unit cost.

All of the above studies set the stage for a CII national study conducted by Sanvido and Konchar (1998). In the CII study, DB showed a superior performance over CMR, which in turn performed greater than DBB. The authors studied 351 projects in the following U.S. general building market sectors: light-industrial, multi-story dwelling, simple and complex office, heavy industrial and high technology. The metrics studied for which the results were statistically significant included unit cost, construction speed and delivery speed. Other metrics had less statistical significance including cost growth, schedule growth, turnover quality and systems quality. It is important to note that in addition to univariate analyses, multivariate analyses were also performed in Konchar and Sanvido's study to understand significant differences in delivery performance and to identify the variables that have the biggest impact on project performance.

Molenaar et al. (1999) studied DB performance in the public sector by looking at 104 projects. The study considered numerous project variables: owner experience, level of design completion, design-builder selection, contract type, method of award, and DB process variations in the public sector. Performance metrics were both quantitative, including cost and schedule growth, and qualitative, including the measurement of quality with respect to the user's expectations, construction administrative burden, and owner satisfaction with the overall project. Quantitative results show that 59% of the DB projects experienced less than 2% cost growth, and 77% of the DB projects experienced less than 2% schedule growth. Qualitative results show that most owners were satisfied with the performance of DB.

Another CII study funded by the National Institute of Standards and Technology (NIST) was conducted in 2002 by Thomas et al. to compare the impact of DB and DBB on project performance by using CII benchmarking data from 617 projects. The study confirmed Sanvido’s findings from 1998 in terms of DB superiority to DBB, this time looking at schedule, changes, and rework. The results for these three metrics were statistically significant for owner-submitted projects, while contractor-submitted projects had a significantly better performance only for change performance, although outperforming DBB in rework as well. DB also had a superior “practice use” record, which the authors defined as using best practices, such as pre-project planning and constructability. DBB projects only outperformed contractor-submitted DB projects in schedule, where the difference was statistically significant. However, as previously stated, owner-submitted DB projects outperformed DBB in the schedule metric. Overall, there were no statistically significant differences for any of the cost metrics. This study is interesting because the superior performance was not across the board for DB; rather the results depended on which party submitted the project data.

Ibbs et al. (2003) also studied DB and DBB using data from 67 CII projects by comparing cost growth, schedule growth and productivity as the performance metrics. Schedule growth results confirmed previous findings (e.g. Konchar and Sanvido 1998; Molenaar et al. 1999) on the superiority of DB compared to DBB. However, DB was not found superior to DBB when looking at cost growth and productivity.

Riley et al. (2005) studied the effects of using DB mechanical contractors (DBMC) on green building projects through three case studies. Their research showed that early involvement of DBMC resulted in a significant improvement over the DBB approach through initial cost savings and a more efficient final product. One significant trend the study notes is the DBMC’s willingness to adopt new technologies and innovative solutions, such as lean principles.

Debella and Ries (2006) studied the performance of multi-prime delivery compared to (1) single-prime delivery and (2) multi-prime with CM agent. The scope was limited to data from 105 projects for public school districts. The two main findings were that construction speed for multi-prime with agent was faster than the two other delivery types, and single-prime delivery had less litigation cases than both multi-prime delivery types. Depending on the data set used, the metrics dealing with unit cost, cost growth, schedule growth and change orders, had inconsistent results and most of the performance differences between delivery systems were not statistically significant.

More recently, Rojas and Kell (2008) conducted a study focusing on cost performance of CMR and DBB project delivery systems. Their scope also was limited to delivering public schools, but this time in the U.S. Pacific Northwest. The study looked at 273 DBB and 24 CMR projects. The results show:

- no statistically significant difference between CMR and DBB in construction change order costs;
- CMR school project costs exceeded the GMP in 75% of the cases;
- a statistically significant finding that DBB averages less cost growth than CMR; and
- when comparing school to non-school projects, CMR school projects experienced a lower average change order ratio than non-school projects.

These findings challenge Sanvido's findings regarding CMR cost performance, but only apply to the limited scope of public schools in the U.S. Pacific Northwest.

Korkmaz et al. (2010a) studied the influence of project delivery methods on achieving sustainable high performance buildings. Looking at 12 in-depth case studies covering DBB, DB and CMB, the study investigated the effects of project delivery attributes on project performance at construction completion. Korkmaz et al. found that CMR and DB outperform DBB projects overall; one specific result suggests that projects adopting the DBB method display higher cost growth. Similar to the Pocock (1996) study, the Korkmaz study reveals that the level of integration in the delivery process affects final project outcomes. Interestingly, the results show that project delivery attributes, such as owner commitment and timing of participant involvement, affect the level of integration more than the characteristics of the project delivery method selected.

Another study by Korkmaz et al. (2010b) identified key metrics for sustainable building project delivery in the United States. Their study examined more than 100 variables in 40 projects delivered through CMR, DB and DBB systems. The results show that CMR and DB outperform DBB in the delivery speed metric. This study will be covered in greater depth in the *Section 2.3*.

Most recently, one of the latest studies comparing delivery systems contrasted IPD to other delivery systems. Cho and Ballard (2011) studied (1) whether the Last Planner System, a production control tool that smoothes construction project task workflow, improves project performance, and also (2) whether IPD projects show different project performance than non-

IPD projects. While it was shown that the Last Planner System improves performance, the authors were not able to find significant differences in performance between IPD and non-IPD projects. They performed t-tests on data from 49 projects, but the paper does not provide any information on the dataset (e.g. project types, sizes, and locations). Additionally, the authors' definition of project performance is restricted to reductions in time and cost, which might not prove comprehensive enough when studying the value-adding IPD, as discussed earlier.

Table 3 shows a summary of the studies most relevant to this research. The table highlights major focus areas, such as the delivery systems compared in each study, the type of data collected, the type of comparative analyses made, and the number of projects used for the analysis. The last column of *Table 3* shows how this current research effort compares to the previous studies.

To summarize this section of the *Literature Review* chapter, most studies provide some evidence for more collaborative delivery systems being superior to less collaborative systems. The statistical significance of the results was stronger in some cases, depending on the type of construction and the scope of the studies performed. However, there is no literature showing statistically significant performance differences for the emerging IPD system.

TABLE 3: Literature Summary of Comparing Delivery Systems

Attributes	Oberlender	Molenaar	Songer	Pocock	Bennett	Sanvido	Molenaar	Ernzen	Thomas	Ibbs	Debella	Rojas	Hale	Korkmaz	Cho-Ballard	*El Asmar*
Year	1993	1995	1996	1996	1996	1998	1999	2000	2002	2003	2006	2008	2009	2010	2011	2012
Delivery Systems Compared																
DBB	*	*	*	*	*	*		*	*	*	*	*	*	*		*
CMR	*					*						*		*		*
DB	*	*	*	*	*	*	*	*	*	*			*	*		*
Others				*						*	*				*	*
IPD															*	*
Data Collection																
Quantitative data	*			*	*	*	*	*	*	*	*	*	*	*	*	*
Qualitative data					*	*	*							*		*
Type of Comparative Analysis																
Univariate analysis	*	*	*	*	*	*		*	*	*	*	*	*	*	*	*
Multivariate analysis					*	*								*		*
Number of variables	33	15	7	4	45	58	5	7	29	6	7	16	11	>100	3	304
Private sector			*		*	*			*					*		*
Public sector	*	*	*	*	*	*	*	*	*		*	*	*	*		*
Facility classification					*	*		*	*		*	*	*	*		*
Number of projects	106	NA	NA	209	332	351	104	2	617	67	105	297	77	40	49	35

Parts of this table are adapted from Sanvido and Konchar (1998).
Only the name of the first author was included in the table for sizing issues; the full citations can be found in Appendix A – Bibliography.

2.3 Key Variables

While reviewing the literature, special attention was given to highlighting the key variables used in each previous study. These variables can be classified in three groups: (1) performance metrics, the dependent variables which are after-the-fact measures of project success, such as project cost and schedule growth; (2) independent variables that affect the performance metrics and are manageable in a given project, such as the project delivery system, which will be tested against performance; and (3) control variables, such as building type or size, which also affect the performance but are unmanageable and therefore will be standardized or kept constant throughout the analyses.

2.3.1 Performance Metrics

As introduced in the *Methodology* section earlier, dependent variables are the factors studied and expected to change whenever the independent variables are altered. The project performance metrics will act as dependent variables for this study. Previous comparative studies, including some of those discussed earlier, included several performance metrics in their analyses. Additionally, several researchers specifically studied performance metrics, which also can have different names, such as: project success factors, success criteria, or key performance indicators.

Songer et al. (1996) investigated owners' attitudes towards both success criteria and selection factors for DB procurement by surveying 137 owners in the United Stated and the United Kingdom. The success criteria identified include: *On Budget*, *On Schedule,* and

Conforms to Users Expectations. The DB selection factors identified include: *Establish Cost*, *Reduce Cost*, *Establish Schedule*, *Shorten Duration*, and *Reduce Claims*.

The metrics used by Sanvido and Konchar (1998) include: unit cost, construction speed, delivery speed, cost and schedule growth, turnover quality and systems quality. The study introduced a hybrid *cost per schedule* measure called *Intensity*; however, this metric was not highlighted in the final results. The metrics studied included some summary metrics. For example, turnover quality was a qualitative summary metric that combined individual ratings received for the following factors: facility startup, the number and magnitude of call backs and the operation and maintenance cost for the building. Each of these turnover quality metrics was scored on a scale of 10, the aggregate score being 30. System quality was another summary metric that combined three indices: (1) the performance of the envelope, roof, structure, and foundation; (2) the interior space and layout; and (3) environmental systems. Similarly each was scored on a scale of 10 leading to a maximum score of 30 for system quality. The final qualitative metric was process equipment and layout quality.

Chan et al. (2002) surveyed construction literature from 1990 to 2000 to develop a framework for DB success criteria. The authors grouped the criteria under three project phases: preconstruction, construction, and post-construction. The criteria also were grouped based on two types of factors: objective and subjective. The most significant measures found were time, cost, quality, and satisfaction of key project participants. Other measures that also were significant include: profitability, technical performance, productivity, and environmental sustainability. While these success criteria were developed for DB, they also can be used to measure the performance of projects delivered through other delivery systems.

The research study on project success completed by Menches and Hanna (2006) also was reviewed for performance metrics. Menches and Hanna quantified project success from the project manager's perspective. The authors evaluated the performance of electrical contractors on 55 projects in the United States. The performance measurement index included six variables: actual percent profit, percent schedule overrun, amount of time given, communication between team members, budget achievement, and change in work hours.

Rankin et al. (2008) first identified a set of performance metrics for measuring performance of the Canadian construction industry. The authors combined metrics for both the construction phase and for an extended timeline of building life. In doing so, they covered seven performance areas: cost, time, scope, quality, safety, innovation, and sustainability. A pilot study was conducted to verify the metrics, and, as expected, results showed that cost, time, scope, and safety information was readily available, while quality, innovation and sustainability metrics required considerable additional effort to obtain.

Sands (2010) explains how commodity-based standards and procurement practices prohibit innovation and lead to high-cost and low-performance buildings through unintended consequences. He presented two performance-based whole-building standards and measures, the cost effectiveness index (CEI) and the building performance index (BPI), and argued that these will lead to innovation and high performance buildings. Scenarios based on both commodity-based and performance-based standards are shown in *Figure 10* to demonstrate the need for including building lifecycle metrics when assessing project performance. While Sands' exact two indexes were not used for this study, several components of these indexes were included.

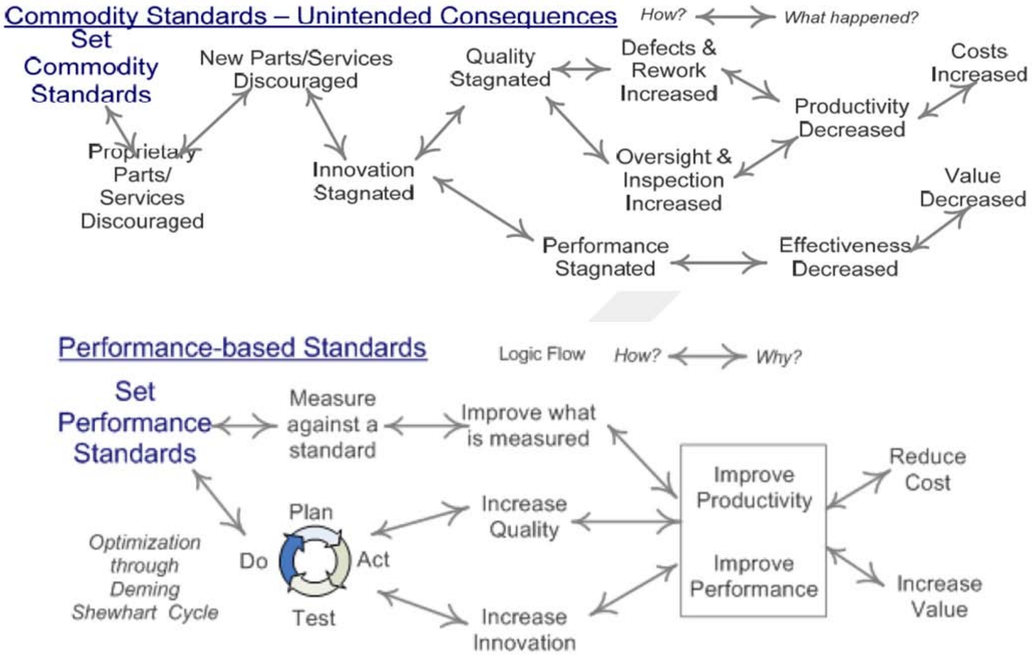


FIGURE 10: Commodity versus Performance-based Standards (Sands 2010)

Molenaar and Navarro (2011) looked at key performance indicators in highway design and construction through four DB case studies. The authors identified six different performance areas, some of which were applied to this study: cost and schedule, quality, safety, and environmental stewardship. The remaining two areas, public information and traffic reliability, while important for public horizontal construction, do not apply to the scope of this study, which will be detailed in *Section 2.6*.

Table 4 displays a sample of important performance metrics used to compare construction projects in the literature. The approaches of the different studies varied from a general level to a more detailed level of metrics; one example is using the broad metric *On Schedule* versus the more specific set of metrics *Construction Speed*, *Delivery Speed*, and *Schedule Growth* when measuring project schedule. The type of data is also a differential factor between the studies, with some advocating objective evaluations and others using subjective evaluations for the same metric.

TABLE 4: Literature Summary of Performance Metrics

Performance Area	**Performance Metric**	Pocock	Sanvido	Debella	Rojas	Cho/Ballard	Songer	Menches	Chan	Molenaar	Sands	EISA
Cost	Unit Cost		*	*	*		*		*	*		
	Cost Growth	*	*	*	*	*	*	*	*			
	Budget Factor				*							
Schedule	Construction Speed		*	*			*	*	*	*		
	Delivery Speed		*									
	Schedule Growth	*	*	*		*	*	*	*		*	
	Schedule Factor											
Safety									*	*	*	
Productivity	Productivity Factor								*		*	
	% Plan Complete											
Business	Profit							*	*			
Quality	Systems		*						*	*	*	
	Turnover		*									
	Defects								*			
	Building Ownership								*			
Occupants	Satisfaction								*			*
Use-ability / Value	Program Spaces										*	
	Functionality								*		*	*
	Suitability for Purpose						*		*			
	Durability											*
Operations & Maintenance	Maintenance		*								*	*
	Service Cost										*	
	Energy Consumption										*	*
Other	Claims			*			*		*			
	Changes / Modifications	*		*				*			*	
	Material Waste								*		*	

Only the name of the first author was included in the table for sizing issues; the full citations can be found in Appendix A – Bibliography.

After reviewing the table, an interesting occurrence is visible: many studies focus heavily on schedule and cost performance metrics – only two out of a minimum of 10 identified performance areas – while largely disregarding the remaining performance metrics. It is only recently that authors stressed on the importance of measuring several additional variables.

A comprehensive list of performance metrics that was originally used for this research is shown in *Table 5*. The table is organized to represent 13 performance areas throughout the three mega-phases of a building lifecycle: Making, Using and Operating, and Changing and Demolishing. The metrics cover major performance areas important to the different project stakeholders, ranging from cost and schedule to project waste and business-related performance. Each of these performance areas may include more specific performance metrics. Performance metrics were grouped depending on the type of data to be collected: 31 quantitative metrics and 11 qualitative metrics, totaling 42 original dependent variables for the study.

Initially, the plan was to collect data for each of the metrics listed in *Table 5*. However, after receiving feedback from the industry panel regarding which metrics will be accessible in a reasonable amount of time and effort, the list was modified. Additionally, after the data was collected, some metrics needed to be removed because of missing fields across several projects in the dataset. The final list of performance metrics will be discussed in *Chapter 3*.

TABLE 5: Performance Metrics Across Life Cycle Stages

Life	Performance Area	Performance Metric	Quantitative	Qual.
Making (Design and Construction)	Cost	Unit Cost	*	
		Cost Growth	*	
		Budget Factor	*	
	Schedule	Construction Speed	*	
		Delivery Speed	*	
		Schedule Growth	*	
		Schedule Factor	*	
	Safety	OSHA recordables	*	
		Lost Time Injuries	*	
		Fatalities	*	
	Productivity	Productivity Factor	*	
		Labor Factor	*	
		PPC	*	
	Waste	Rework, RFIs, Claims, Changes and Modifications, Resubmittals, Field Conflicts, Scoping Issues, Deficiency Issues, Material Waste	*	
	Businesses-specific	Gross Profit	*	
		Image & Return Business		*
		Innovation		*
	Quality	Program Spaces	*	
		Systems		*
		Punchlist	*	
Using, Operating and Maintaining		Turnover		*
		Warranty Costs	*	
		Latent Defects	*	
		Building Ownership		*
	Occupants-related Factors	Occupants Satisfaction		*
		Absenteeism		*
	Functionality, Suitability for Purpose, Durability			*
	Operations & Maintenance	Maintenance Requests	*	
		Service Cost	*	
		Energy Consumption	*	
		Water Consumption	*	
Change / Demo.	Changeability Requirements			*
	Demolition Requirements			*

2.3.2 Independent and Control Variables

This study includes a multivariate analysis, which analyzes the effect of factors other than the delivery system. It is important to identify these factors or attributes because they also might affect project outcome. These factors are clustered in two groups.

The first group consists of independent variables. Independent variables are the selected variables of interest for which relationships with the performance metrics are to be determined. In addition to the project delivery system, independent variables include project team characteristics, involvement and collaboration, contract conditions, and other manageable aspects of the project.

The second group consists of control variables or factors that strongly influence the outcomes but are mostly unmanageable by project leadership, such as project type, project size, acces to the site, and external conditions. Control variables are held constant or used to normalize the data in order to more effectively test the relative impact of independent variables.

Previous studies have reviewed and analyzed independent and control variables of interest to this study. For example, Pocock et al. (1996b) demonstrated that project team integration affects performance. In this dissertation, factors also are identified to quantify integration, based on the different stakeholders' involvement in various phases of the surveyed projects.

Chan et al. (2004) developed a conceptual framework for critical success factors. The authors conducted a comprehensive survey of the construction literature, identifying five groups of independent variables that can affect project success: project-related factors, project procedures, project management actions, human-related factors, and external environment. *Figure 11* summarizes the identified variables.

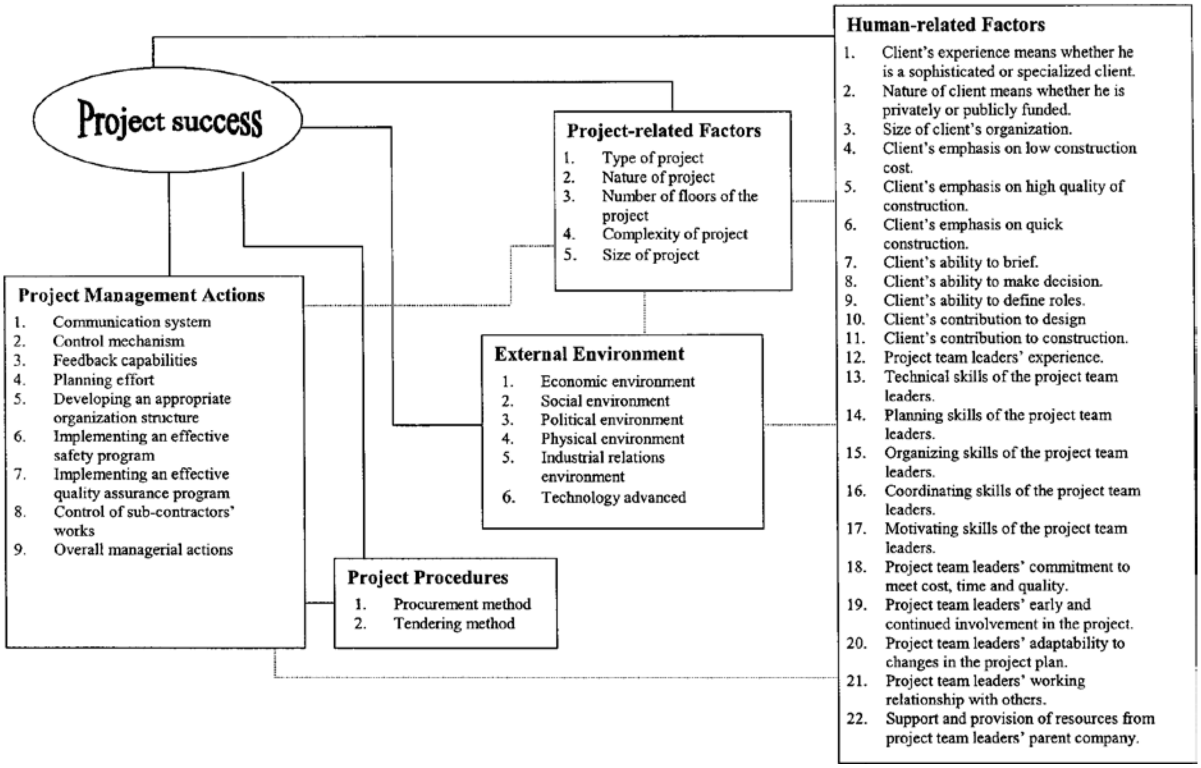


FIGURE 11: Critical Success Factors (Chan 2004)

Konchar and Sanvido (1998) also identified several project characteristics that influenced project performance. The most important factors are shown in the first column of *Table 6*. Korkmaz et al. (2010b) identified seven independent variables and several control variables in their high-performance building study. The variables that resulted in the most significant differences in the analyses are summarized in the second column of *Table 6*. One can see there are common variables identified in the results of both studies, such as the project size, percent of design complete, and stakeholders past experience.

TABLE 6: Independent and Control Variables

Konchar and Sanvido (1998)	Korkmaz et al. (2010)
Project size	project size
Percent design complete before the construction entity joined the project team	Level of completion of the construction document
Subcontractor experience with the facility type	Owner, design-builder, and subcontractor experience with a similar facility
Project team communication	Timing of involvement of contractor, and commissioning agent
Facility type	Owner type
Contract unit cost	Use of construction mockups
Project delivery system	Quality of workmanship for the different building systems
Presence of onerous clauses within the team contracts	
Level of new construction	
Project complexity	

The variables in *Table 6* are among the factors highlighted in the analysis to ensure a comprehensive understanding of the differences in performance related to IPD projects. Questions related to these and other variables were included in the survey to quantify these variables and account for each in the analysis. A complete list of all variables for this study is presented in *Chapter 3*, and also can be found in the survey shown in *Appendix C*.

2.4 Research Opportunities

Conducting a review of the relevant literature was beneficial in understanding what has been previously accomplished in the area of project delivery systems. It also helped uncover gaps and reveal three main research opportunities: (1) there is no consistent IPD definition; (2) there is a lack of studies showing performance differences for IPD; and (3) a comprehensive project performance metric does not exist.

The finding that there is no consistent definition of IPD presents an opportunity to unify and standardize a definition for IPD, based on the literature as well as the data collected from IPD projects. The lack of a consistent definition is not unique to IPD. As with any new concept, different definitions can exist and vary depending on the source. For example, long before IPD, the definition of DB also was changing. As for IPD, the literature shows several definitions, including some definitions that vary and evolve within the same organization. As with any new system, the definition of IPD needs to be standardized in order to provide a comparable baseline that others also can use.

Second, there is an obvious lack of studies showing significant performance differences for the emerging IPD system. Other delivery systems that have been around for a longer period of time have enjoyed several studies that compared their performance depending on the type of projects and the various performance metrics of interest. The lack of studies targeting IPD is most likely due to the very recent existence of IPD. Because IPD is a recent development, there is only a small number of IPD projects completed. The only study that attempted to compare IPD and non-IPD projects shows no differences in cost and time performance. This dissertation fills the void by studying the significance in performance differences across several metrics to provide a more comprehensive understanding of IPD performance through a thorough comparison of IPD and non-IPD projects.

Lastly, this dissertation seizes upon another opportunity by developing a comprehensive performance metric for AEC projects. For example, how do project stakeholders decide whether a project should be classified as successful or not if it is completed on schedule and on budget, but has inadequate building quality or poor safety statistics? Currently, there is no single metric by which to judge a successful project. The *Project Quarterback Rating* was created to address this important research opportunity. Many performance metrics were identified throughout the literature review, and by combining these metrics in a single comprehensive performance metric, a simple comparison can be made using overall project performance.

2.5 Research Objectives

This study evaluates the performance of IPD projects by comparing them to projects delivered using other systems, such as CMR, DB and DBB. The focus extends beyond the commonly analyzed metrics of cost and time to include safety and quality, as well as less commonly studied metrics, such as changes, process inefficiencies, communication, and profit. This expansive analysis also allows the identification of the delivery characteristics that are responsible for causing differences in performance, such as contractual incentives, stakeholder involvement, and practices used during the delivery process. This study has six specific objectives:

1. Provide a standardized definition of *Integrated Project Delivery*, which can be used in future studies as a common basis of comparison.

2. Compare IPD projects to other delivery types based on the several metrics identified in the literature review, supplemented by new metrics and finalized through industry interactions. This comparison is used to understand if IPD provides a superior performance and is worth the use,

research and investment. There are two specific questions that are answered by the comparison results:

   a. Is IPD superior to other delivery types based on each individual performance metric?

   b. And if so, how much improvement can be seen with IPD?

3. Develop a unique comprehensive project success rating that combines key performance metrics; the *Project Quarterback Rating* (PQR) will allow overall performance comparisons of AEC projects.

4. Determine what delivery characteristics are responsible for superior overall performance for IPD projects. Several delivery characteristics grouped under *Contractual Terms*, *Social Tone*, and *Tools and Techniques* (3Ts) will be investigated to understand which characteristics ultimately lead to superior performance.

5. Develop explanatory models for project performance based on the project delivery system and the delivery characteristics applied in a given project.

6. Based on the previous analyses, provide recommendations regarding specific issues, such as what characteristics to apply in order to maximize performance. IPD benchmarks also will be highlighted to help achieve the expected superiority.

2.6 Research Scope

The scope of this study is shaped by the data collected and focuses on the contractors' perspective. Few public sector projects are included, since special legislation is typically needed to deviate from the traditional procurement method of picking the lowest bidder. Therefore, and as anticipated, most data was received from projects with progressive private owners that are willing to experiment with new techniques. However, universities were a noticeable exception as use of private funds may give them some flexibility in project delivery methods.

Vertical construction was targeted, as opposed to horizontal construction. Based on the current use of IPD, the most common types of buildings are large-scale high complexity institutional facilities, such as hospitals and research laboratories. These tend to be relatively large complex projects, justifying the rather intense initial effort required to carry out IPD. Therefore, hospitals and research laboratories were some of the first types of buildings on which IPD was applied.

Data was collected only from projects that were completed after 2005, due to IPD being a recent system, and to provide a fair comparison between IPD and non-IPD projects. It should be stressed that data from non-IPD projects was collected as a means of comparison. For research efficiency purposes, most of the data was collected from the construction manager or general contractor (CM/GC), because the CM/GC typically has access to most of the key project data. This study specifically focuses on metrics related to the delivery phase from the CM/GC's perspective.

Chapter 2 provided a review of the literature, highlighted research opportunities, and presented the objectives and scope of this study. Next, *Chapter 3* discusses the *3T's*, a classification of the final delivery variables used in the analysis, as well as the performance metrics, and presents an overview of the collected data characteristics to lay the ground for the statistical analysis.

Chapter 3. Exploring the Research Variables: The 3Ts and Performance Metrics

The previous chapters introduced this study and its motivation, reviewed the literature, stated the problem statement and the research objectives, and described the methodology used to address the research questions. This chapter, which builds on the earlier discussions, serves two purposes. First, it presents the final variables used in this study. These include delivery characteristics grouped in three distinct areas, as well as performance metrics grouped under nine areas. Second, this chapter provides an initial look at the collected data by providing some general information about the dataset and a discussion of the responses and distributions of several delivery characteristics.

3.1 Project Delivery Characteristics

3.1.1 3T Motivation

This chapter begins with an introduction about project delivery systems and their associated characteristics. The introduction will pave the way for an understanding of how these delivery characteristics evolved into the *3Ts*, as introduced in this study.

A project delivery system defines the relationships, roles and responsibilities of parties involved in a project. It establishes an execution framework sequencing design, procurement, and construction activities required to provide a facility. The project delivery system also can describe practices and techniques of management that are used by the project team (Ireland 1984; Sanvido and Konchar 1999; Oyetunji and Anderson 2006).

Lean construction literature suggests a different definition of project delivery systems with three basic domains: commercial terms, an operating system, and a project organization (Thomsen et al. 2010; Smith et al. 2011; Cho et al. 2010). The project commercial terms reflect the legal relationships between the various parties participating in the process, and some authors suggest relational contracts will result in more successful projects than transactional contracts. The project operating system consists of the Last Planner System and other supporting project management techniques. The project organization domain creates a unified project culture.

There are two concerns with this triple-factor definition. First, the project commercial terms usually dictate the project organization, and are historically embedded in the *AIA Conventional A201 Family* framework. The responsibilities of the parties signatory to the contract include how these parties are organized. Second, and more significantly, this definition does not sufficiently recognize the human and behavioral aspects embedded in team environments and project-based organizations. In fact, previous research (Nelson et al. 2008) states “social, political and cultural change is needed to effectively address today’s complex engineering and environmental challenges,” and several industry interactions indicate that two of the major impediments to a systemic improvement in AEC project delivery are social:

- Professional *Cognitive Impairments*: this is true especially for designers and constructors. Many professionals in the capital project industry have a preexisting bias of deeply entrenched industry mindsets.

- Institutional *Adaptive Challenges*: institutions including owners, the American Institute of Architects, the Association of General Contractors, the National Society of Professional Engineers, suppliers, and many others, face adaptive challenges. Large adaptive changes in the delivery process are required to remove the waste associated with traditional project delivery methods (Tradeline 2009).

A more systematic and inclusive manner for characterizing the three delivery domains can be modeled after the triple bottom line, as introduced by Elkington (1998): social, economic, and physical. Therefore, AEC project delivery systems can be characterized by the following three domains:

1. Social: the culture by which the team normalizes and operates, including the team experiences, the vocabulary used, the human accountability system, and the timing and degree of stakeholders' involvement – essentially the *tone* by which the project team functions to address cognitive impairments and adaptive challenges.

2. Economic: the financial and legal terms by which the project is executed, including the procurement, contract type, compensation, incentives, and risk management – essentially the *terms* by which the project is delivered.

3. Physical: the operating system or functional domain, including Building Information Modeling and Lean construction tools and techniques – essentially the *tools* used by the team to deliver the project.

Accordingly, these three delivery domains can be referred to with an easy acronym, the *3Ts*: tone, terms and tools. The identification of these domains facilitates the upcoming taxonomy discussion. The manner in which the *3Ts* (inputs) affect the performance of a project (output) can be represented as:

$$\boxed{\text{Project Performance} = f\,(\text{terms, tone, tools})}$$

This formulation will be revisited and will serve as a basis for the development of performance models throughout *Chapter 6*. Since several new practices are being used in project delivery today, definitions of several terms that fall under the 3Ts, such as specific tools and techniques, are provided in *Appendix B*.

3.1.2 A Changing Industry

The need for a taxonomy of terms and a list of new definitions is evidence that the construction industry is changing. *Figure 12* shows how each of the *3Ts* is evolving. The project terms are shifting from the linear and sequential design-bid-build to the more integrated types of project delivery systems, first with CMR and DB, and now with IPD. The tone and social characteristics of projects are moving from the traditional segregated stakeholders to a more collaborative approach that relies on the early involvement of all players. The tools and technology used on projects also are advancing: one example is the move from two-dimensional drawing to three-dimensional designs with Building Information Modeling (more dimensions when including cost and schedule, etc.).

Although the intent of this study is to evaluate the performance of IPD, it is important to understand that there might be synergies between the different changing factors discussed.

For example, the use of IPD involves stakeholders early in the process, and could facilitate the use of BIM and increase collaboration between the several project teams. Therefore, the multivariate analysis will be key in understanding how these different characteristics, along with IPD, are impacting performance together. In fact, enhancements across all three domains might be needed to address the industry chronic problems introduced earlier in *Chapter 1*.

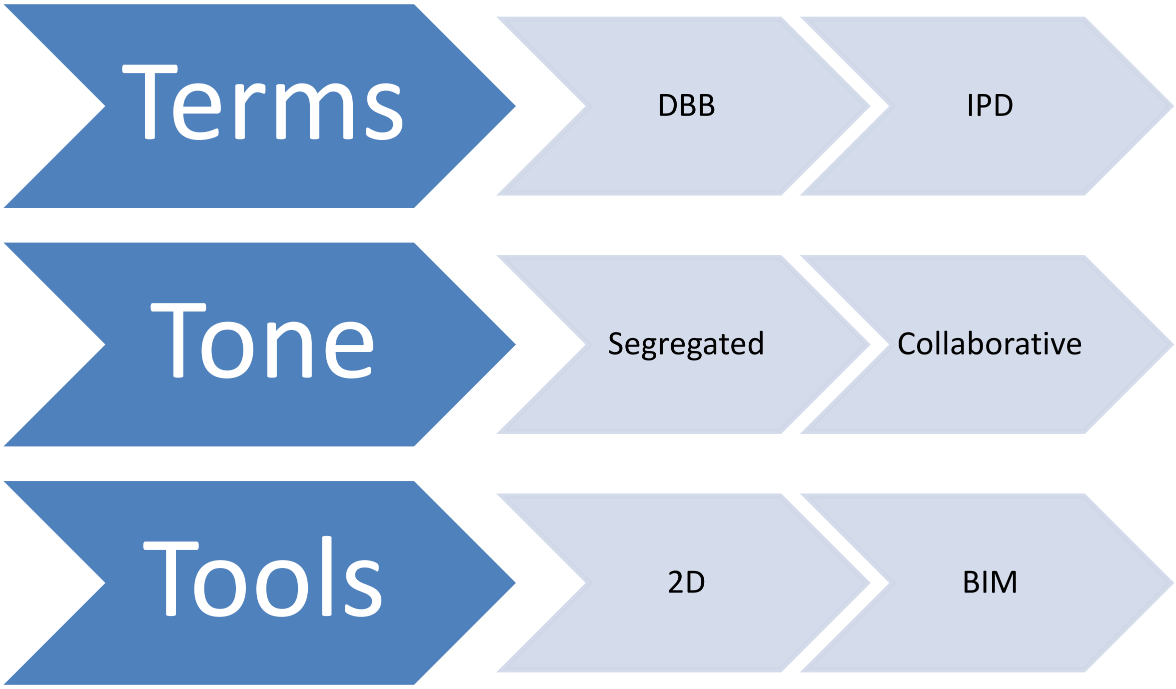


FIGURE 12: A Changing Industry

3.2 Taxonomy and Definitions of the 3Ts

In several instances throughout the literature, terms like IPD, lean construction, and lean delivery are used interchangeably, which might cause confusion for the reader. For example, in some instances lean construction was incorrectly defined as a project delivery system. This confusing terminology is not limited to one article, but is found in a number of publications relating to IPD.

An incorrect or confusing taxonomy makes it difficult to identify clear and effective solutions. Therefore, the development of this taxonomy section and a list of definitions available in *Appendix B* explains some of the key terms and their relationships, adding clarity to the discussion. *Table 7* shows the *3T's*: *Terms*, *Tone*, and *Tools*, as discussed earlier, with selected characteristics listed in the columns under each domain.

TABLE 7: 3T Taxonomy Table (not comprehensive)

Terms	**Tone**	**Tools**
Project delivery system (IPD, DB, DBB, etc.)	*Project Management Structure*	BIM (use, systems, reliance, etc.)
Contract type (IFoA, AIA, ConsensusDOCS 300, etc.)	• Core Group	*Lean construction tools*
Procurement Type	• Executive Team	• Last Planner System
Compensation Type	• Cluster Teams	• Target Value Design
Contractual Incentives	Stakeholder Involvement in various project phases	• Set-Based Design
Risk Allocation	Vocabulary	• Pull Planning
Fiscal Transparency	Team Prior Experiences	*Innovative Techniques*
	Physical Co-location	• Prefabrication
		• Mock-ups

The examples of delivery characteristics shown in *Table 7* are meant to illustrate the breadth of the three domains; all of these need to be investigated in addition to focusing solely on the highlighted project delivery system. *Appendix B* complements the taxonomy table and includes definitions of terms related to IPD and delivery systems. The definitions were aggregated from several sources, including CII (2005 and 2007), AIA (2007 and 2011), LCI (2011), Ballard (2008), Forbes et al. (2011), and Kim et al. (2011). These definitions were discussed with a research group and a panel of industry experts, and then adapted in order to provide a common baseline for defining the new concepts. This combination of taxonomy and definitions aims to clarify the confusion related to terms used when dealing with IPD.

When discussing definitions, one specific key definition needs to be discussed: the definition of IPD. The literature review clearly demonstrates there are different IPD definitions being used in the construction industry, including definitions evolving over time. For this study, none of the discussed definitions are adopted exclusively. Rather, characteristics existing in the the LCI, AIA, and especially NASFA definitions are combined to create a simple standard definition that will be used for this study, which also conforms with the core definition of a project delivery system (i.e. defining the relationships between project stakeholders and their timing of engagement). From these considerations, IPD is appropriately defined as:

> ***Integrated Project Delivery (IPD)****: a project delivery system distinguished by a multiparty agreement and the very early involvement of the key participants, ideally at 0% design but definitely before 10% design complete.*

In fact, two-thirds of the "IPD" projects surveyed have a multiparty contract, all of which had the contractors engaged before 10% of the design was complete. Moreover, 75% of these projects had the contractors engaged at 0% design complete, before the design had even started. This definition can be used to standardize IPD. Standardization is important because it allows everyone in the AEC industry to “talk the same language”, and therefore it will be easier to assess the performance of IPD projects.

As stated earlier, data was collected for all of the 3T delivery characteristics, which were later used in both univariate and multivariate analyses to obtain a more comprehensive picture of project performance. *Table 7* shown earlier was not meant to be comprehensive and only presents a few representative variables to illustrate the *3Ts*. Next, *Tables 8*, *9*, and *10* display a more inclusive list of the *3T* delivery characteristics, highlighting the terms, tone, and tools, respectively. Different levels of detail were included in order to enhance the organization of the variables presented: the *3T* high level that has three branches, an intermediate level, and the detailed survey level. The variables shown in *Tables 8*, *9*, and *10*, also can be found in the survey shown in *Appendix C*. The definitions and a glossary of terms for some of these variables can be found in *Appendix B*. One example of delivery characteristics is the use of incentives in the contract: the respondents are asked whether incentives were used, what they were based on, how they were funded, their total value and how it was allocated among project participants. Another example is measuring the use of BIM by asking whether the contract allows 3D models to be relied upon and whether the different parties use joint servers, as well as identifying specific building systems for which BIM was used. Additionally, the respondents are asked to rate the use of BIM for a list of 17 potential tasks, such as visualization and clash detection.

TABLE 8: 3T Delivery Characteristics (Contractual Terms)

3T	Intermediate Level	Survey Level
ORGANIZATIONAL: TERMS	Project Delivery System (PDS)	Type of PDS
		If IPD, Type of IPD contract
		If IPD, Multiparty contract?
		If IPD, Parties in contract?
		If IPD, Liability waivers?
	Compensation	CM/GC
		Subcontractors
		Architect/Engineer
		Design-Builder (if applicable)
		GMP establishment time
	Incentives	Incentives
		Based on?
		Metrics
		Funded with GMP savings?
		Incentives Value
		Incentives Distribution: Owner, CM/GC, A/E, DB, Subs, Other
	Risk Allocation	Contingency
		Formal Risk Review Process
		Subcontractors Participation
		Risks Allocated
		OCIP
		CCIP
		Onerous clauses
		Problematic?
		Regulatory constraints
		Problematic?
	Fiscal Transparency	Change Orders
		Bidding and Procurement
		Contingency Usage
		All project Costs
	Team Selection	CM/GC Selection
		CM Competition
		Subcontractors Selection
		Sub Competition

TABLE 9: 3T Delivery Characteristics (Social Tone)

3T	Intermediate Level	Survey Level
SOCIAL: TONE	Past Team Experience: Construction Type	CM/GC
		Sub
		Owner
		A/E
		DB (if app.)
	Past Team Experience: Construction Size	Same categories as above
	Past Team Experience: Delivery System	Same categories as above
	Past Team Experience: BIM	Same categories as above
	Past CM/GC Experience: Stakeholders	Sub
		Owner
		A/E
		Team Past Experience as a Unit
	CM/GC Current Experience with O/A/E	
	CM/GC Current Experience with Subs	
	Project Management Structure: Project Leadership Team	Project Leadership Team
		Number of representatives
		Parties Represented
		Authority
		Jointly develop goals
		Collaborative Decision-Making
		Periodic Reviews
		Preplanning Meetings
		Construction Meetings
		Commissioning Meetings
		Lessons Learned
	Project Management Structure: Cluster Teams	Percent of Project
		Preplanning Meetings
		Construction Meetings
		Commissioning Meetings
	Project Management Structure: Executive Teams	Preplanning, Construction, & Commissioning Meetings
		Conflict Authority
	Stakeholders Involvement	% Design Complete
		Contractor Familiarity
		Owner Participation
		A/E Support
		CM/GC in Design
		Subs in Design
		Co-location

TABLE 10: 3T Delivery Characteristics (Tools and Techniques)

3T	Intermediate Level	Survey Level
FUNCTIONAL: TECHNOLOGY AND TOOLS	Pull Planning	Frequency
		Effectiveness
	Lean Construction Tools	LPS
		SBD
		Weekly Commitments
		VSM
		5S
		Just-In-Time
		TVD
		Visual Mgmt
		Daily Huddles
	Other Tools and Techniques	Prefabrication
		Point Cloud
		Mockups
		Project Training Sessions
		Constructability Reviews
		Safety Trainings / Awareness / Commitment
		Innovation/ Cutting Edge
		Tracking % Complete
		CO Management Process
	BIM	BIM Protocol Manual, Right of Reliance, and Joint Servers
	Use of BIM	Visualization, Space Validation, Site Logistics, Environmental Analysis, Early Design Coordination, MEP Coordination, Design Collaboration, Clash Detection, Submittals, Estimating, 4D Scheduling, Digital Fabrication Construction Simulation, Project Turnover, Facilities Management, Rule/Code Checking
	BIM Systems	Foundation
		Structure
		Interior Finishes
		Exterior Enclosure
		Roofing
		Mech. Syst.
		Elec. Syst.
		Site
		Process Equipment
		Conveying Syst. and Specialties

In addition to the 3T characteristics, general project information also was collected. These range from questions asking for the project name and stakeholders' contact information, to questions regarding the project size, type, complexity and as-built quality. *Tables 11* and *12* display this information.

TABLE 11: General Project Information

High level	Intermediate level	Survey level
GENERAL PROJECT INFORMATION		Project Name
		Location
	STAKEHOLDERS	CM/GC
		CM/GC/DB
		Owner
		Public/Private
		Non-Profit/For Profit
		Owner Contact
		Architect
		Architect Contact
		Mechanical Sub
		Mech. Contact
		Electrical Sub
		Elec. Contact
	General Information	Type
		Program
		GSF Planned
		GSF Final
		Site Size
		Floors
	Labor	Manhours Total
		Manhours CM/GC
		%MH supervisory
		%MH craft
	Construction Type	New
		Addition
		Renovation
	Special Conditions	Special Conditions
		LEED
		Seismic
		Site Access
		Other

TABLE 12: General Project Information (Project Systems)

High level	Intermediate Level	Survey Level
PROJECT SYSTEMS	Complexity	Overall Complexity
		Foundation
		Structure
		Interior Finishes
		Exterior Enclosure
		Roofing
		Mech. Syst.
		Elec. Syst.
		Site
		Process Eqpt
		Conveying Syst.
		Specialties
	Quality as built	Overall Quality
		Foundation
		Structure
		Interior Finishes
		Exterior Enclosure
		Roofing
		Mech. Syst.
		Elec. Syst.
		Site
		Process Eqpt
		Conveying Syst.
		Specialties

The variables presented so far are the 3T delivery characteristics and the general project information. These variables are considered inputs to the project performance function. Next, the performance metrics, or outputs, are discussed.

3.3 Performance Metrics

Some performance metrics were introduced in the literature review. These were compiled and revised by panels of industry experts and researchers, and the resulting list was used to develop questions for the survey. The metrics then were organized in the survey to collect data in a consistent manner over several projects. *Tables 13* and *14* present the performance metrics that were gathered in the survey, spanning the nine areas of cost, quality, schedule, safety, changes, communication, labor, recycling, and business performance. The metrics highlighted in the table are the ones that were calculated based on other data that was obtained. For example, construction unit cost is calculated as the final construction cost divided by the final gross square footage of the facility. The next section demonstrates how the data was used once collected.

TABLE 13: List of Performance Metrics

Performance Area	Performance Metric	Performance Area	Performance Metric
COST	*Unit Cost*	**LABOR**	Self Labor Cost
	Cost Growth		Self Total Cost
	Budget Factor		*Labor Factor*
	Overall Growth		PPC Trend
	Total Cost – Initial (I)		Overtime
	Total Cost – Award (A)		Second Shift Work
	Total Cost – Final (F)		Overmanning
	Design Cost – Initial		*Extra Labor*
	Design Cost – Award		Avg # Workers
	Design Cost – Final		Peak # Workers
	Construction Cost – (I)	**MATERIAL WASTE**	*Waste Ton / million $*
	Construction Cost – (A)		Material Waste (tons)
	Construction Cost – (F)		% Recycled
	Site % Cost		% Sent to Landfills
QUALITY	*Systems Quality*	**BUSINESS**	OH&P
	Deficiency Issues		Return Business
	Deficiencies / million $	**COMMUNICATION**	*RFI per million $*
	Number of Punchlist Items		Number of RFI’s
	Punchlist / million $		RFI Processing Time
	Punchlist % of cost		% Early RFIs
	Warranty Costs		% Field RFIs
	Latent Defects		RFI Work-Arounds
SAFETY	OSHA recordables		*Resubmittals/million $*
	Lost Time Injuries		Number of Re-submittals
	Fatalities		Rework Percentage
	Recordable Incident Rate		Claims
	Lost Time Rate		Cost of Claim
	Recordable per $100 million		
	LTI per $100 million		

TABLE 14: List of Performance Metrics (Continued)

Performance Area	Performance Metric
SCHEDULE	*Construction Speed*
	Delivery Speed
	Construction Schedule Growth
	Delivery Schedule Growth
	Intensity
	Project Advertised Target
	Project Advertised Actual
	Design Start Date Target
	Design Start Date Actual
	Design End Date Target
	Design End Date Actual
	Construction NTP Target
	Construction NTP Actual
	Substantial Comp. Target
	Substantial Comp. Actual
	End of Commissioning Target
	End of Commissioning Actual
	Occupancy Target
	Occupancy Actual
CHANGE	Total Changes / Modifications
	Changes Due to Program Additions and Deletions
	Changes Due to Design Issues and Deficiencies
	Changes Due to Major Regulatory Agencies
	Change Order Processing Time

3.4 Combining Variables

The project information and delivery characteristics, combined with the performance metrics, add up to a total of 304 variables. Some of these original variables were combined into more comprehensive variables to gauge specific aspects of project delivery. A good example to illustrate this concept is the *Project Management Structure* variable, as shown in *Figure 13*. This variable shows how well a project team follows a specific IPD management structure made up of three distinct levels: the core team, executive team, and cluster teams.

Detailed survey-level questions were asked to obtain specific information about the meetings frequency of each of these teams, the number of parties that are involved, the number of representatives, whether they have the authority to make project decisions, whether they develop goals collaboratively, among other questions. All of this information for this distinct management structure is combined into one variable that is used in the analysis.

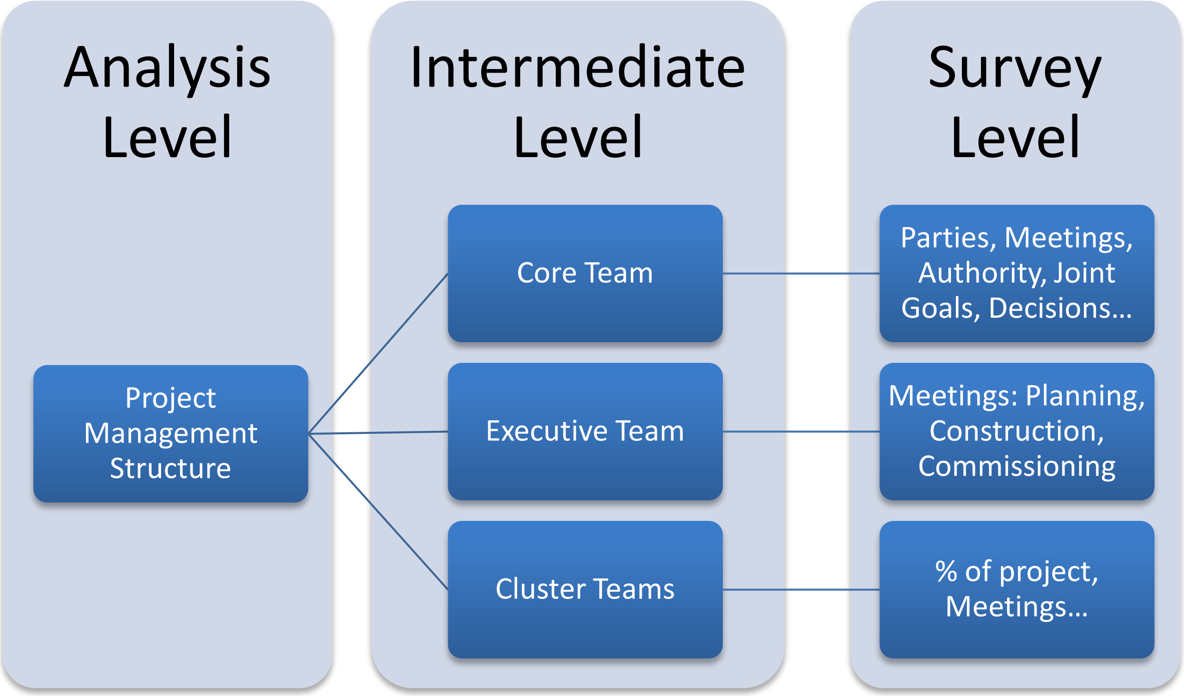


FIGURE 13: Project Management Structure

Another example is fiscal transparency, as shown in *Figure 14*. The four variables (1) fiscal transparency with respect to change orders, (2) fiscal transparency with respect to bidding and procurement, (3) fiscal transparency with respect to the use of contingency, and (4) fiscal transparency with respect to all project costs, were all combined into one variable that measures the degree of fiscal transparency in the project.

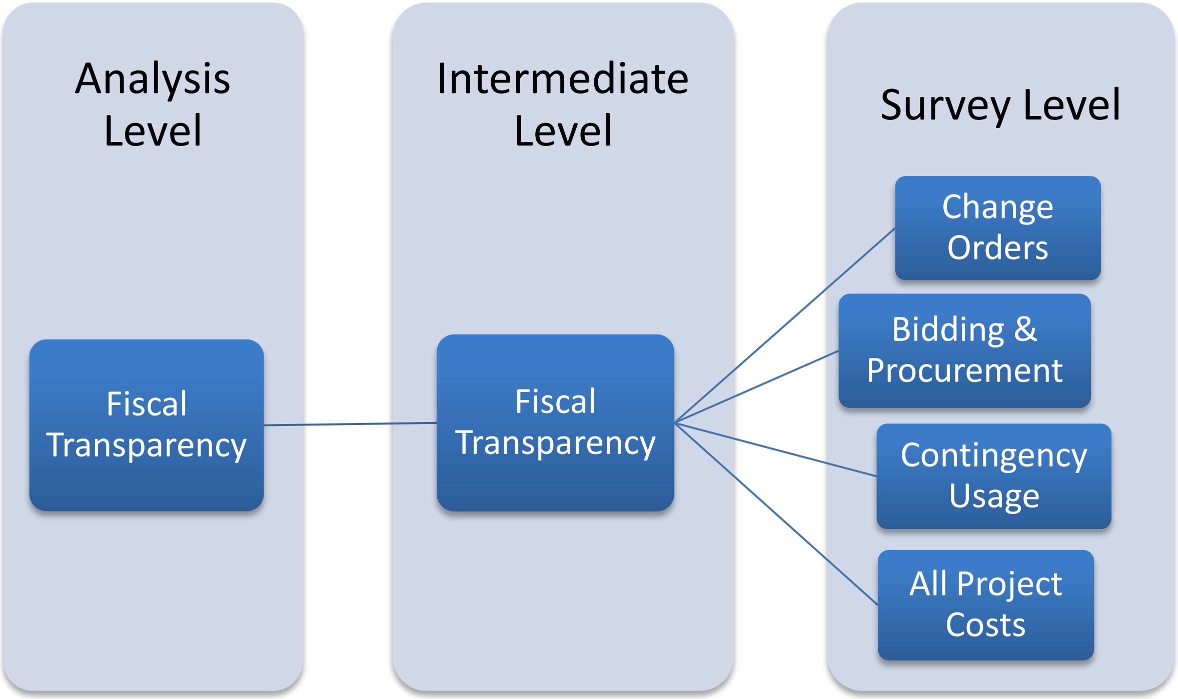


FIGURE 14: Fiscal Transparency

Figure 15 shows a similar summative variable created for the stakeholders' past experiences as a team. The experience of the CM/GC with the subcontractors, the owners, and the designers, are gauged individually, and then the team experience as a unit is measured. All these variables are combined into one comprehensive variable that reflects the stakeholders' past experience as a team.

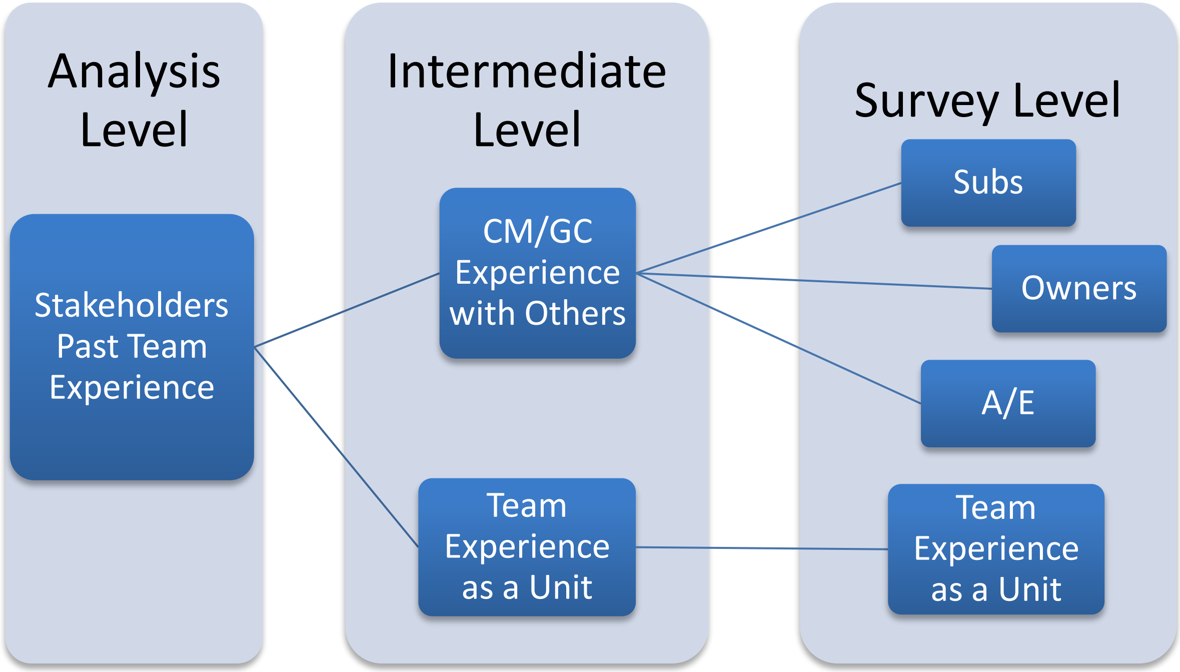


FIGURE 15: Stakeholders Past Experience as a Team

As shown in *Figure 16*, similar summative variables were created for the use of BIM and other tools, techniques, and processes. The figure also shows a variable that represents the team past experiences with the type and size of the project.

Many more delivery variables also were combined when adequate. Other variables were used individually, such as the percent of design complete when the contractor joins the project team, the type of project delivery system, the use of owner-controlled insurance programs, and the use of physical co-location.

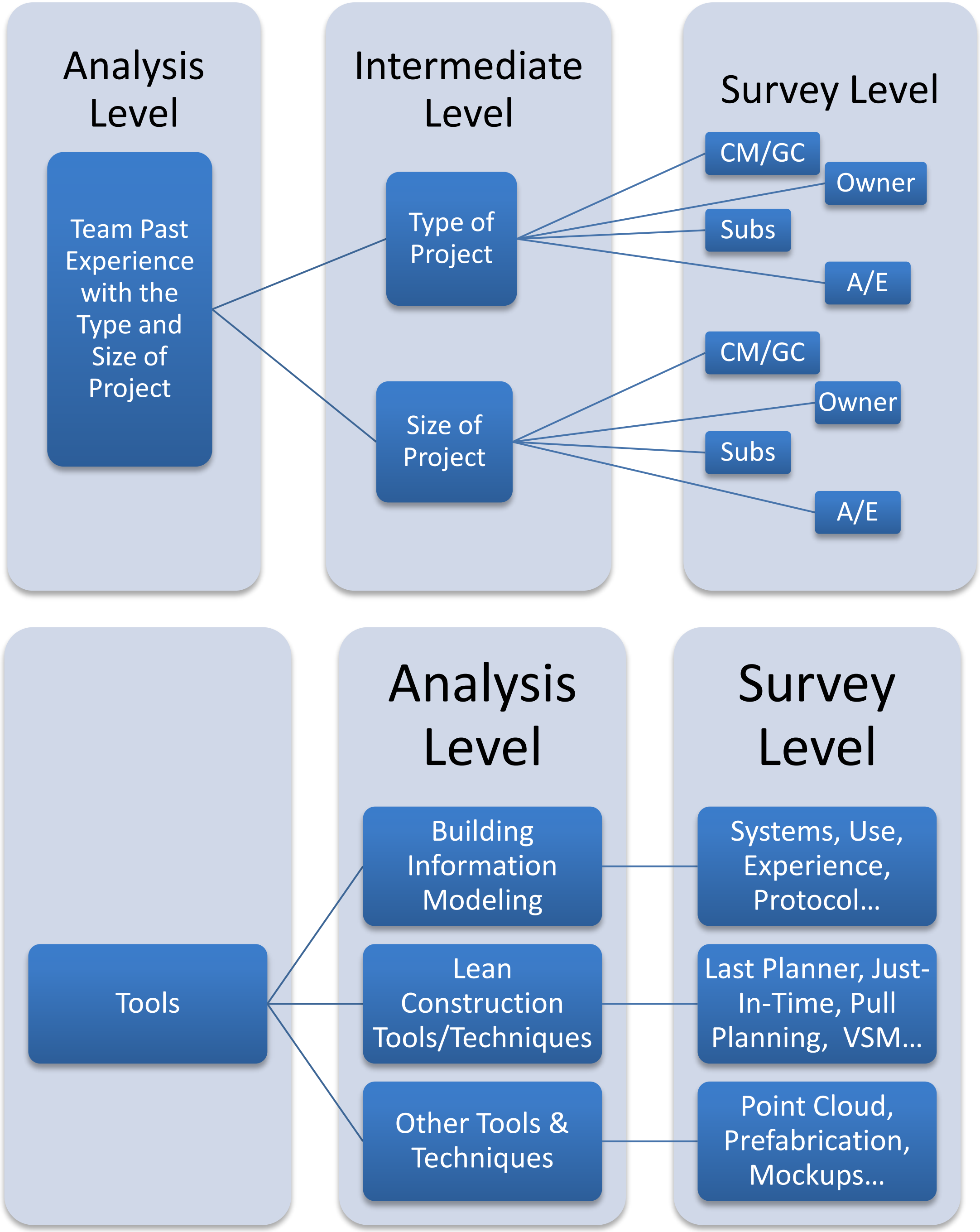


FIGURE 16: Other Summative Variables

3.5 Coding Non-numerical Variables

The types of data collected varied significantly. Some data was binary, such as the use of incentives for which the answer was *yes* or *no* (converted to one or zero), while other fields were discrete, like the number of parties signing the contract. Other data was continuous, i.e. the construction unit cost in dollars per square foot, and there was also categorical data like the type of project delivery system or procurement. Moreover, another type of collected data was ordinal, on a scale of poor to excellent, such as the past experience with the other stakeholders on the team, while interval data also existed, such as the labor overtime (zero to 5%, 6 to 10%, etc.). Finally, some data was based on an actual ratio or percentage format, i.e. the construction cost growth.

Coding of the several variables had to be completed in order to convert them to numerical values that can be statistically analyzed. Following is an explanation of the coding steps for some key variables used in this study.

Both quality and complexity were averaged for all the major building systems, on a scale of 0 to 2 for complexity and a scale of 1 to 5 for quality. The absolute values do not hold a special meaning; however, the relative differences between these values allow for a comparative analysis to be performed. Responses for PDS were categorical (i.e. DB, IPD, etc., with no particular order); the use of a multiparty contract or liability waivers are both binary variables, coded to zero and one; incentives also are binary, although more detailed data is collected in case the findings demanded more investigation. OCIP is binary as well, and the number of stakeholders signing the multiparty contract is discrete so it does not need coding.

Contractors and subcontractors procurement or selection was coded as 0 for open bidding, 1 for prequalified bidding and 2 for negotiated contracts. These values were averaged for CM/GC and subcontractors. Competition, which is binary, was also averaged for CM/GC and subcontractors. Then a comprehensive variable is created for contractors procurement by combining the competition score from the procurement score, then adding one to get a positive range of 0 to 3 : from 0 with all open bidding and competition, to 3 with all negotiated and no competition.

Regarding the compensation variables, there are different potential compensation mechanisms, most commonly used are lump sum, cost plus, guaranteed maximum price, and negotiated compensation. There are different stakeholders that are compensated, most notably the CM/GC, subcontractors, and designers. Each compensation method for each stakeholder is treated as a binary variable, with responses of “yes” and “no” converted to “1” and “0”. Then the variables are summed over key stakeholders so that each compensation method has one value. For example, lump sum compensation for each stakeholder is treated as a binary variable, with responses yes and no converted to one and zero. Then the three variables are summed over key stakeholders to get one value representing lump sum compensation, varying between 0 and 3. When more than one compensation method is used for the same group of stakeholders, the lump sum score for this stakeholder group will be divided by the total number of methods. For instance, if some subcontractors are compensated with lump sum, and others are compensated with cost plus, then the lump sum score for subcontractors will be 0.5 instead of 1.

Past Experience of the key stakeholders is measured across several fields: experience with the type of the facility, size of the facility, experience with the project delivery system, and experience with the other stakeholders and the project team. Information was collected regarding the experience of each group of project participants for each type of experience, on a scale of {0, 1, 3, 9} and then responses across the several stakeholders were averaged to get one value for each type of past experience. This specific scale is used to clearly distinguish high levels of experience. The four different types or prior experience are averaged to get a score for the past experience of the team as a whole on a scale of 0 to 9.

The current experience of the CM/GC with the project team and their chemistry with the other stakeholders is originally rated on a scale of one to five from poor (1), to fair (2), good (3), very good (4), and excellent (5). Two values are averaged to get this score. First their experience with the subcontractor team is gauged, and then their experience with the owner and designer team is gauged.

The project management structure is measured along three different levels: the core leadership team, the executive group, and cluster teams. For projects that had a core group, the number of representatives and the number of parties represented are measured, their meetings frequency is assessed and converted to a scale of 0 to 4, from no meetings to daily meetings. Then the core group’s role is assessed by asking whether they had the authority to make daily project decisions, whether they developed joint goals, made decisions collaboratively, if they performed periodic project reviews and discuss lessons learned. These variables were coded to a scale of 0 to 3, and then added up to the values for meetings frequency and team makeup to create a comprehensive variable for the core group. A value of 0 means there was no core group, and the higher the values the more active and inclusive the core group was on the project. Cluster teams were evaluated based on the percentage of the project they were active, multiplied by their meeting frequency for the planning phase, construction phase, and commissioning phase of the project. Executive teams first had a binary variable for whether they were used or not, again multiplied by their meeting frequency. In case of conflict, the team was asked whether they voted or whether the owner made the decision, and the responses were binary. Then these were added to the scores from the three project management structure levels to form a comprehensive variable for the project management structure.

The percent of design complete when the construction entity joined the team is used as is, and so is the percent of design complete when the GMP was established for GMP project. The involvement of the different stakeholders is measured by combining different aspects of their participation in the project. The familiarity of the CM/GC with the owner’s objectives is gauged, and so is their (and the subcontractors’) degree of involvement in the preplanning and design stage. The participation of the owner’s staff and the support of the A/E during the construction phase also are measured. All these aspects are rated on a scale of 0 to 3 and then added up in a variable measuring the degree of involvement of the different project parties. Also included in this variable is the involvement of the subcontractors in the project risk review process.

Fiscal transparency is measured separately for change orders, bidding and procurement, contingency usage, and all project costs. For each area, the respondents can answer *None* (0), *A little* (1), *Some* (3), or *A lot* (9). Scores for these four aspects are averaged to get a comprehensive fiscal transparency variable. Another characteristic, physical co-location of the key project stakeholders, is also rated on a scale of 0 to 9.

The scores for BIM were scored differently that the remaining characteristics because of the abundant BIM data collected. There were two main sets of variables: extent of BIM use and potential of effectiveness.

1. BIM use includes modeled systems and BIM functions. The building systems that were modeled using BIM (as shown on *Page 7* of the survey in *Appendix C*) are determined individually on a scale of 0 to 3, and then averaged across all systems. The functions for which BIM was used (as shown on *Page 12* of the survey in *Appendix C*) are rated on a scale of 0 to 9, and then averaged to obtain a variable that reflects the extent of BIM functions in the project. These two BIM use variables, BIM systems and BIM functions, are multiplied to gauge the extent to which BIM was used in the project.

2. The potential for BIM effectiveness includes the BIM experience of stakeholders and the BIM infrastructure for the project. The experience of the different stakeholders with BIM (as shown on *Page 8* of the survey in *Appendix C*) is gauged on a scale of 0 to 9 and then averaged across all stakeholders to get one score for BIM experience. The BIM infrastructure questions are binary and include the existence of a BIM protocol manual, the contractual right of reliance on 3D models, and the use of joint servers,

which are all added together. The resulting value is averaged with the score for stakeholders' BIM experience to result in the potential for effectiveness of BIM use; i.e. the more the team is experienced with BIM and has the infrastructure for it, the more the use of BIM is likely to result in better performance.

Finally, (1) and (2) are multiplied to compute a single variable that represents BIM characteristics in a given project. This method is similar to methods used to account for risk items, by multiplying their probability and impact scores.

The use of lean construction tools and processes is also measured. Each tool or process is scored on a scale of 0 for "no use" to 9 for "used a lot". The list of lean tools measured for this study includes the Last Planner System (LPS), pull planning and scheduling (PP), just-in-time delivery of materials (JIT), 5S, set-based design, value stream mapping, target costing or target value design, daily huddles, and the use of visual management devices. LPS, PP, and JIT were analyzed separately as opposed to being combined with the remaining tools and techniques because they are more predominant. More than one question was related to the use of each of these techniques. For pull planning, both the frequency and effectiveness of PP were gauged and their scores multiplied. For LPS, the respondents were asked to what extent they used LPS, whether they tracked weekly commitments from the project teams, and whether they tracked reliable promises or percent plan complete. For JIT, respondents were asked to rate the extent to which JIT was used, and then to provide a JIT definition based on their understading of it. The definitions varied from the materials being taken off the truck and installed on the building right away, to a minor storage of small batches for a short period of time, to a site warehouse for long batches stored for a long period of time. These definitions were scored 9, 3, and 1, respectively, and then multiplied by the scores reflecting the extent to which JIT was used.

The scores for the use of innovative tools and techniques averaged scores of six variables, all measured on a scale of 0 to 9 as follows: *None* (0), *A little* (1), *Some* (3), or *A lot* (9). The only exception was prefabrication, where the number of trades that used prefabrication was counted, and happened to also range from 0 to 9. The remaining variables are the use of a point cloud technology, physical mockups, project training sessions, constructability reviews, and safety trainings.

Some performance metrics were also coded. For instance, the use of extra labor is gauged through assessing the amount of overtime labor, second shift work, and overmanning in terms of ranges of percentage values.

3.6 Dataset Summary

As discussed earlier, collecting IPD data was no easy task. First, given that IPD is a very recent delivery system, the total number of IPD projects available for data collection is scarce. Second, since IPD is considered an advantageous process by the companies using it, convincing them to release IPD data to be analyzed and shared also would be difficult. However, generous industry collaborators granted access to 35 projects: 12 IPD projects and 23 comparable non-IPD projects. *Figure 17* shows the makeup of delivery systems for the 35 projects. As stated earlier, two-thirds of the IPD projects (as identified by their project managers) had a multiparty contract signed by three or more parties. The upside from having strong industry commitment is that it allowed a very thorough data collection effort to take place, gathering information about 304 variables for each project. The data collection effort took up to two days for a single project.

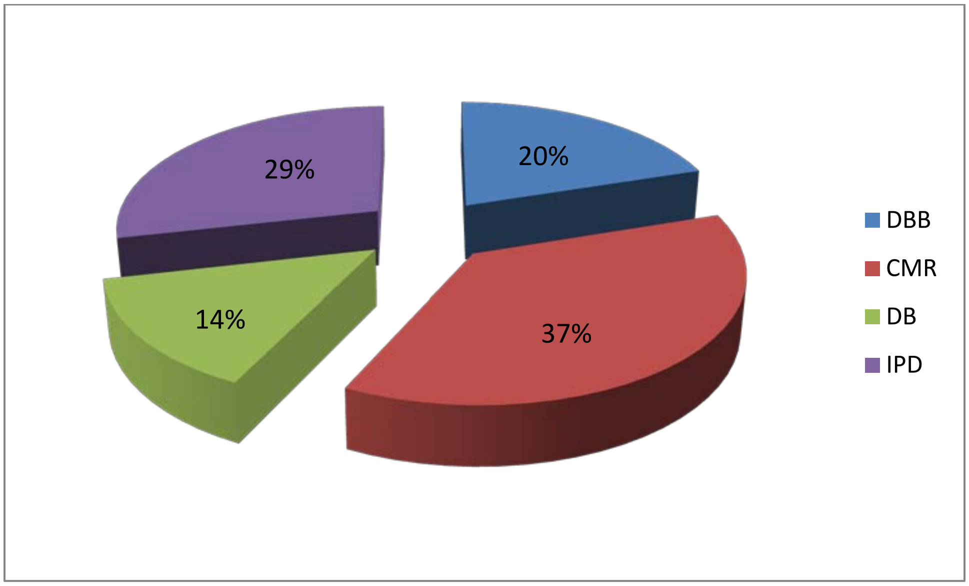


FIGURE 17: Projects Makeup among Major PDS

The projects used for this research were predominantly in two geographic locations, as shown in *Figure 18*. The first is the U.S. Midwest region, and the second is the state of California. Most IPD work is being conducted in these two geographic locations, mostly because some companies that are leading IPD efforts are involved with projects in these two parts of the country. Data was collected in these areas to provide fair comparisons between the IPD and non-IPD projects. Some projects analyzed were also located in the states of Colorado and Massachusetts.

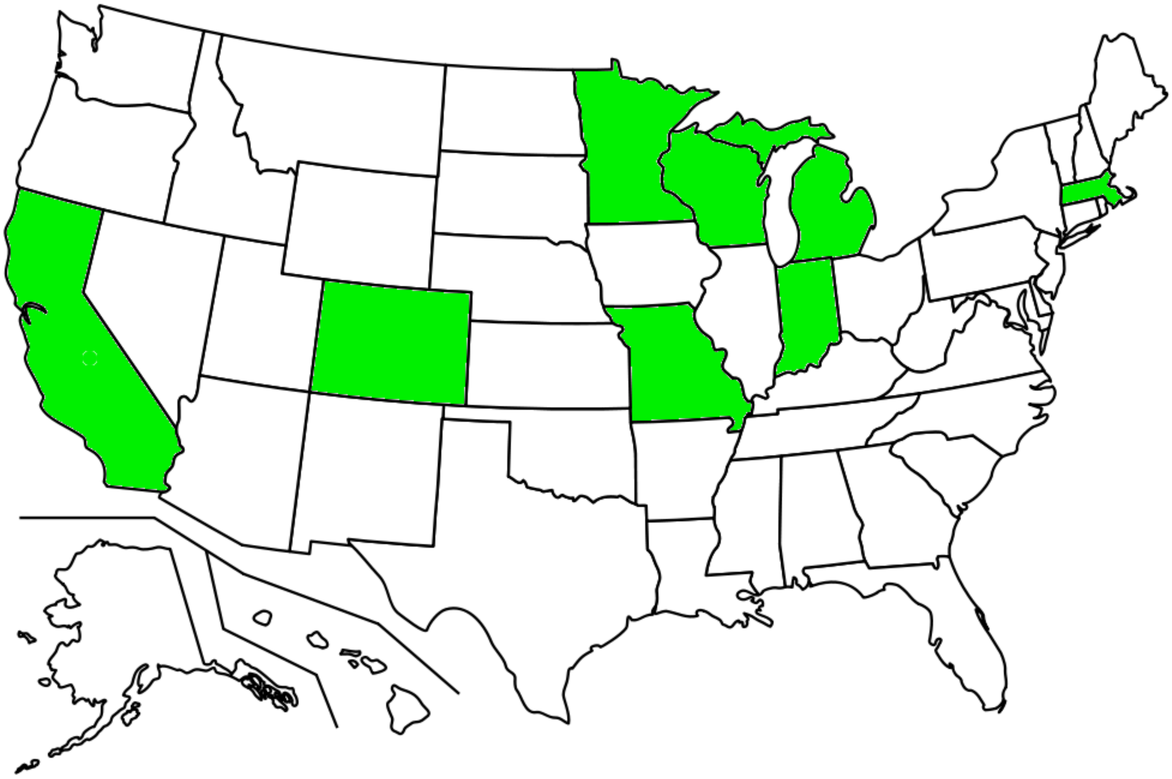


FIGURE 18: U.S. Map of Respondents

The types of projects, as discussed in the *Research Scope* section, were generally complex institutional vertical construction facilitites, with a few commercial facilities. In fact, about 50% of the projects in the database were healthcare facilities and about 25% were university research laboratories. The total dollar amount of construction work for all projects combined was close to $3 billion. The cost distribution, shown in *Figure 19*, included project costs from $5 million to $400 million.

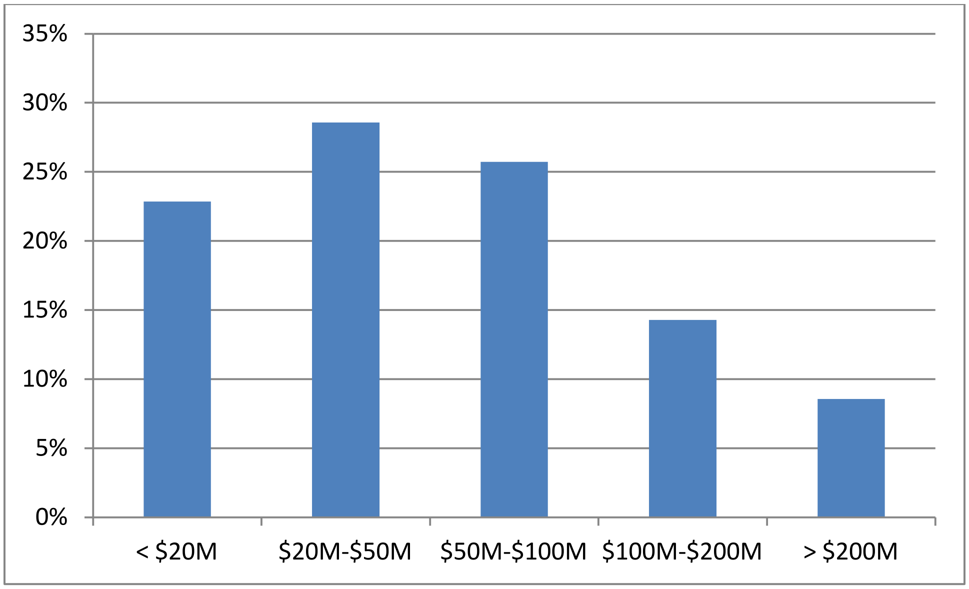


FIGURE 19: Projects Distribution in Construction Cost

Next, *Figure 20* shows the projects size distribution in terms of final gross square footage. All projects except one were recently completed, between the years 2005 and 2012.

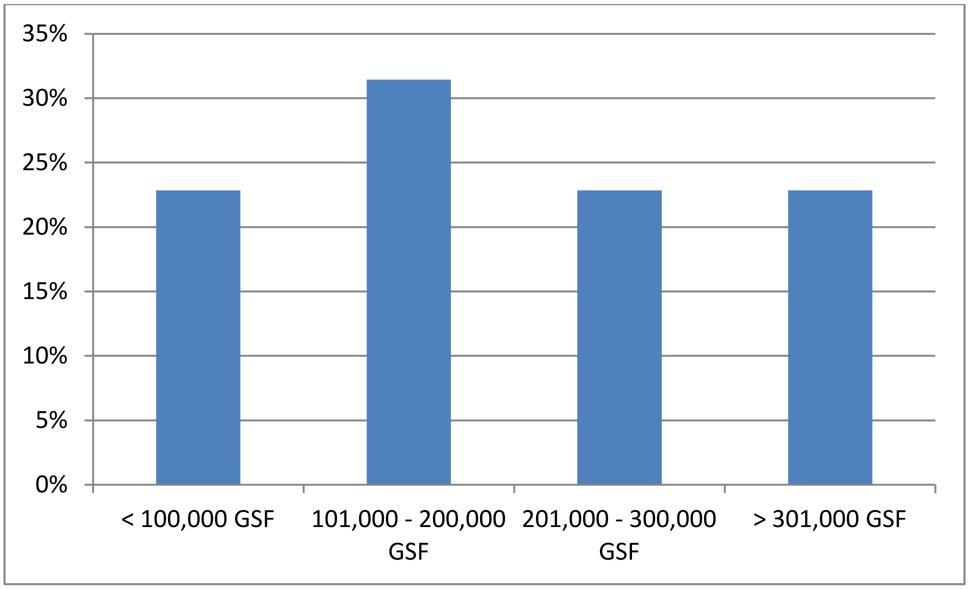


FIGURE 20: Projects Distribution in Final GSF

The distribution of projects by number of floors is shown in *Figure 21*, ranging from a single story all the way to 13 stories. Then *Figure 22* displays the distribution of projects in number of total construction labor hours (including subcontractors).

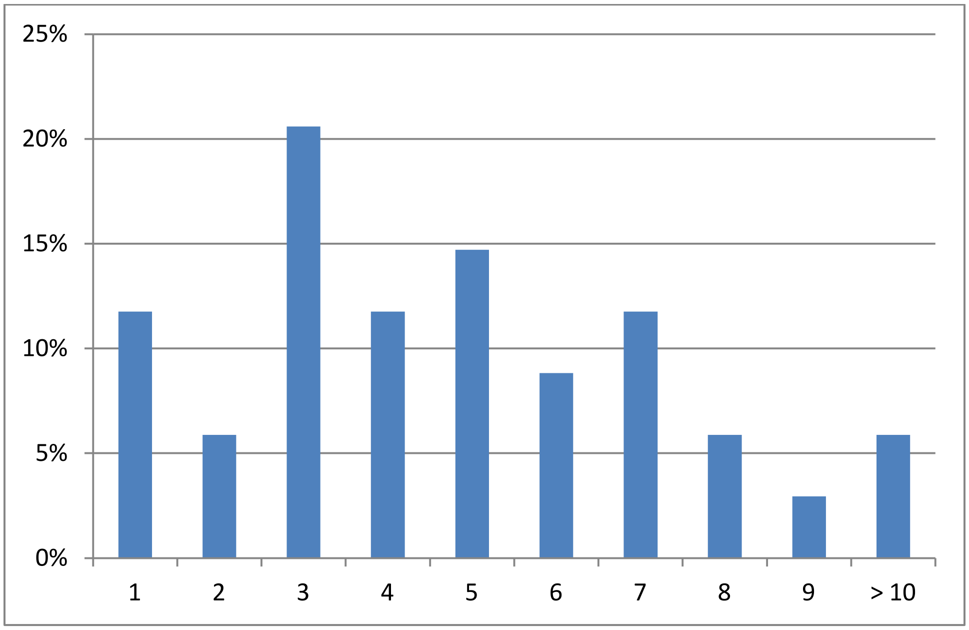


FIGURE 21: Projects Distribution in Number of Floors

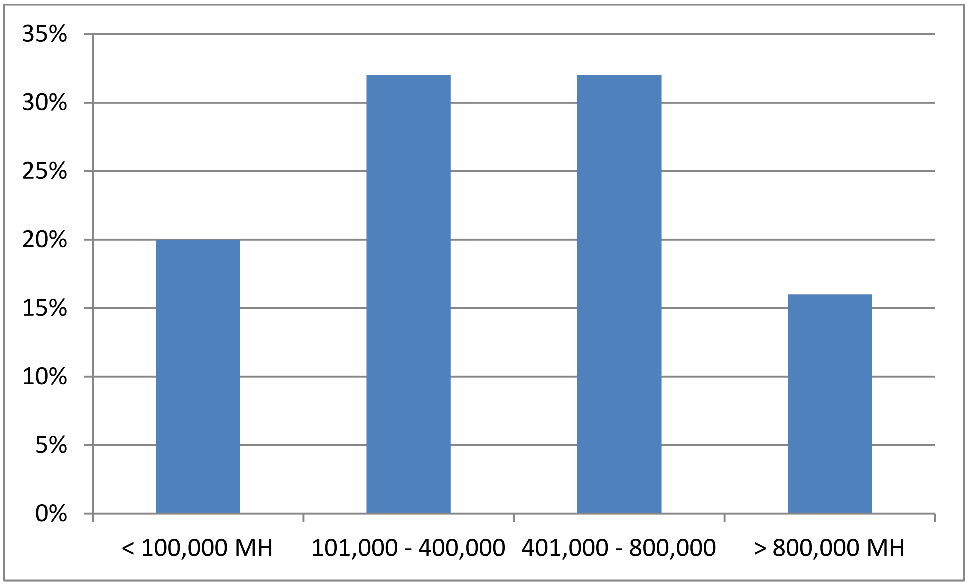


FIGURE 22: Projects Distribution in Number of Construction Labor Hours

Over all projects, approximately 70% of the total work consisted of new construction, with 15% of the work for additions and another 15% renovation work, as show in the pie chart in *Figure 23* below.

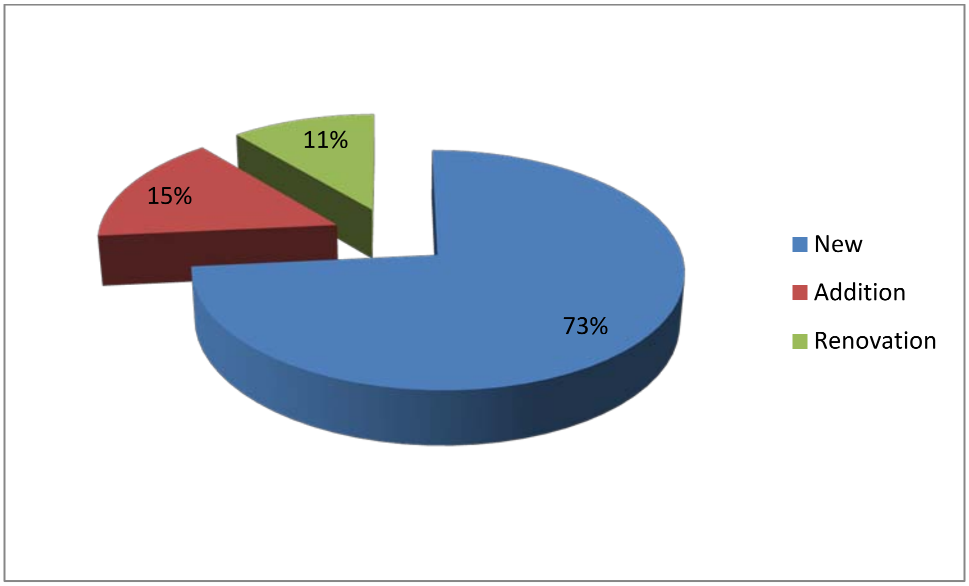


FIGURE 23: Projects Makeup by Type of Construction

About 40% of the projects obtained a Leadership in Energy and Environmental Design (LEED) rating, as shown in *Figure 24*. However, the largest portion of the projects in the database did not pursue a LEED rating. These numbers are shown here; however, it is arguable whether a facility that obtained a LEED rating will achieve higher performance than a comparable non-LEED facility. In fact, studies show that there are no significant differences between the energy consumption of LEED and non-LEED buildings (Menassa et al. 2011).

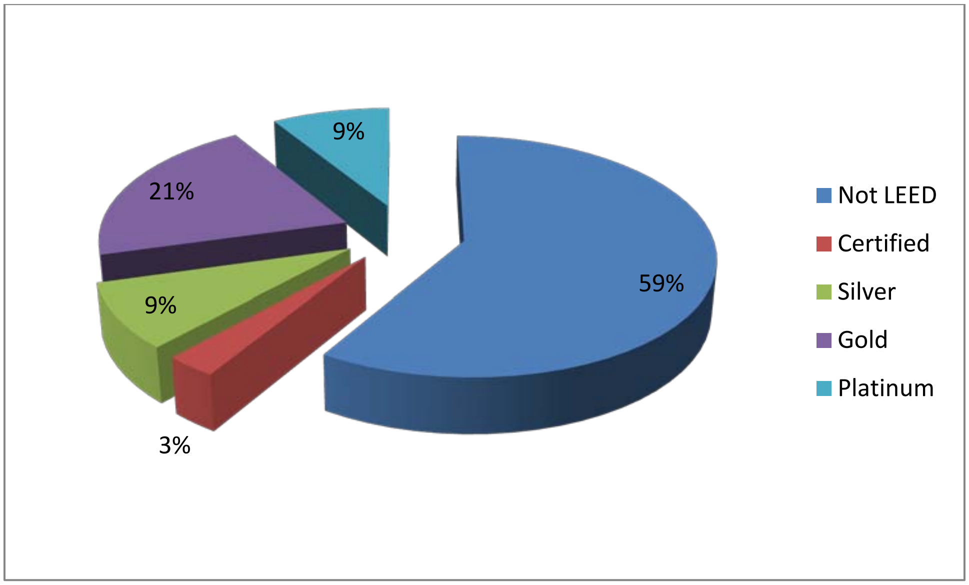


FIGURE 24: LEED vs. Non-LEED Projects Distribution

When asked about procurement and contractor selection, the responses for construction manager or general contractor (CM/GC) selection were different than the responses for subcontractor selection. *Figure 25* highlights these differences. Competitive open bidding was not common when selecting a CM/GC in this database of projects (only 6%); rather prequalified bidding was used in 26% of the cases, and the remaining majority of 68% was negotiated work. The results are somewhat different for the selection of subcontractors: about half of the subcontractors were selected based on prequalified bidding, while the remaining were split halfway between open bidding and negotiated procurement.

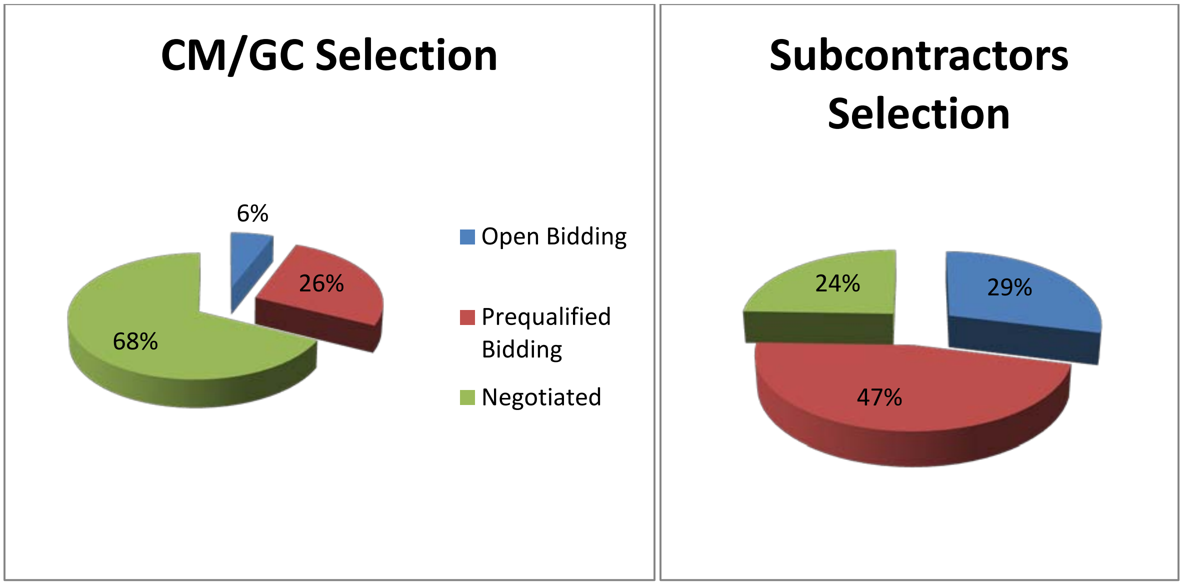


FIGURE 25: Contractor Procurement

What is interesting here is that in the entire IPD projects sample, all but one CM/GC were selected in a negotiated environment. However, this only applied to about a third of subcontractors' selection on IPD projects. It is remarkable that the IPD philosophy of negotiating with your project partners is being applied at the level of owners and general contractors, but not at the level of subcontractors. This new philosophy did not fully trickle down the whole supply chain yet, and it might just be a matter of time until all project stakeholders get used to this new environment.

The compensation type of the CM/GC for the projects in the database is mostly based on a Guaranteed Maximum Price (GMP), while the compensation for subcontractors was mostly based on lump sum. It is interesting that no IPD projects used a lump sum compensation for their CM/GC, and mostly relied on Cost Plus and in some instances (about 30%) GMP compensation.

Incentive clauses are included in some contracts to incentivize collaboration by sharing risk and reward. About a third of the total number of projects used incentives, some of which were not IPD projects. This number doubles to about two thirds for IPD projects using incentive clauses. The typical incentive clauses used include are sharing the profits and the savings, as well as the unused contingencies. The incentives were typically split between the owner, CM/GC, A/E and often subcontractors. The amount of shared incentives varied from half a million dollars up to $8 million.

Chapter 3 introduced the delivery characteristics and performance metrics that are studied, and discussed the characteristics of the data collected for this research. The following three chapters analyze the collected data to evaluate the performance of IPD projects. Next, *Chapter 4* discusses univariate analyses where the performance of IPD projects is compared to that of non-IPD projects based on each individual performance metric.

Chapter 4. Improvements of IPD on Key Performance Metrics: A Univariate Analysis

The previous chapter introduced several delivery characteristics and performance metrics used for this study. This chapter focuses solely on the project delivery system, and investigates its effect on all identified performance metrics for which data was available. IPD and non-IPD projects are compared for each performance metric individually using a univariate analysis, which allows for a clear comparison of IPD and non-IPD project performance. A glossary of terms, abbreviations and definitions can be found in *Appendix B*.

4.1 Statistical Methods

Two statistical tests were used for each performance metric to provide a comprehensive look at the comparisons between IPD and non-IPD projects. The two types of analyses used for this part of the study are t-tests and the non-parametric Mann-Whitney-Wilcoxon (MWW) tests. The t-test is an analysis that can be used to assess the statistical significance of the difference between two sample means. In general terms, a t-test is optimal when each population in the dataset is normally distributed. MWW is a non-parametric statistical hypothesis test used when the data cannot be assumed to be normally distributed. In fact, when the data is normally distributed, the MWW test has 86% of the power of the t-test; however, when the data is not normally distributed, the MWW test has a much larger power, at times up to infinity (Lehmann 2006). It is more conservative to interpret the results of the MWW test because the normality assumption is not needed, and therefore it is less likely to draw the wrong conclusions.

Data for certain variables can be assumed to be normally distributed, whereas this assumption does not hold for other variables. *Quantile-to-Quantlile plots*, or Q-Q plots, were used to determine whether the normality assumption was met for t-tests, although this assumption is not necessarily required. These Q-Q plots will be discussed in greater detail in *Section 4.3*. Both t-tests and MWW tests were used throughout for each performance metric to provide a more comprehensive look at the comparisons, and both sets of results are presented in this chapter.

In addition to comparing each metric individually, the metrics belonging to each general performance area will be aggregated in one comprehensive metric which also will be tested. For example, all the different metrics related to quality (systems quality, deficiency issues, warranty costs and latent defects, etc.) will be combined into one comprehensive metric representing quality, which will be compared for IPD and non-IPD projects. The statistical technique used for this purpose is Principal Component Analysis (PCA), a dimension reduction technique for quantitative data. PCA aims to reduce the number of dimensions in a dataset, while keeping as much information as to the inherent variability within the data. In fact, PCA linearly combines the original variables into new variables that are uncorrelated with each other, such that a few of these new variables will explain most of the variation in the original data. PCA is a completely general method that does not assume any model for the data, such as multivariate normality (Cox 2005). A more detailed explanation of PCA can be found in Jolliffe (2002). PCA and its role to combine several related metrics for the purpose of this study will be discussed further in the next section.

A discussion of univariate results for individual performance areas follows. After conducting the analyses, most of the results were identical whether t-tests or MWW tests were used. Therefore, *Section 4.2* presents the MWW results, then *Section 4.3* complements these results with the t-test results for a clear comparison.

4.2 Metric by Metric Univariate Results

This section looks at each performance area individually. The following nine subsections of *Section 4.2* are split by performance areas that cover related metrics. For instance, the first area consists of metrics related to cost performance, including construction unit cost and cost growth.

4.2.1 Cost Performance Metrics

Data for two standard cost performance metrics was available for most of the projects: (1) unit cost and (2) construction cost growth. *Unit cost* is measured in dollars per square foot. *Construction cost growth* is measured in percentage terms by comparing the final construction costs to the original estimated construction costs. Finally, a Principal Component Analysis (PCA) was conducted to combine these two metrics into one comprehensive cost performance metric, which will be compared for IPD and non-IPD projects.

Project costs were adjusted to account for location and time. RS Means City Cost Indexes were used to adjust costs based on location. To reduce effect of location and provide a fair comparison of construction costs across the U.S., unit costs were divided by the index. Additionally, the Engineering News Record (ENR) Construction Cost Indexes, available from 1908 to date, were used to adjust the unit costs based on the time of each project. For consistency, the dates for substantial completion (month and year) were used for all projects. ENR builds its indexes by combining average common labor rates and the prices of structural steel, portland cement, and lumber for each period.

IPD and non-IPD projects are compared. Before discussing the results of the statistical analysis, boxplots of the data are presented. A boxplot is a non-parametric graphical summary of data, displaying the sample minimum, lower quartile, median, upper quartile, and maximum. The median value is represented by a thick black line, dividing the dataset in half, and the box represents the 50% of the data around the median, while the remaining 50% of the data are divided equally above and below the box. Boxplots give a visual representation of the dataset and provide insights as to the distribution of the data.

Figure 26 includes two boxplots depicting cost performance. The boxplots on the left side of *Figure 26* show the construction unit cost data in dollars per square foot. The horizontal axis separates the IPD projects in green and non-IPD projects in red, while the *"IPD-ish"* projects shown in the middle are projects identified by the respondents as IPD because they used many IPD delivery characteristics. However, by not using a multiparty contract, these projects in yellow did not fit the IPD definition used for this study. Therefore they were not considered as part of the two groups of interest for this study, and plotted in yellow between IPD and non-IPD projects.

The vertical axis corresponds to construction unit cost, and the boxplots show that IPD projects seem to have a median unit cost slightly higher than the non-IPD projects, as represented by the thick horizontal line around the middle of each boxplot. These findings are only based on plots of the raw cost data, and any visual differences need to be tested for statistical significance. However, the discussion of project costs is incomplete without considering project quality, as will be discussed in the next subsection.

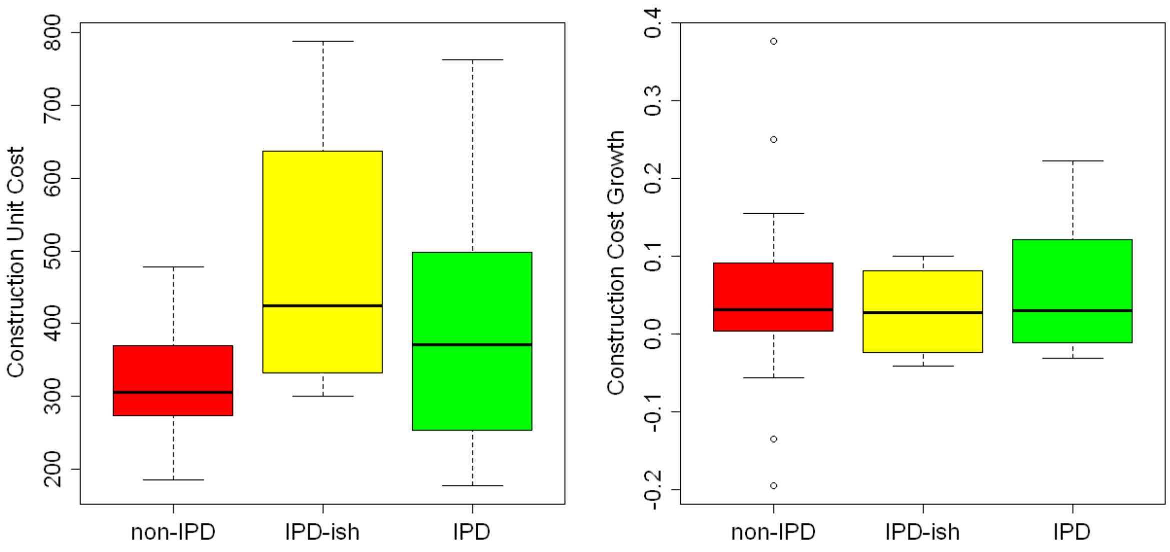


FIGURE 26: Boxplots for Cost Metrics

The boxplots on the right side of *Figure 26* represent the data for construction cost growth in percentage of the initial cost estimates. One cannot see major differences in medians here.

Next, statistical tests are used to determine whether there are any statistically significant differences in cost performance between IPD and non-IPD projects. To compare construction unit cost and construction cost growth for the IPD projects sample and the non-IPD sample, both two-sided and one-sided MWW tests were conducted.

For each individual test, the smaller the p-value, the more significant the performance differences are between IPD and non-IPD projects. The commonly used threshold is 0.05 below which the performance differences between the two samples are considered statistically significant. The null hypothesis used for these tests is that the performance metric (i.e. unit cost, cost growth, etc.) is equal for both the IPD and the non-IPD samples. The alternative hypothesis for two-sided tests is that the IPD sample performs differently than the non-IPD sample. The alternative hypothesis for one-sided tests is that the IPD sample performs better than the non-IPD sample. For example, in the case of unit cost, superior performance means a lower cost. When the tests do not reject the null hypothesis, one can assume there are no significant differences in performance between IPD and non-IPD projects. When the tests reject the null hypothesis, the alternative hypothesis is true and IPD has a dissimilar or superior performance based on the specific performance metric. In the four tests conducted here, all the p-values were higher than 0.05, denoting no significant differences in cost performance, as shown in the first two rows of *Table 15*.

To confirm that cost performance differences between IPD and non-IPD projects were not statistically significant, the last analysis uses a comprehensive metric that combines the two cost metrics by using PCA. As discussed earlier, PCA linearly combines the original variables into new variables that are uncorrelated with each other, such that a few of these new variables will explain most of the variation in the data. These new variables are called principal components and will be compared for IPD and non-IPD projects.

The first principal component of the cost performance metrics ($PC1_c$) combines the two original cost metrics (construction unit cost and cost growth) and explains 56% of their total variance. This might not be the best example of PCA since there are only two metrics; the reader will appreciate this method even more in the coming subsections. $PC1_c$ can be used as a new variable that summarizes cost performance in one metric, expressed as:

- $PC1_c = 0.707*UnitCost + 0.707*CostGrowth$

The component loadings or coefficients of each variable (here 0.707) ensure that $PC1_c$ is in the direction that maximizes the portion of variance explained by the component. The cost data is standardized before using PCA, which results in the average project scoring zero, while half the dataset obtains positive scores and the other half negative scores. $PC1_c$ scores for the projects varied from the best cost score of -2.02 to the worst score of 2.36. When $PC1_c$ was compared for IPD and non-IPD projects, the p-value was 0.216, which denotes no significant differences between the two samples. This last result concludes the cost performance analysis for IPD projects, and a summary of the tests discussed are presented in *Table 15*. This subsection confirms the findings of previous literature (Cho and Ballard 2011) that found no significant differences in cost performance for IPD projects. The following subsection complements the cost discussion by evaluating IPD project quality.

TABLE 15: Hypothesis Tests for Cost Performance

Hypothesis Test Number	Hypothesis	p-value	Outcome at 95% level
1	IPD projects result in a different construction unit cost than non-IPD projects	0.659	Fail
2	IPD projects experience a different cost growth than non-IPD projects	0.941	Fail
3	IPD projects see an overall superior cost performance over non-IPD projects ($PC1_c$)	0.216	Fail

4.2.2 Quality Performance Metrics

The previous subsection showed there are no statistically significant differences in cost performance between IPD and non-IPD projects. However, as mentioned earlier, the cost discussion is incomplete without considering project quality in order to conduct a fair comparison. Since quality is hard to measure, both qualitative and quantitative performance metrics were evaluated to provide a comprehensive understanding of quality performance. The quality performance metrics include (1) the quality of major building systems, (2) the number of deficiency issues, (3) the number and cost of punchlist items, and (4) the costs of warranty and latent defects.

Major building systems include structural, mechanical, and finishes. Respondents were asked to provide the quality of each system on a scale of 1 to 5, representing Economy, Standard, High Quality, Premium, or High Efficiency Premium. Deficiency issues are issues that arise during the course of construction, and can be related to numerous reasons, such as failed field inspections and jurisdiction problems related to code observance. Punchlist items are the uncompleted or unsatisfactory items remaining after the substantial completion of a project, such as components needing minor repairs or replacement. Warranty costs are measured in the first year of occupation. Latent defect costs are measured after the end of the one-year warranty period.

All of the above items can serve as indicators of the building quality. In order for these items to be compared across a number of projects of different sizes, their values have been normalized. For example, the number of deficiency issues per million dollars was obtained by dividing the total number of deficiency issues for a project by the final construction cost of the project. The number of punchlist items per million dollars was calculated in a similar manner. The relative costs of punchlist items, the costs of warranty, and the costs of latent defects, were all obtained in percentage of the total construction cost.

The upper left corner of *Figure 27* shows the boxplots for overall project quality combining all major building systems. One can see a clear superiority in quality performance for the IPD projects in green when compared to the non-IPD projects in red, while the quality scores for IPD-ish projects were in between. The upper right corner shows the boxplots for the number of deficiency issues per million dollars. Even before performing any statistical analyses, one can see that IPD projects experience considerably less deficiency issues than their non-IPD counterparts. Additionally, IPD projects in this sample have considerably less punchlist items than non-IPD projects, as shown in the lower left corner of the figure. Finally, the interpretation of the warranty costs and latent defects is not very obvious and will need statistical testing. One can see that 25% of the non-IPD sample scored “one”, as illustrated by the missing portion of the box below the median. As discussed earlier, these findings are only based on plots of the raw cost data, and any visual differences need to be tested for statistical significance.

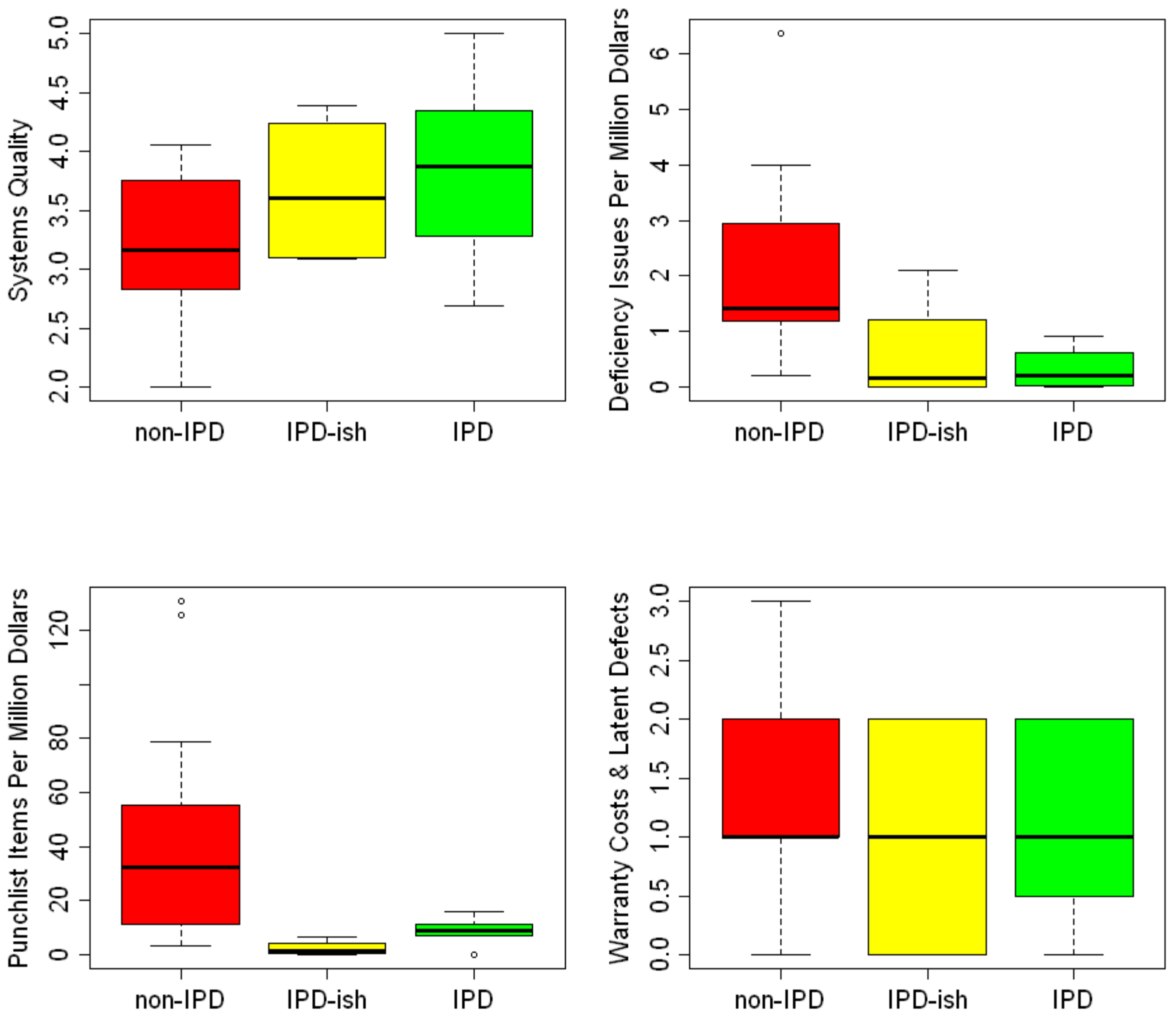


FIGURE 27: Boxplots for Quality Metrics

Both two-sided and one-sided MWW tests were conducted to statistically verify the significance of the differences observed when comparing quality metrics between the IPD projects sample and the non-IPD sample. Most tests showed significant differences; the one-sided test for systems quality showed a p-value of 0.032 indicating IPD projects have evident superior quality over non-IPD projects. This result is statistically significant at the 0.05 level and proves that IPD projects have a higher systems quality than their non-IPD counterparts.

Similarly, both two-sided and one-sided MWW tests were conducted for deficiency issues, and both tests showed significant differences between IPD and non-IPD projects. The one-sided test resulted in a p-value of 0.001 indicating IPD projects have significantly less deficiency issues than non-IPD projects. This result is significant at the 0.05 level and the more conservative 0.01 level. In fact, the median value for non-IPD projects is 1.4 deficiency issues per million dollars versus 0.2 deficiency issues for the IPD projects. The point estimate for the difference is 1.4 issues per million dollars with a 95% confidence interval ranging between 0.5 and 3.0 issues.

Additionally, both two-sided and one-sided MWW tests were conducted for the two metrics measuring punchlist items: (1) the number of punchlist items per million dollars and (2) the cost of punchlist items in percentage of total construction cost. All four tests show significant differences between IPD and non-IPD projects. The one-sided test for number of items per million dollars shows a p-value of 0.013 indicating IPD projects have significantly less punchlist items than non-IPD projects. This result is significant at the 0.05 level, and the median value for non-IPD projects is 32.39 items per million dollars, versus 8.98 for the IPD projects. The point estimate for the difference is 23.05 items per million dollars, with a 95% confidence interval for the difference ranging between 2.82 and 48.18 items. The width of this confidence interval is a function of the sample size and the variance of the data.

While the tests conducted for the latent defects variable do not show significant differences between IPD and non-IPD projects, the individual tests for warranty costs show differences in performance. The one-sided test shows a p-value of 0.040 indicating IPD projects have lower warranty costs than non-IPD projects, and this result is significant at the 0.05 level. A point estimate for warranty costs is not provided here because the data collected for this metric was ordinal, where a value of zero indicates no warranty costs, a value of one denotes warranty costs of zero to 0.5% (relative to project cost), and a value of two denotes 0.6% to 1% of project cost.

To confirm that the differences in quality performance metrics between IPD and non-IPD projects are statistically significant, the last analysis used a comprehensive metric that combined all quality metrics using PCA. The first two principal components explain about 73% of the variance in the original quality metrics. On one hand $PC1_q$ highlights systems quality which needs to be maximized, and on the second hand it underlines other areas of quality performance that need to be minimized, such as deficiencies, punchlist items, warranty costs and latent defects. $PC2_q$ highlights warranty costs and latent defects, which are long-term quality problems, versus deficiency issues which are dealt with during the course of construction. $PC1_q$ and $PC2_q$ are presented here:

- $PC1_q$ = 0.613*SystemsQuality - 0.666*PunchlistItems - 0.356*Deficiencies - 0.232*WarrantyAndLatent
- $PC2_q$ = - 0.639* Deficiencies + 0.766* WarrantyAndLatent

There were several missing values for the deficiency issues variable, and therefore an additional principal component analysis is completed without this variable. As shown below, the new principal components can be interpreted similarly to the previous ones, with $PC1_q$' focusing on systems quality and punchlist items, and $PC2_q$' highlighting the long-term problems of warranty costs and latent defects. Both sets of principal components are tested.

- $PC1_q$'= 0.630*SystemsQuality - 0.653*PunchlistItems
  - 0.420*WarrantyAndLatent

- $PC2_q$'= 0.371*SystemsQuality - 0.221* PunchlistItems
  + 0.902* WarrantyAndLatent

When the principal components were compared for IPD and non-IPD projects, the first component denoting overall quality performance showed significant differences. $PC2_q$ which represents long-term quality performance did not show significant differences at the 0.05 level. However, the p-value of 0.076 is considered significant at the less stringent 0.10 level. Additionally, the test results of $PC1_q$' and $PC2_q$'were very similar to those of $PC1_q$ and $PC2_q$; therefore only the latter are included in *Table 16*. These results confirm the previous findings, which tested individual performance metrics and showed significant quality differences for all quality metrics except latent defects. This last result concludes the quality performance analysis for IPD projects, and a summary of the tests discussed are presented in *Table 16*. The sample size used for each test can be found in *Appendix E*, along with more information regarding the test results.

This subsection provides the first quantitative proof that the IPD delivery system has superior performance as compared to traditional delivery systems. Although higher quality systems typically have added complexity, IPD projects do not result in more long-term issues with these systems. In fact, depending on the significance level wanted, IPD projects could even decrease the long-term quality issues with building systems.

TABLE 16: Hypothesis Tests for Quality Performance

Hypothesis Test Number	Hypothesis	p-value	Outcome at 95% level	Outcome at 90% level
1	IPD projects result in a higher systems quality than non-IPD projects	0.032	Pass	Pass
2	IPD projects result in less deficiency issues than non-IPD projects	0.001	Pass	Pass
3a	IPD projects result in less punchlist items than non-IPD projects	0.013	Pass	Pass
3b	IPD projects result in less punchlist percentage costs than non-IPD projects	0.003	Pass	Pass
4	IPD projects result in lower warranty costs than non-IPD projects	0.040	Pass	Pass
5	IPD projects result in lower latent defects costs than non-IPD projects	0.442	Fail	Fail
6	IPD projects result in higher overall project quality than non-IPD projects (PC1q)	0.021	Pass	Pass
7	IPD projects result in higher long-term project quality than non-IPD projects (PC2q)	0.076	Fail	Pass

Combined with the previous subsection on cost performance, these results provide a better understanding of IPD project performance by demonstrating that IPD delivery systems result in higher quality projects at no significant cost premiums. The next subsection investigates IPD schedule performance.

4.2.3 Schedule Performance Metrics

Data for three standard schedule performance metrics were available for most of the projects: *construction speed*, *delivery speed*, and *construction schedule growth*. Construction speed is measured in square feet per day, starting from the construction Notice to Proceed and ending at the project Substantial Completion. Delivery speed is also measured in square feet per day, starting from the design start date and ending at the occupancy date. Construction schedule growth is measured in percentage terms by comparing the final construction schedule to the original estimated construction schedule.

In addition to these typical schedule performance metrics, a supplementary metric was used to gauge the intensity of the construction schedule by measuring the average dollar value of construction work completed per day. This metric is called *intensity*. The rationale behind measuring schedule intensity is the fact that schedules are based on estimates, and some estimates are more aggressive than others. The intensity metric will provide another comparison of construction speed by normalizing with respect to the amount of construction work put in place during the same timeframe.

The boxplots on the upper left corner of *Figure 28* show data for construction speed, and the boxplots on the upper right corner show data for delivery speed. In both cases, one can see that the median represented by the thick black line in the middle of the green IPD sample, is higher than the median in the middle of the red non-IPD sample. The boxplots show IPD projects have a slightly superior schedule performance over the non-IPD projects.

Furthermore, the boxplots on the lower left side show data for the schedule intensity metric, and the boxplots on the lower right side show data for construction schedule growth in percent of the initial schedule estimate. Based on these boxplots, IPD projects seem to have a higher intensity but also a larger construction schedule growth. A statistical analysis is conducted to examine these claims.

Similar to the analysis conducted for cost and quality performance, both two-sided and one-sided MWW tests were conducted to compare schedule performance for the IPD projects sample and the non-IPD sample. Both the two-sided and one-sided tests for construction speed and for construction intensity showed no significant differences. However, the one-sided MWW test conducted for delivery speed shows a p-value of 0.057. Although this result is not significant at the 0.05 level, it is nevertheless interesting because it provides some evidence (at the 0.10 level) validating assumptions that IPD enhances overall project schedules by overlapping the design and construction phases. This specific result will be discussed in more detail in *Section 4.3*, which shows that the normality assumption holds for the delivery speed data, resulting in the t-test being more appropriate than the MWW test in this situation. When the t-test is performed to compare the delivery speed for IPD and non-IPD samples, the p-value equals 0.046 which means differences in delivery speed are statistically significant. Moreover, the estimate for the difference is about 54 square feet per day, and the 95% confidence interval for the difference is (-9.7, 117.5). Although the interval includes zero and possible negative values, most of it is positive and can get up to 117.5

additional square feet per day for IPD projects. Delivery speed is arguably the most important schedule metric because it is at the highest level of scheduling, and encompasses the whole construction phase and all schedule growths.

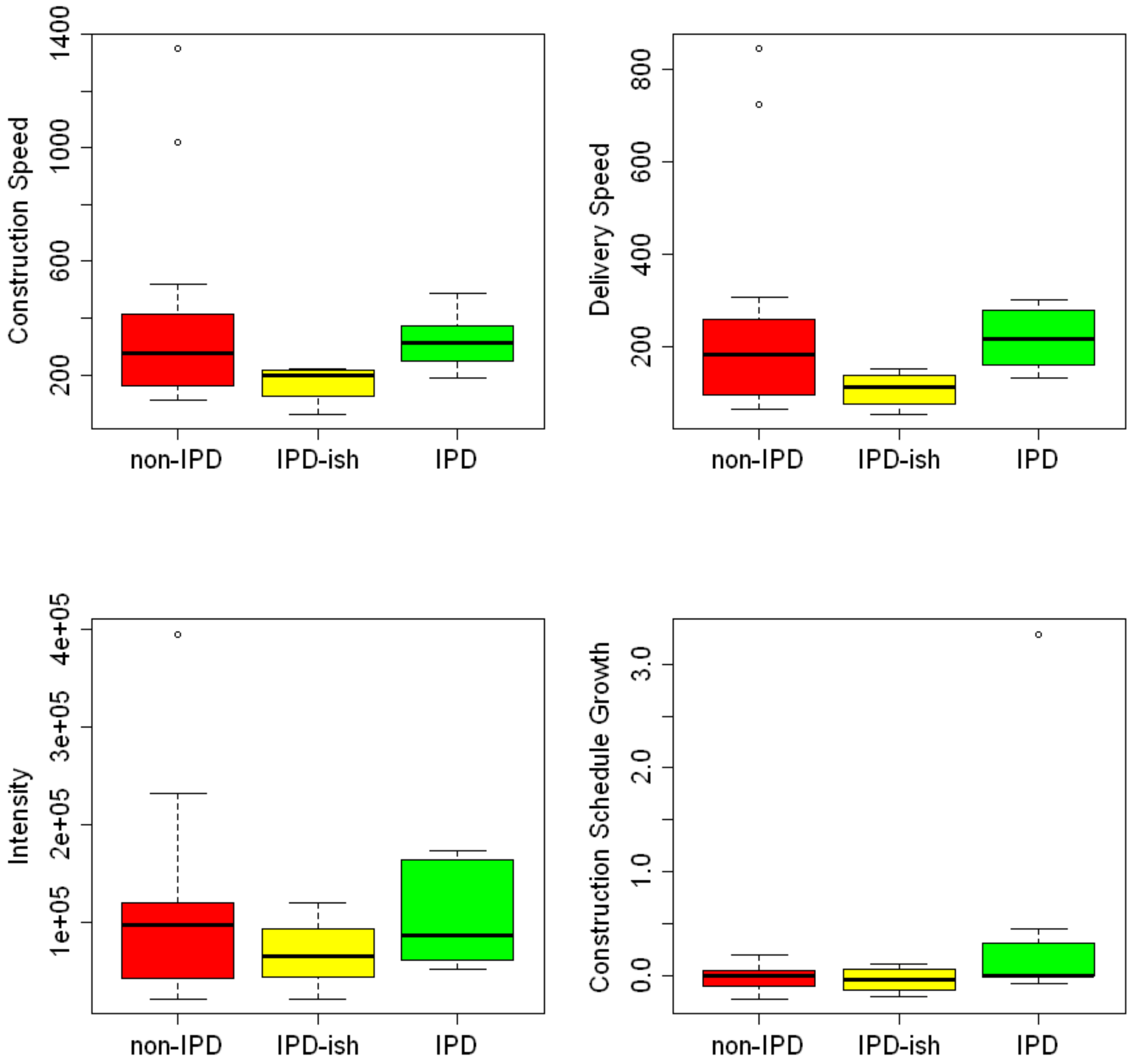


FIGURE 28: Boxplots for Schedule Metrics

Tests also were conducted for delivery schedule growth and construction schedule growth. The first showed no significant differences between IPD and non-IPD projects, while the second showed a p-value of 0.076 also not significant at the 0.05 level used in this study, but nevertheless interesting because it suggests IPD projects experience a higher construction schedule growth than non-IPD projects. However, even with possible higher construction schedule growth, IPD projects still have a superior delivery speed.

After having found that none of the schedule performance metrics were statistically significant at the 0.05 level between IPD and non-IPD projects, the last analysis attempts to confirm these findings by using a comprehensive metric that combines schedule metrics using PCA. The first principal component of the schedule performance metrics explains more than 69% of the variance in the original metrics (construction speed, delivery speed, schedule intensity, and schedule growth). $PC1_s$ was able to summarize the three schedule speed metrics in one metric, and $PC2_s$ denotes schedule growth alone. Together they explain more than 94% of the variance in the original metrics. PC1 and PC2 are expressed as:

- $PC1_s$=-0.585*ConSpeed-0.591*DelivSpeed-0.555*Intensity
- $PC2_s$=0.997*ScheduleGrowth

The interpretation of signs for the different variables making up a principal component mostly reveal the contrast between positive and negative variables, and the value of the coefficient typically shows to what extent a specific variable is responsible for the total component score. When the values for the principal components were compared for IPD and non-IPD projects, the p-values were larger than 0.05 which denotes no significant differences between the two samples. This last result concludes the schedule performance analysis for IPD projects, and a summary of the tests discussed are presented in *Table 17*.

TABLE 17: Hypothesis Tests for Schedule Performance

Hypothesis Test Number	Hypothesis	p-value	Outcome at 95% level
1	IPD projects result in a higher construction speed than non-IPD projects	0.168	Fail
2	IPD projects result in a higher delivery speed than non-IPD projects	0.046	Pass
3	Non-IPD projects experience a smaller construction schedule growth than IPD projects	0.076	Fail
4	Non-IPD projects experience a smaller delivery schedule growth than IPD projects	0.281	Fail
5	IPD projects result in a higher schedule intensity than non-IPD projects	0.141	Fail
6	IPD projects see an overall faster schedule performance over non-IPD projects ($PC1_S$)	0.288	Fail
7	IPD projects see an overall larger schedule growth over non-IPD projects ($PC2_S$)	0.063	Fail

4.2.4 Safety Performance Metrics

Three safety metrics were measured: (1) the number of OSHA recordables, (2) the number of lost-time-injuries (LTI), and (3) the number of fatalities. Fortunately, there were no fatalities on any of the projects surveyed; therefore, fatalities are not included in this analysis.

Incidence rates were calculated for both the recordables and LTI. Per the Bureau of Labor Statistics formula, incidence rates are computed by multiplying the number of recordables or LTI by 200,000, and then dividing by the total hours worked (BLS 2012). The 200,000 hours represent the equivalent of 100 employees working 40 hours per week. This computation provides a means to normalize the values for projects of different sizes. Since the total project hours were not always available, another type of normalization was used by dividing the number of recordables or LTI by the cost of the project. The analyses for both metrics are discussed below.

The upper left corner of *Figure 29* shows the boxplots for the OSHA Recordables Incidence Rate, and the lower left corner shows boxplots for the number of OSHA recordables per million dollars of construction, for IPD and non-IPD projects. It is not easy to spot a difference between the IPD and non-IPD OSHA recordables, but one can notice that the distributions are much wider for non-IPD projects (in red), allowing for higher values.

The upper right corner of the figure shows LTI incidence rates, while the lower right corner shows boxplots for the number of LTI per million dollars. It is fairly easy to see that IPD projects have a considerably smaller range for LTI as compared to their non-IPD counterparts.

Two-sided and one-sided MWW tests were conducted to statistically compare OSHA recordables for the IPD and non-IPD project samples. Both the incidence rates and the number of recordables per million dollars were compared. As expected based on the boxplots, none of the four tests showed significant differences between the IPD and non-IPD samples. The p-values varied from 0.425 to 0.982.

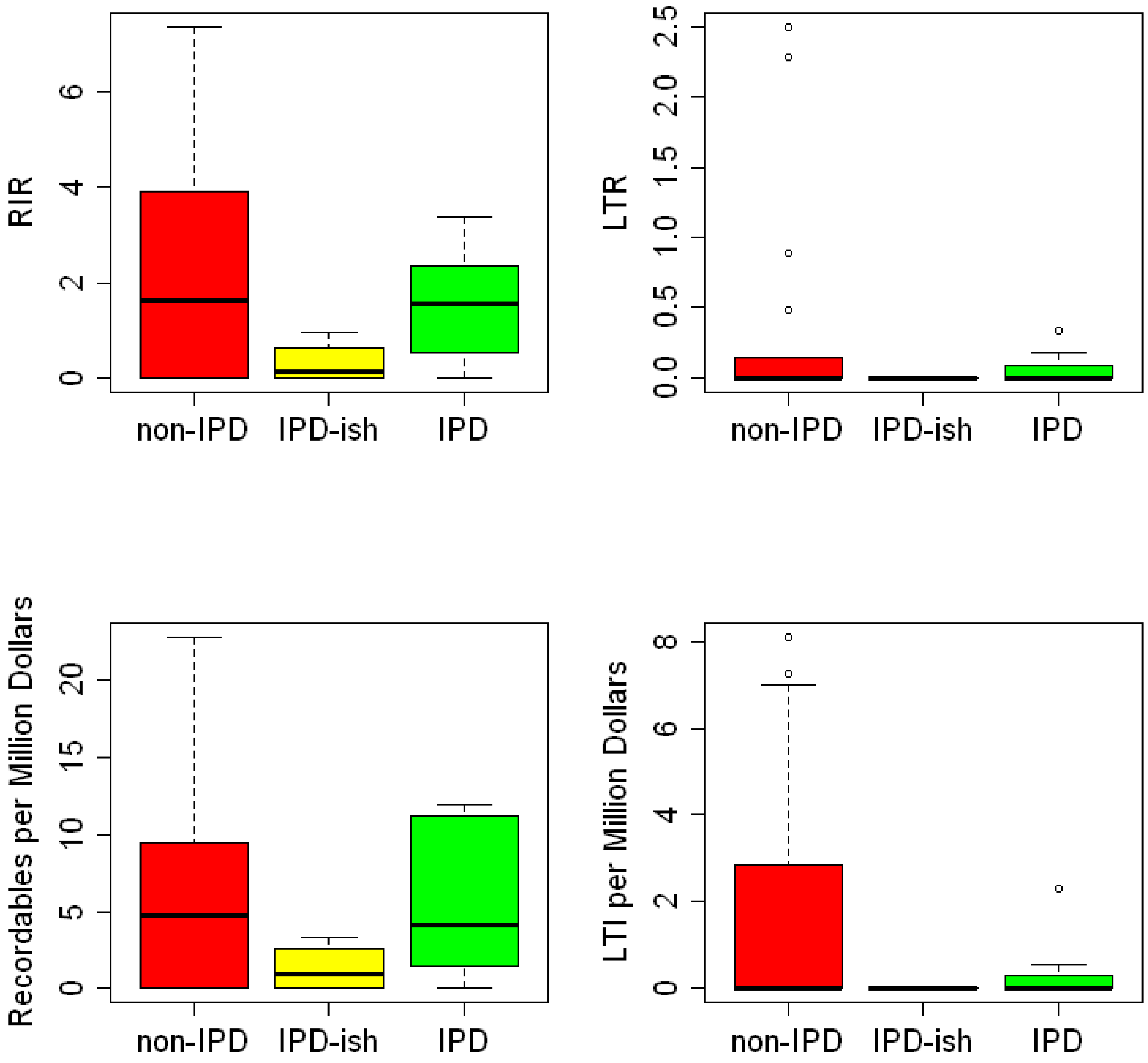


FIGURE 29: Boxplots for Safety Metrics

Similarly, two-sided and one-sided MWW tests were conducted for the two metrics measuring LTI: incidence rate and LTI per million dollars. Tests on the incidence rate did not show significant results. However, tests on the LTI per million dollars, for which more data was available, showed a p-value of 0.083. Although this value is not enough to show significant differences between IPD and non-IPD projects at the 0.05 level, this finding warrants further discussion. This result is significant at the more lenient 0.10 level. In fact, the 95% confidence interval for the difference ranges from 0.001 to 2.739 LTI per million dollars. A larger dataset could possibly substantiate these claims at the 0.05 significance level. In reality, LTI were found to have statistically significant differences when an early analysis was performed with half the dataset, before completing all the data collection for the study. For the moment, a column was added to *Table 18*, showing that LTI differences are significant at the less stringent 0.10 significance level.

After seeing that some of the safety performance metrics were statistically significant at the 0.10 level between IPD and non-IPD projects, the last analysis used a comprehensive metric that combined the two safety metrics with PCA. This analysis gives an overall picture of safety performance. The first principal component of the safety performance metrics explains 67% of the variance in the original quality metrics, and therefore can be used to summarize safety metrics. PC1 is presented here:

- $PC1_f$= -0.707*Recordables-0.707*LostTimeInjuries

When $PC1_f$ was compared for IPD and non-IPD projects, the p-value was 0.226, which shows no significant differences in safety performance between the delivery systems. This result means that unlike quality, the comprehensive safety metric was not different for IPD and the improvements in safety performance are limited to LTI. This last result concludes the safety performance analysis for IPD projects and a summary of the tests discussed are presented in *Table 18*.

TABLE 18: Hypothesis Tests for Safety Performance

Hypothesis Test Number	Hypothesis	p-value	Outcome at 95% level	Outcome at 90% level
1	IPD projects result in a less OSHA recordables than non-IPD projects	0.425	Fail	Fail
2	IPD projects result in less LTI than non-IPD projects	0.083	Fail	Pass
3	IPD projects see an overall superior safety performance over non-IPD projects ($PC1_f$)	0.226	Fail	Fail

4.2.5 Project Change Performance Metrics

In addition to cost, quality, schedule, and safety metrics, several project change performance metrics were targeted for data collection. Overall, change performance data included three types of metrics:

1. Total percent of change in the project;

2. Reason for the changes: CII Research Report 158-11 (2001) shows the two key reasons for changes are project additions and design-related changes (including design changes, design coordination, and design errors). Data was collected to assess these two types of change for each project. In addition, the industry panel for this research requested that data be collected for changes due to code or major regulatory agencies;

3. Average change order processing time, defined as the period of time between the initiation of the change order and the owner's approval of the change order.

Data for total change and reasons for the change orders were gathered in percentage terms, while data for the average change order processing time was collected in days (i.e. 1 to 7 days, 8 to 14 days, etc.).

The boxplots in the upper left corner of *Figure 30* show data for the overall percent change experienced by the IPD and non-IPD projects. One can notice the wider distribution and the larger median for non-IPD projects, so the figure clearly shows IPD projects experienced fewer changes than their non-IPD counterparts.

The boxplots in the upper right corner show the data for the changes related to design issues, again showing IPD projects experience considerably less design changes. The lower left corner shows program changes related to additions and modifications, and IPD projects seem to have a slightly lower median, which remains to be tested statistically.

The change order processing time is displayed in the boxplots in the lower right corner. The units for the x-axis are weeks, and the difference between the IPD and non-IPD processing times are clearly visible. In fact, the value for IPD projects is approximately one week, whereas change orders need four times longer to be processed for non-IPD projects.

Two-sided and one-sided MWW tests were conducted to statistically compare project change performance metrics for the IPD and non-IPD project samples. All metrics introduced above were compared. The differences in total percent change were not significant at the 0.05 level with a p-value of 0.224. The differences in changes due to additions and deletions also were insignificant with a p-value of 0.334.

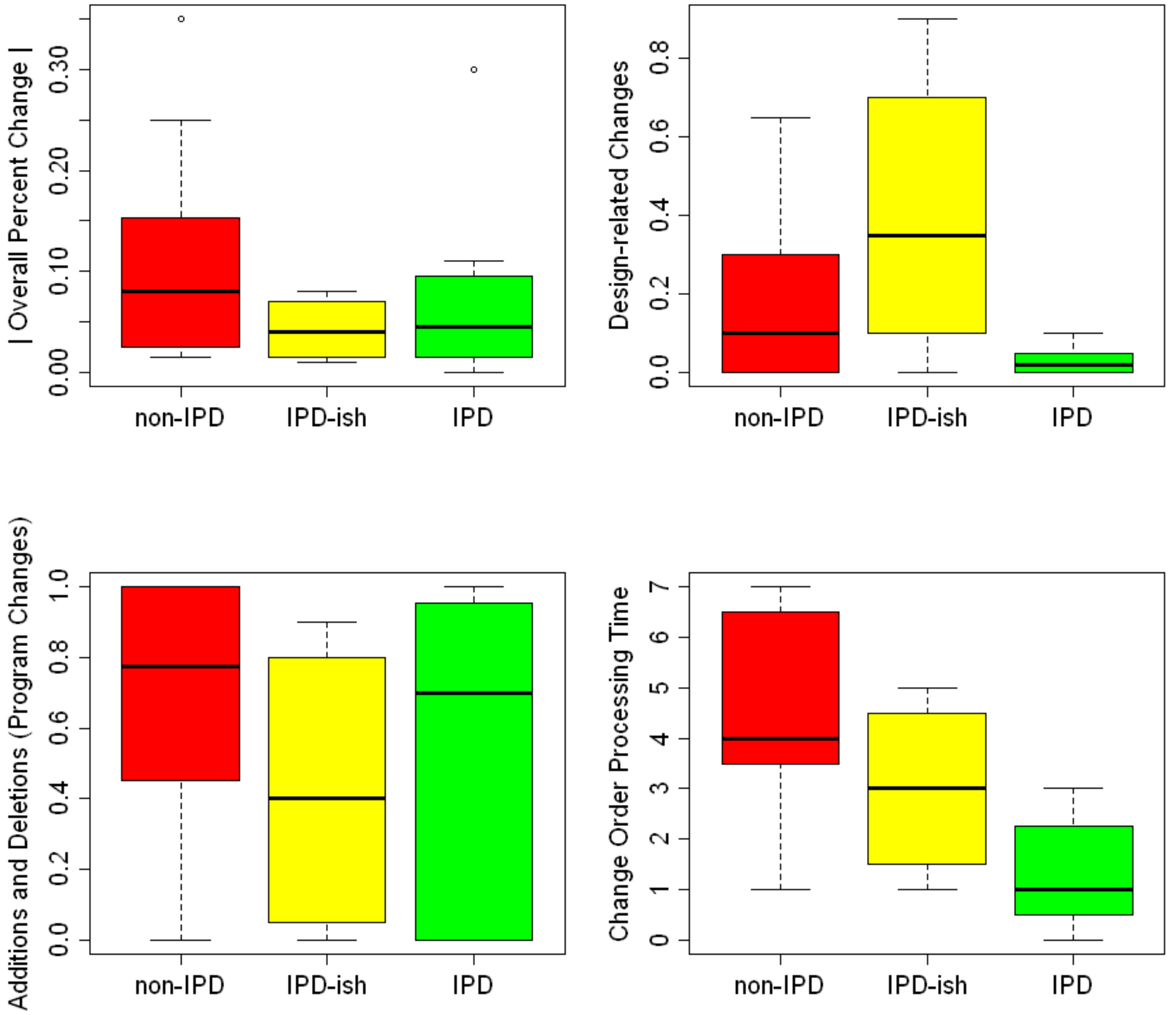


FIGURE 30: Boxplots for Change Metrics

However, the differences in design changes were significant at the 0.05 level with a p-value 0.029. The median value for non-IPD projects is 10% versus 1.8% for the IPD projects. The point estimate for the difference is only 5%, with a 95% confidence interval ranging between zero and 25%.

The difference in changes stemming from major regulatory agencies also are significant at the 0.05 level with a p-value of 0.049. The median value for non-IPD projects is 5% versus zero for the IPD projects. The point estimate for the difference is only 1.2%, with a 95% confidence interval ranging between zero and 5%.

Even more noteworthy are the differences in change order processing times, which also are statistically significant at the 0.05 level and the more conservative 0.01 level, with a p-value of 0.000. As stated earlier, the median value for non-IPD projects is four weeks versus one week for the IPD projects. The point estimate for the difference is three weeks with a 95% confidence interval ranging between two and five weeks.

After determining that some of the project change performance metrics were statistically significant between IPD and non-IPD projects, the last analysis used a comprehensive metric that combined all change metrics using PCA. The first two principal components of the change performance metrics explain 65% of the variance in the original change metrics (shown in the equations below). $PC1_g$ highlights additions or program changes, while $PC2_g$ mostly highlights change order processing time.

- $PC1_g$ =-0.408*OverallChange -0.556*Additions +0.486*Design +0.494*Regulatory -0.209*ProcessingTime
- $PC2_g$ =-0.472*OverallChange -0.312*Design+0.383*Regulatory -0.730*ProcessingTime

When the principal components were compared for IPD and non-IPD projects, $PC2_g$ showed very significant differences, while $PC1_g$ did not. This result confirms the previous findings testing individual change performance metrics, showing significant change improvements for IPD projects on some metrics, and no significant differences based on other metrics. In summary, IPD improves design-related changes and regulatory changes, as well as the change order processing time. This last result concludes the change performance analysis for IPD projects, and a summary of the tests discussed are presented in *Table 19*.

TABLE 19: Hypothesis Tests for Change Performance

Hypothesis Test Number	Hypothesis	p-value	Outcome at 95% level
1	IPD projects result in less project changes than non-IPD projects	0.224	Fail
2	IPD projects result in less program-related changes (additions/deletions) than non-IPD	0.334	Fail
3	IPD projects result in less design-related changes than non-IPD projects	0.029	Pass
4	IPD projects result in less changes due to regulatory issues than non-IPD projects	0.049	Pass
5	IPD projects result in a faster change order processing time than non-IPD projects	0.000	Pass
6	IPD projects see an overall superior change performance over non-IPD projects ($PC1_g$)	0.318	Fail
7	IPD projects see an overall faster change performance over non-IPD projects ($PC2_g$)	0.001	Pass

The decrease in design changes as well as regulatory changes for IPD can be due to the high level of involvement of key project stakeholders throughout the entire project timeline. The contractors’ involvement in the design phase and the designers’ involvement in the construction phase can result in an increased common understanding of the project, and therefore, a reduction in design-related changes. The same phenomenon happens when regulatory officials are involved in the project and cause less changes due to code and regulations.

Additionally, the much faster processing time for changes can be associated with the weekly meetings of the core groups leading IPD projects, which have the authority needed to make project-related decision such as approving and processing changes. One other performance area that might be similarly affected by IPD is communication performance.

4.2.6 Communication Performance Metrics

This study offers a broad definition of project performance, beyond the typical triangle of cost, schedule and quality. So far, safety and changes have been discussed, and next are communication performance metrics. Communication performance refers to direct means of communication as well as process inefficiencies and work that needed to be redone. This subsection focuses on requests for information (RFI), rework and resubmittals.

RFI are considered a communication performance metric because they can be an important source of waste for projects. The reason is simple: crews lose productivity while waiting for information, especially when it takes weeks for other project parties to respond. Often, these crews have to demobilize and remobilize more than once, which can add costs to the project. Additionally, there are crew morale and learning curve effects that have been studied and documented when such events happen (Hanna et al. 2004b).

RFI data includes two metrics: (1) the number of RFI and (2) the RFI processing time. To normalize the RFI values in order to compare projects of different sizes, the number of RFI is divided by the project construction cost. The boxplots in the upper left corner of *Figure 31* show data for the number of RFI per million dollars for IPD and non-IPD projects. The difference in the medians is straightforward: approximately ten RFI for non-IPD projects, compared to approximately two RFI for IPD projects. The upper right corner of the figure shows the boxplots for the RFI processing time. Again, IPD projects have much lower values than their non-IPD counterparts.

The lower left corner of the figure shows comparable median values for rework, and the lower right corner shows considerably less resubmittals for IPD projects. These findings need to be confirmed with statistical testing.

Two-sided and one-sided MWW tests were conducted to statistically compare RFI for the IPD and non-IPD samples. Both the number of RFI per million dollars and the RFI processing times were compared. The differences in number of RFI were significant at a 0.05 level and the more stringent 0.01 significance level, with a p-value of 0.001. The median for IPD projects was about 9.61 RFI per million dollars, compared to 1.81 RFI for non-IPD projects. The point estimate for the difference is 8.23 RFI with a 95% confidence interval ranging between 4.10 and 14.88 RFI per million dollars. Data was also collected for *work-arounds,* or alternative means used to avoid RFI, such as phone calls or emails. There were no significant differences in work-arounds between IPD and non-IPD projects, which further strengthens the RFI results.

The differences in the RFI processing times were significant at a 0.05 level with a p-value of 0.025. The median for non-IPD projects was two weeks compared to one week for IPD projects. The point estimate for the difference is one week with a 95% confidence interval ranging between zero and two weeks.

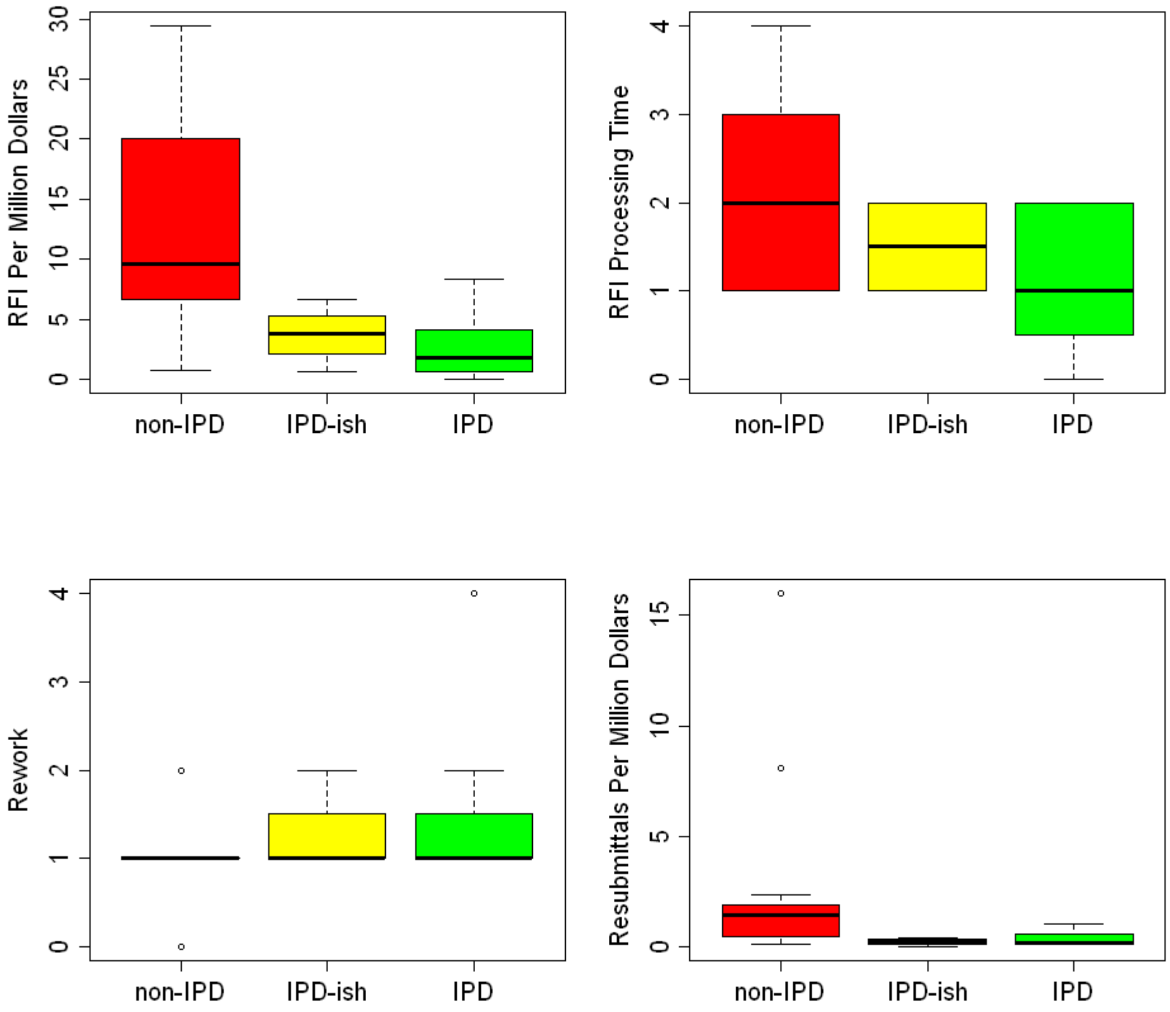


FIGURE 31: Boxplots for Communication Metrics

Rework was tracked in percent of overall cost, and two-sided and one-sided MWW tests were conducted to statistically compare performance between the IPD and non-IPD samples. The differences in rework are not statistically significant at a 0.05 level with a p-value of 0.173.

The number of resubmittals was divided by the project construction cost in million dollars to normalize the data for projects of different sizes. The differences in the number of resubmittals between IPD and non-IPD projects are statistically significant at the 0.05 level with a p-value of 0.018. The median for IPD projects is about 0.20 resubmittals per million dollars compared to 1.44 resubmittals for non-IPD projects. The point estimate for the difference is 0.94 resubmittals with a 95% confidence interval ranging between 0.01 and 1.80 resubmittals per million dollars.

Claims and litigations can also be discussed as part of this subsection. In the whole dataset, no IPD projects experienced any claims, while three non-IPD projects experienced claims.

To confirm that communication performance metrics are statistically significant between IPD and non-IPD projects, the last analysis used a comprehensive metric that combined all communication metrics using PCA. The first two principal components of the communication performance metrics explain 71% of the variance in the original metrics. $PC1_m$ highlights the number of RFI and their processing time, as well as the number of resubmittals. $PC2_m$ highlights construction rework. $PC1_m$ and $PC2_m$ are presented here.

- $PC1_m$ =-0.577*RFI -0.506*ProcessingTime -0.129*Rework
  -0.539*Resubmittals -0.322*Claims
- $PC2_m$ =-0.219*ProcessingTime +0.948*Rework-0.175*Resubmittals

  +0.129*Claims

There were several missing values for the resubmittals variable, and therefore as discussed earlier for the quality area, an additional principal component analysis is completed without this variable. As shown below, the new principal components can be interpreted similarly to the previous ones, with $PC1_m$’ mostly highlighting the number and processing time of RFI as well as the number of resubmittals, and $PC2_m$’ highlighting rework. Both results are tested, and since they are similar only $PC1_m$ and $PC2_m$ are presented in *Table 20*.

- $PC1_m$’=+0.640*RFI+0.577*ProcessingTime +0.176*Rework +0.476*Claims
- $PC2_m$’=-0.324*ProcessingTime +0.938*Rework

When the principal components were compared for IPD and non-IPD projects, the results were similar to the previous analyses. $PC1_m$’ showed very significant differences, while $PC2_m$’ did not. This result confirms the previous findings testing individual communication performance metrics, which demonstrated significant improvements for IPD projects with respect to all communication metrics except rework. This last result concludes the communication performance analysis for IPD projects. A summary of the tests discussed above are presented in *Table 20*.

TABLE 20: Hypothesis Tests for Communication Performance

Hypothesis Test Number	Hypothesis	p-value	Outcome at 95% level
1	IPD projects result in less RFI's than non-IPD projects	0.001	Pass
2	IPD projects result in a faster RFI processing time than non-IPD projects	0.025	Pass
3	IPD projects result in less rework than non-IPD projects	0.173	Fail
4	IPD projects result in less resubmittals than non-IPD projects	0.018	Pass
5	IPD projects see an overall superior communication performance over non-IPD projects ($PC1_m$)	0.009	Pass
6	IPD projects see a superior rework performance over non-IPD projects ($PC2_m$)	0.188	Fail

4.2.7 Labor Performance Metrics

Labor is often one of the high risk items on construction projects, especially given that labor costs can make up to half of the total project cost. Therefore, labor performance is an important aspect of overall project success. Three labor performance metrics were available for data collection: (1) extent to which additional labor is used, in terms of overtime, second shift work, and over-manning; (2) trend of Percent Plan Complete (PPC), or the measure of work flow reliability, which is calculated by dividing the number of actual task completions by the number of planned tasks; and (3) labor factor, measured as a ratio of the total cost of self-performed work divided by the labor cost of self-performed work.

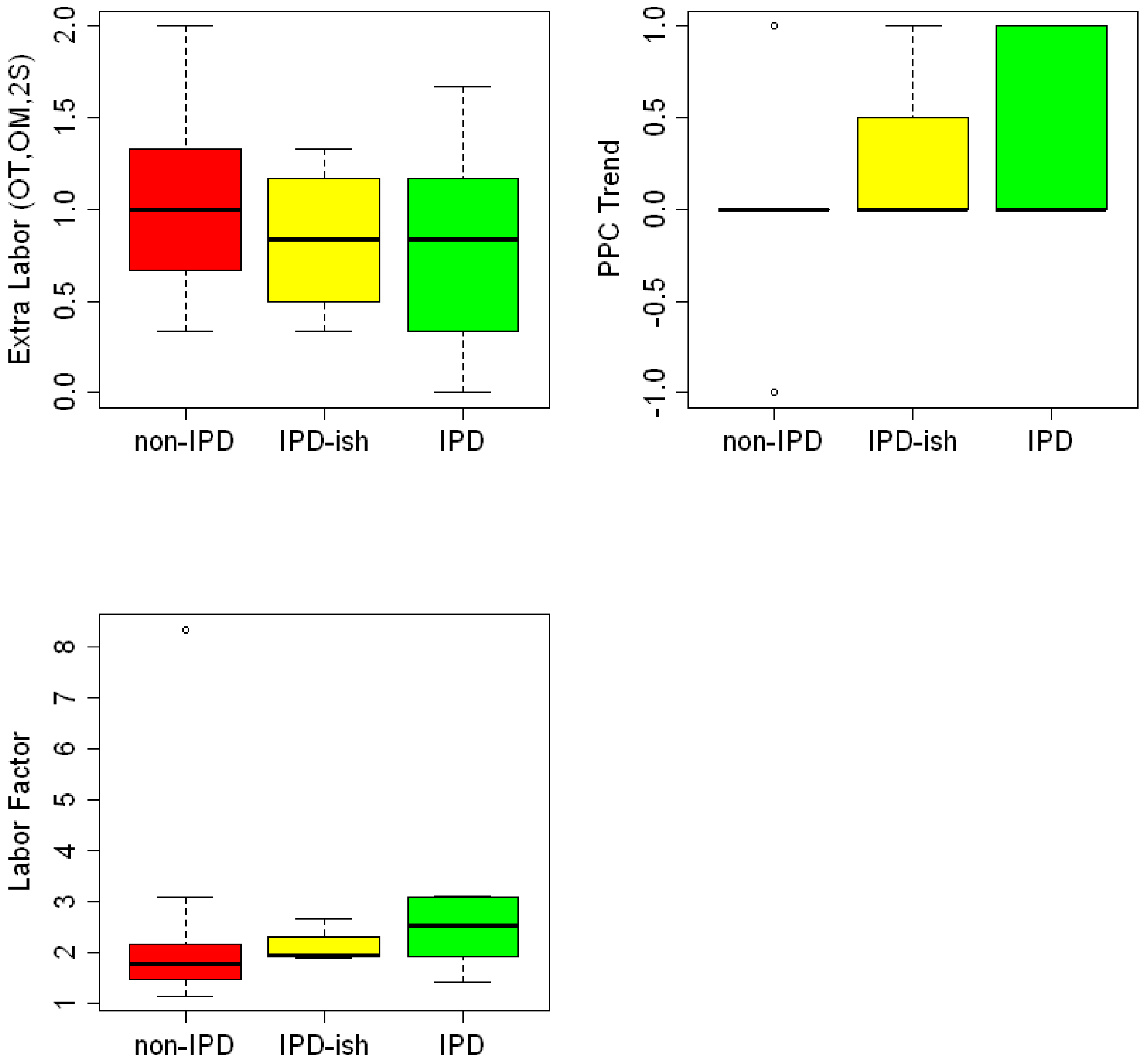


FIGURE 32: Percentage of Labor Cost (left) and Labor Factor (right)

The boxplots in the upper left corner of *Figure 32* show that IPD projects use less extra labor than non-IPD projects. The boxplots in the upper right corner show that IPD projects have a positive PPC trend, as compared to the not so encouraging stagnant PPC for non-IPD projects. Finally, one can see that the Labor Factor for IPD projects is higher than for non-IPD projects, potentially meaning IPD projects use labor more efficiently. Statistical testing is needed to examine the significance of these findings.

Two-sided and one-sided MWW tests were conducted to statistically compare labor performance metrics. None of the tests resulted in significant differences between the IPD projects sample and the non-IPD sample at the 0.05 level. However, PPC and labor factor results were significant at the 0.10 level. The one-sided test for labor factor gave a p-value of 0.094 (with 1.76 the median for non-IPD and 2.53 the median for IPD projects) and the test for PPC trend gave a p-value of 0.072, indicating superior labor performance for IPD projects, but only at the 0.10 significance level.

The last analysis used a comprehensive metric that combined all labor metrics using PCA. The first and second principal components of the labor performance metrics, $PC1_L$ and $PC2_L$, together explain 81% of the variance in the original labor metrics. $PC1_L$ highlights the labor factor and the use of extra labor, while $PC2_L$ mostly highlights the PPC trend. $PC1_L$ and $PC2_L$ are presented here:

- $PC1_L$ =0.720*LaborFactor +0.250*PPCtrend +0.647*ExtraLabor
- $PC2_L$ =-0.901*PPCtrend +0.427*ExtraLabor

When the principal components were compared for IPD and non-IPD projects, no significant differences were observed at the 0.05 level used for this study. However, the results for both $PC1_L$ and $PC2_L$ show relatively small p-values that are significant at the 0.10 level. This finding could suggest that the overall labor performance for IPD projects might be superior, and more data is needed to validate this result. A summary of the tests discussed is presented in *Table 21*.

TABLE 21: Hypothesis Tests for Labor Performance

Hypothesis Test Number	Hypothesis	p-value	Outcome at 95% level	Outcome at 90% level
1	IPD projects use less extra labor (overtime, second shift work, and overmanning) than non-IPD projects	0.230	Fail	Fail
2	IPD projects experience a better PPC trend than non-IPD projects	0.072	Fail	Pass
3	IPD projects experience a larger labor factor than non-IPD projects	0.093	Fail	Pass
4	IPD projects see a superior labor intensity over non-IPD projects ($PC1_L$)	0.063	Fail	Pass
5	IPD projects see a superior labor reliability over non-IPD projects ($PC2_L$)	0.079	Fail	Pass

4.2.8 Recycling Performance Metrics

In addition to project-specific performance metrics, the impact a construction project has on the environment also needs to be reported. One such environmental metric is the material recycling rate. The available recycling data includes three metrics: (1) total value of material waste (in tons, normalized per million dollars); (2) percentage of waste recycled; and (3) percentage of waste sent to landfills. *Figure 33* shows the boxplots for tons of material waste and for percentages of waste recycled for IPD and non-IPD projects. The difference in the medians is quite visible for total material waste, where non-IPD projects produce about twice as much waste as IPD projects. The distribution is also much wider for non-IPD projects. The difference is not that obvious for the recycling rate, with IPD projects recycling only slightly more.

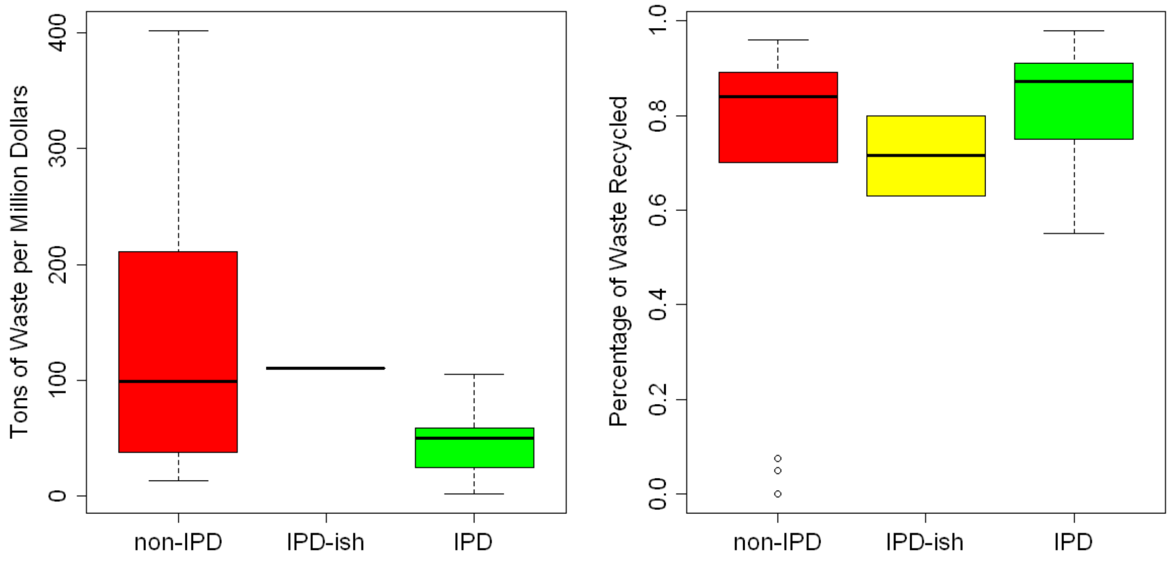


FIGURE 33: Boxplots for Recycling Metrics

Both two-sided and one-sided MWW tests were conducted to statistically compare material waste and recycling performance metrics. The tests for recycling rate did not result in significant differences between the IPD sample and the non-IPD sample at the 0.05 level, with a p-value around 0.242. The tests for tons of material waste also did not result in significant differences at the 0.05 level, with a p-value around 0.077. However, this p-value is considered significant if looking at the less conservative 0.10 significance level. A larger sample size might bring this value down to the significance level used in this study. In fact, just like for delivery speed, Section 4.3 will show that a t-test is more appropriate in this specific situation, because the recycling data is normally distributed. The t-test will result in a p-value of 0.022 which means the differences are significant and IPD projects produce significantly less material waste than non-IPD projects.

To follow-up on these results for recycling performance metrics, the last analysis used a comprehensive metric that combined all recycling metrics using PCA. The first principal component of the recycling performance metrics explains more than 65% of the variance in the original metrics, and can be expressed as:

- $PC1_r = 0.707*WasteTons + 0.707*Recycled$

When $PC1_r$ was compared for IPD and non-IPD projects, the p-value was 0.134, which is not significant at the 0.05 level. A summary of the tests discussed are presented in *Table 22*.

TABLE 22: Hypothesis Tests for Recycling Performance

Hypothesis Test Number	Hypothesis	p-value	Outcome at 95% level
1	IPD projects generate less material waste than non-IPD projects	0.022	Pass
2	IPD projects have a higher recycling rate than non-IPD projects	0.242	Fail
3	IPD projects see a superior recycling performance over non-IPD projects ($PC1_r$)	0.134	Fail

4.2.9 Business Performance Metrics

Like any for-profit organization, contractors only can afford to remain in the construction business if they make a reasonable monetary profit. Therefore, profit is a key performance metric from the contractors' perspective. However, it is impractical to ask contractors how much profit they made on specific projects. The simplest way to avoid blank responses (and awkward moments) is to group job overhead and profit (OH&P) together into the same metric. Obviously, this metric would be sensitive to the changes in overhead from project to project, but holding constant the company variable one can assume overhead to be relatively stable over several projects.

Business performance metrics included one additional metric: the effect of the project on the company image and potential for return business. Although qualitative, this metric identifies projects that lead to immediate return business and others that lead to a bad working relationship with clients.

The left side of *Figure 34* shows the boxplots for OH&P for IPD and non-IPD projects. The values on the vertical axis represent windows of values (0 for *negative OH&P*, 1 for *less than 5%*, 2 for *5 to 10%*, and 3 for *11 to 15 %*.) The median for non-IPD projects was *less than 5%*, while the median for IPD projects was *5 to 10%*. A few IPD projects had *11 to 15%* OH&P, while non-IPD projects did not have any projects with values above *10%*.

The right side of the figure shows responses for return business. Here the values on the vertical axis represent the potential of the project for return business, from *-2* for “very negative” to *+2* for “very positive.” The non-IPD projects experiences some *negative* and *very negative* responses, while the lowest response for IPD projects was *positive*.

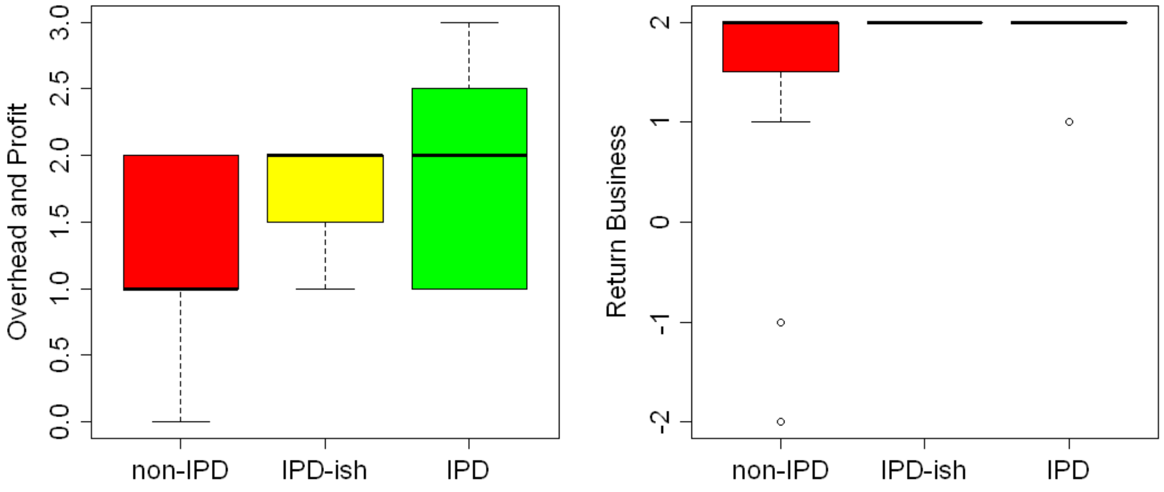


FIGURE 34: Boxplots for Business Metrics

Both two-sided and one-sided MWW tests were conducted for the IPD projects sample and the non-IPD sample to statistically compare OH&P and the project's impact on company image and potential for return business. The differences in OH&P were significant at a 0.05 level with a p-value of 0.0248. The differences in image and return business had a p-value of 0.211 which is not considered significant at the 0.05 significance level.

After seeing that one of the business performance metrics was significant and the other wasn't, when comparing IPD and non-IPD projects, one last analysis used a comprehensive metric that combined the two business metrics using PCA. The first principal component of the business performance metrics explains about 70% of the variance in the two original metrics, and consists of a linear combination of these metrics:

- $PC1_b$=0.707*Overhead&Profit+0.707*Image&ReturnBusiness

When $PC1_b$ was compared for IPD and non-IPD projects, the p-value was 0.038 showing significant differences at the 0.05 level. This result confirms the previous findings testing individual business performance metrics, which showed significant improvements for IPD projects. *Table 23* presents a summary of the tests discussed above.

TABLE 23: Hypothesis Tests for Business Performance

Hypothesis Test Number	Hypothesis	p-value	Outcome at 95% level
1	IPD projects result in a higher overhead and profit than non-IPD projects	0.024	Pass
2	IPD projects result in a better company image and return business than non-IPD projects	0.211	Fail
3	IPD projects see a superior business performance over non-IPD projects ($PC1_b$)	0.029	Pass

This section provided a first quantitative understanding of IPD performance as compared to non-IPD projects. There were significant differences at the 0.05 level in performance metrics belonging to six out of the nine areas investigated. MWW results were mainly presented in this section, but in the next section the t-test results will be compared and contrasted to those of MWW, leading to the final conclusions of this chapter.

4.3 A Comparative Look at Univariate Results

This section summarizes univariate results of the previous section, and compares the previously discussed MWW results to t-test results for a more complete understanding of performance differences across all different performance metrics. Comparative results for 31 performance metrics are presented in *Table 24*: the metrics are listed first, followed by their respective t-test results, and finally the MWW results. Unlike the previous section organized by performance area, *Table 24* is organized by increasing p-values of the MWW tests regardless of the performance area. As discussed earlier, for each individual test, the smaller the p-value, the more significant the performance differences are between IPD and non-IPD project. The p-values are color-coded to facilitate the interpretation of the results: all p-values highlighted in red are significant at the 0.05 level, while the values highlighted in green are significant at the less stringent 0.1 level. One-sided results are shown in the table.

When observing the MWW results, the table shows very significant differences (p-values less than 0.01) between IPD and non-IPD for the first four performance metrics, which include metrics in three distinct performance areas: project change, quality, and communication. Depending on the significance level chosen, this list can increase to include 14 metrics at the 0.05 significance level, and up to 18 metrics at the 0.10 significance level. This section discusses the differences between the t-tests and the MWW results. A glossary of terms is provided in *Appendix B* to define some of the abreviations and terms used in this table.

TABLE 24: Comparative Univariate Results for IPD vs. Non-IPD (Ordered by MWW p-value)

Performance Metric		**t-test**			**MWW test**	
		t-statistic	**d.f.**	**p-value**	**W-statistic**	**p-value**
1	Change Order Processing Time	5.57	23	0.000	443	0.000
2	Deficiency Issues	3.67	13	0.001	185	0.001
3	Request For Information	4.64	24	0.000	366	0.001
4	Punchlist Cost	4.15	17	0.000	380	0.003
Above are MWW results statistically significant at the 0.01 level						
5	Punchlist Items	3.90	23	0.000	359	0.014
6	Resubmittals	1.99	13	0.034	173	0.018
7	OH&P	-2.01	9	0.038	279	0.024
8	RFI Time	2.36	13	0.017	381	0.025
9	Quality	-2.02	9	0.037	326.5	0.032
10	Design Changes	3.32	24	0.001	381	0.032
11	Warranty Costs	1.66	8	0.068	256	**0.040**
12	Regulatory Changes	1.81	21	0.042	347.5	0.050
Above are MWW results statistically significant at the 0.05 level						
13	Delivery Speed	-1.80	16	**0.046**	214	0.057
14	PPC Trend	-1.46	12	0.085	181.5	0.072
15	Tons of Waste	2.20	14	**0.022**	122	0.077
16	Const. Schedule Growth	-1.24	6	**0.131**	152	0.079
17	Lost-Time-Injuries	2.44	28	0.011	396	**0.083**
18	Labor Factor	-1.49	6	0.093	95	0.094
Above are MWW results statistically significant at the 0.10 level						
19	Schedule Intensity	-1.01	14	0.164	202	0.141
20	Const. Speed	-0.94	18	0.181	204	0.168
21	Rework	-1.05	7	0.164	326.5	0.173
22	Return Business	-1.41	28	0.084	354.5	**0.211**
23	Total % Change	0.65	11	0.266	331	0.224
24	Labor OT/OM/2S	0.90	10	0.194	384.5	0.230
25	Recycling Rate	-1.35	19	0.096	258	**0.242**
26	Deliv. Schedule Growth	-1.16	6	0.289	84	0.281
27	Unit Cost	-1.04	6	0.168	347	0.330
28	Additions / Deletions	0.59	10	0.283	350.5	0.334
29	Latent Defects	-0.18	10	0.432	210	0.442
30	Cost Growth	-0.34	11	0.370	354.5	0.471
31	OSHA Recordables	0.13	14	0.450	369	0.491

When examining the differences in results between the t-tests and the MWW tests, the focus will be on the values that would change the interpretation of the performance differences between IPD and non-IPD. For example, as shown in rows number 22 and 25 in *Table 24*, both the return business and recycling rate metrics have been found to have insignificant differences when using the MWW test, while the t-test shows results significant at the 0.10 level. In order to find out which of these two tests is more rigorous in this situation, the earlier discussion of MWW versus t-tests is useful. A key fact is that t-tests are more powerful when the data is normally distributed, while MWW tests are more powerful when it is not. Therefore, Q-Q plots will be used here to make this distinction. A normal Q-Q plot can be used to compare the sample data to normal theoretical quantiles; when the points almost form a straight line the distribution can be assumed to be normal.

For the example cited earlier, *Figure 35* shows Q-Q Plots for return business data (top), and for recycling rate data (bottom). None of the plots seem to have points following a straight line, so their data cannot be assumed normally distributed. Therefore, the MWW test can be considered more powerful than the t-test for these two variables, and its results are shown in bold font in the table. The MWW tests in rows number 22 and 25 of *Table 24* show that the differences in performance for return business and for recycling rate are not statistically significant.

Whenever the MWW test and t-test results lead to different interpretations, the value with a bold font shows which test is more appropriate based on the distribution of the data. A summary table with the final results will be presented at the end of this section.

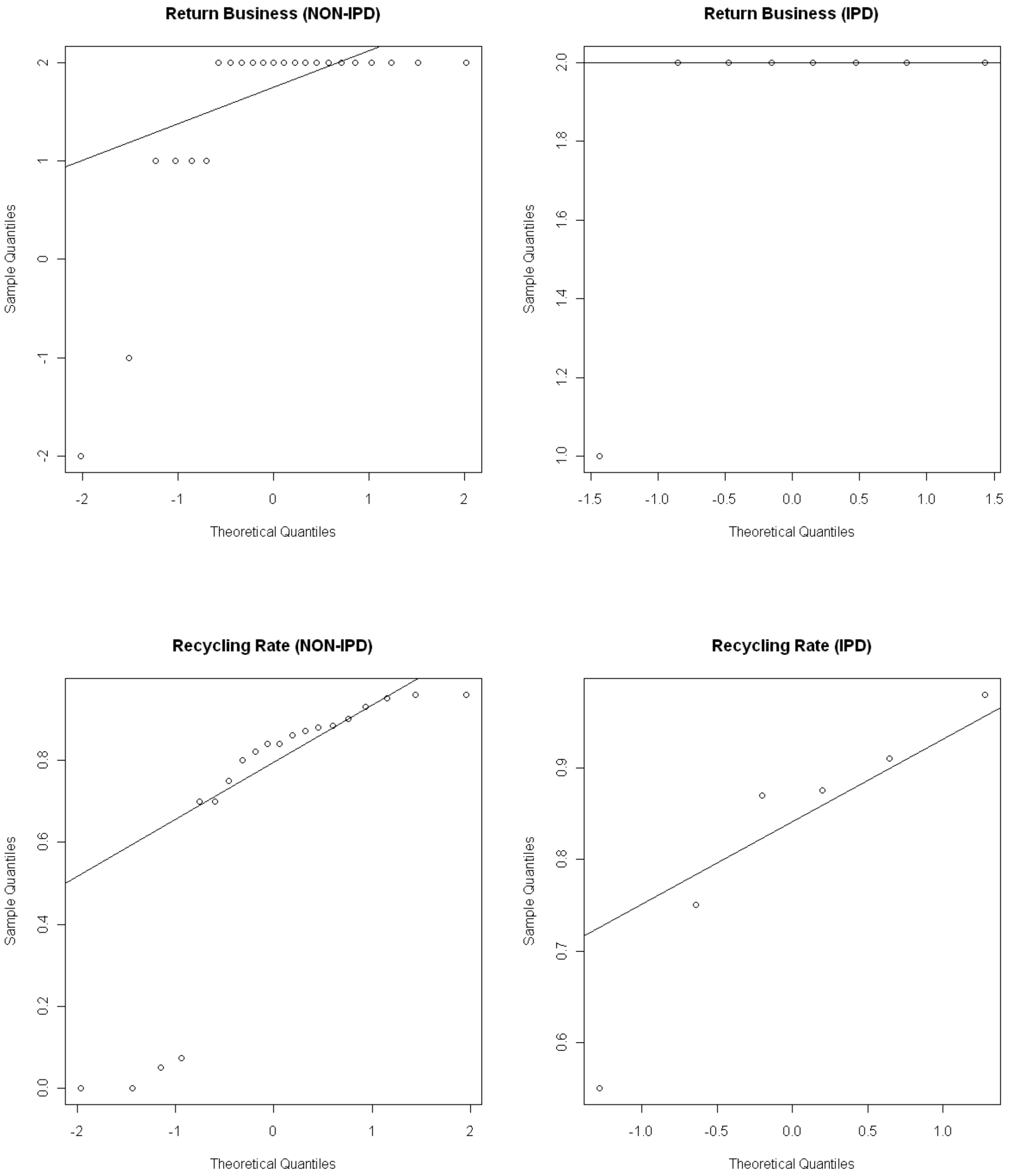


FIGURE 35: Q-Q Plots for Return Business (top) and Recycling Rate (bottom)

A similar instance can be seen for warranty costs and lost-time injuries, in rows 11 and 17 of *Table 24*. *Figure 36* also shows Q-Q Plots for warranty costs data (top), and for lost-time injuries data (bottom).

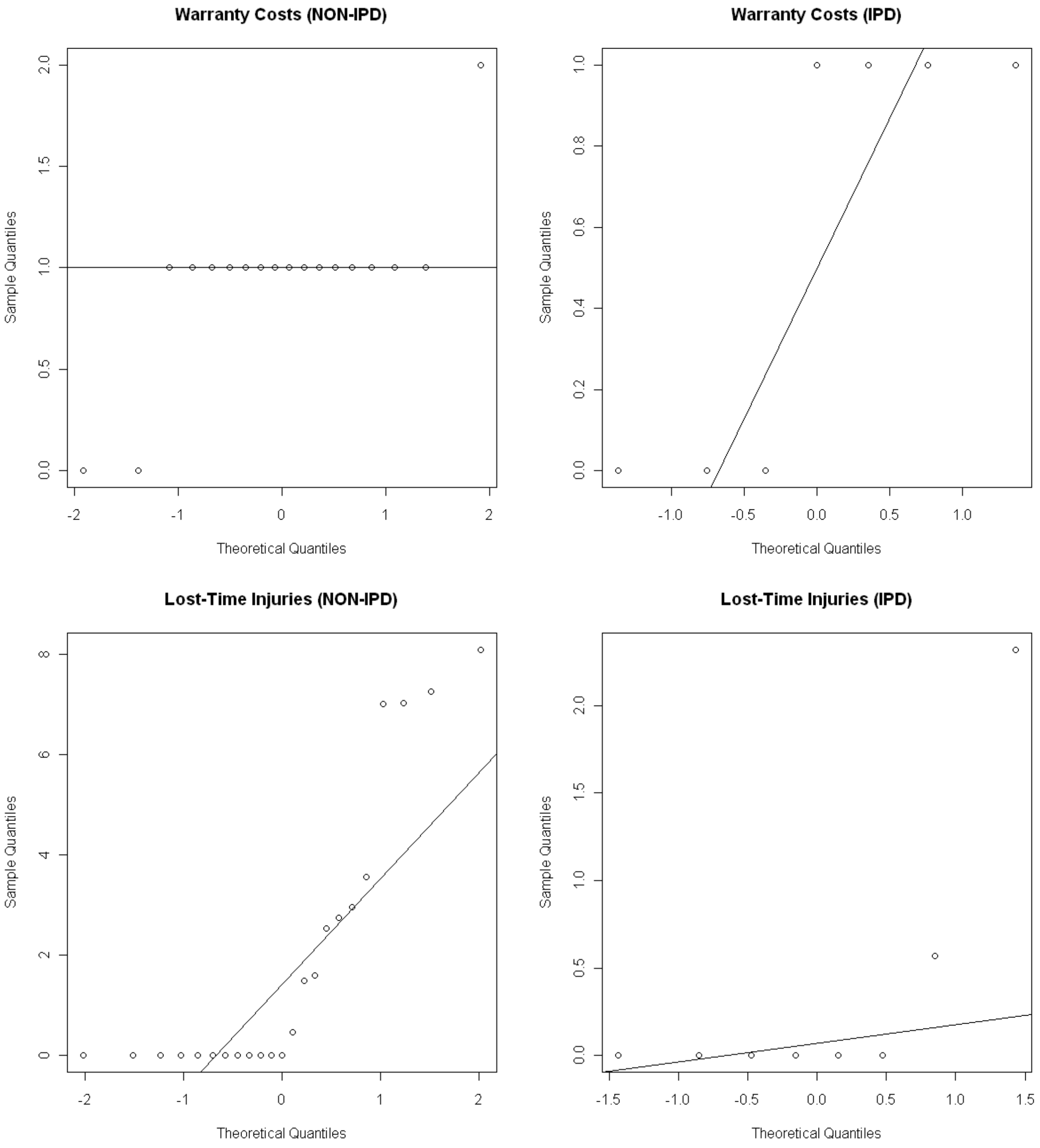


FIGURE 36: Q-Q Plots for Warranty Costs (top) and Lost-Time Injuries (bottom)

Again, none of the plots seem to have points following a straight line, so their data cannot be assumed normally distributed. Therefore, the MWW test can be considered more powerful than the t-test for these two variables. The MWW tests in rows number 11 and 17 of *Table 24* show that the differences in warranty costs performance are statistically significant at the 0.05 level, and the differences in lost-time injuries performance are statistically significant at the 0.10 level.

The opposite scenario can be seen for the construction schedule growth variable, in row 16 of *Table 24*, where the t-test results are favored over the MWW test results. *Figure 37* shows the reason: Q-Q Plots for construction schedule growth data are fairly linear, making possible the assumption of a normally distributed dataset.

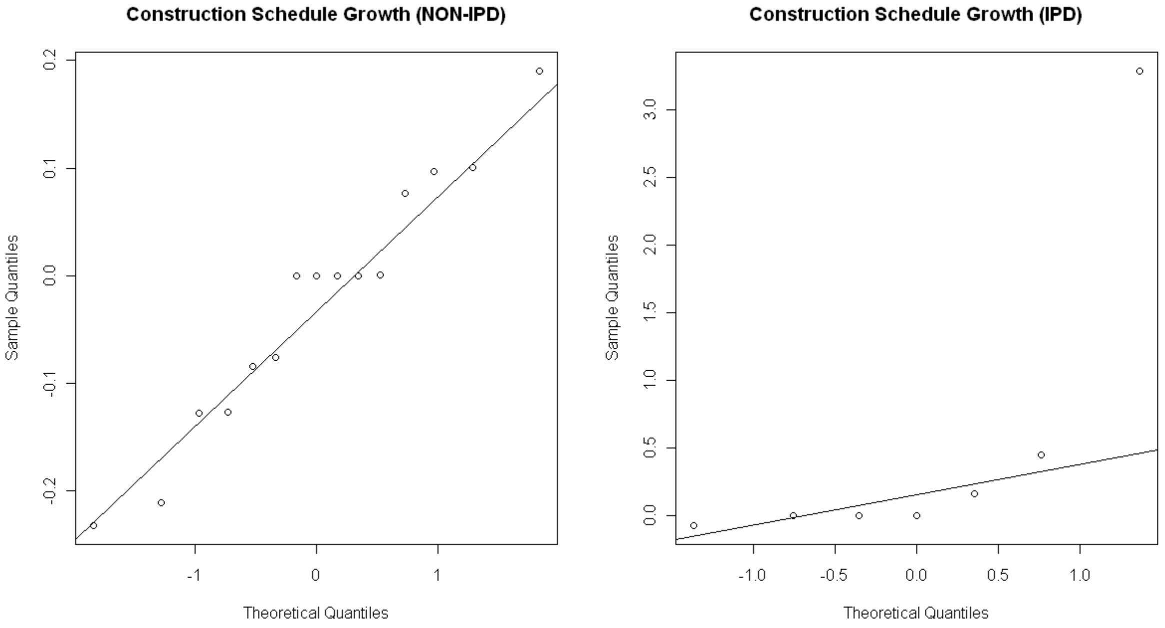


FIGURE 37: Q-Q Plots for Construction Schedule Growth

In this case, the t-test holds more power than the MWW test and its results are considered more rigorous. While the MWW test was showing the performance differences significant at the 0.10 level, the t-test shows a p-value of 0.131 and therefore the differences should not be considered significant.

The last two instances of disagreement between MWW and t-test results occurred for the delivery speed and the material waste variables in rows 13 and 15 of *Table 24*. Here again, the Q-Q Plots shown in *Figure 38* are fairly linear, and therefore the t-test results are favored over the MWW test results. Based on the t-test results, differences for both metrics should be considered significant at the 0.05 level.

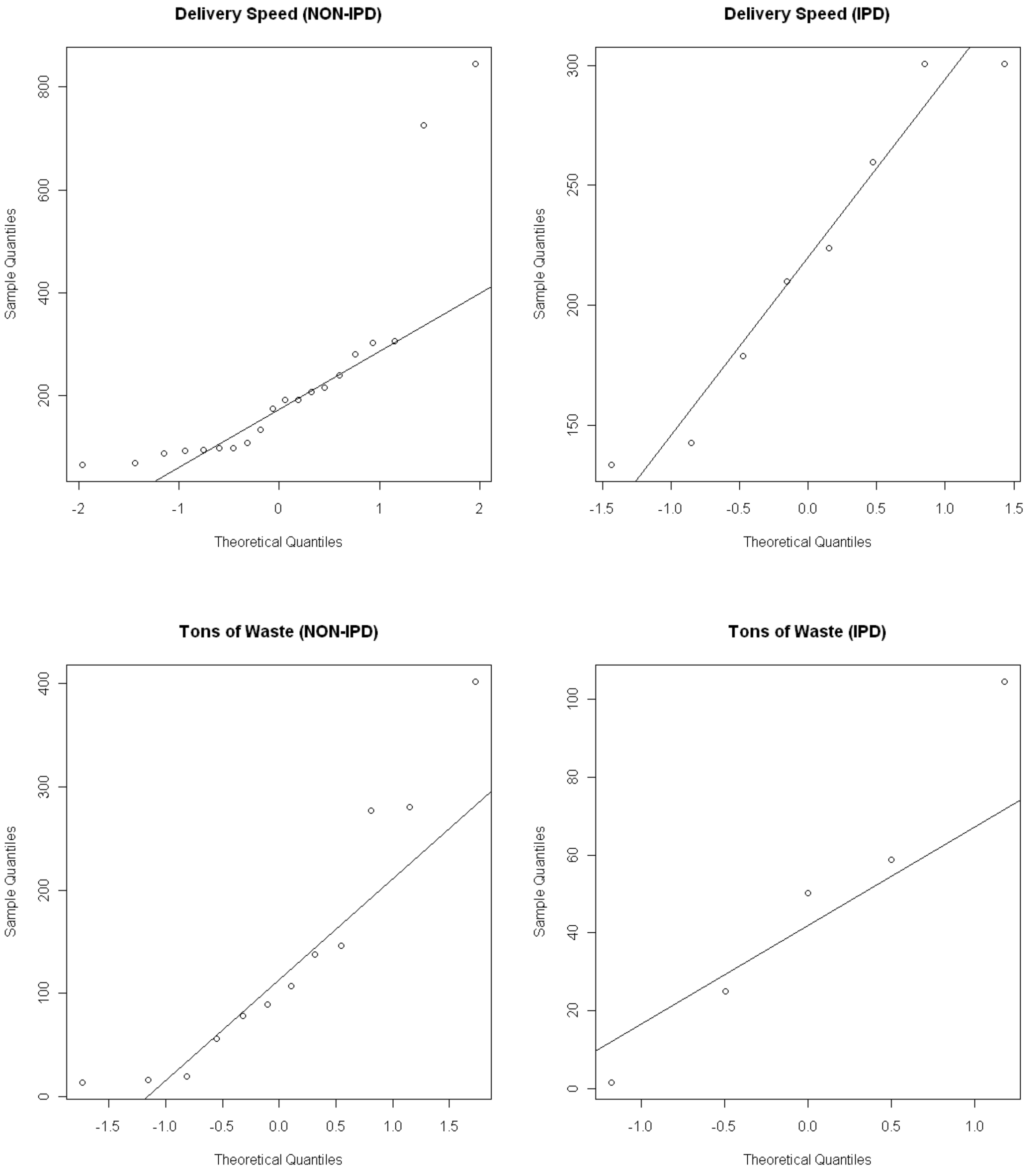


FIGURE 38: Q-Q Plots for Delivery Speed (top) and Tons of Waste (bottom)

In accordance with the previous discussion, the final results are summarized in *Table 25*. Overall, there are 14 performance metrics that showed differences between IPD and non-IPD projects at the 0.05 level. The additional three metrics considered significant at the 0.10 level will not be highlighted as part of the findings to avoid false positives since several metrics are tested in this study. These statistically significant findings help define the performance areas that IPD affects positively, and provide guidance for project stakeholders when choosing a project delivery system.

TABLE 25: Final Univariate Results for IPD vs. Non-IPD (ordered by p-value)

Performance Metric		p-value
1	Change Order Processing Time	0.000
2	Deficiency Issues	0.001
3	Request For Information (RFI)	0.001
4	Punchlist Cost	0.003
Above are results statistically significant at the 0.01 level		
5	Punchlist Items	0.014
6	Resubmittals	0.018
7	Tons of Waste	0.022
8	Overhead & Profit	0.024
9	RFI Processing Time	0.025
10	Systems Quality	0.032
11	Design Changes	0.032
12	Warranty Costs	0.040
13	Delivery Speed	0.046
14	Regulatory Changes	0.050
Above are results statistically significant at the 0.05 level		
15	PPC Trend	0.072
16	Lost-Time-Injuries	0.083
17	Labor Factor	0.094
Above are results statistically significant at the 0.10 level		

4.4 IPD versus: DBB, CMR, DB

In addition to comparing IPD to non-IPD projects as a general group, it is also beneficial to compare IPD to specific project delivery systems separately. The non-IPD projects for which data was collected were delivered using DBB, CMR, and DB. In this subsection, the differend PDS will not be combined; instead they will be treated as separate groups to provide another comparison between IPD and other PDS performance.

For this step, the Kruskal-Wallis test is used. The Kruskal-Wallis test is an extension of the MWW test to three or more groups. It is a nonparametric test equivalent to the one-way analysis of variance (ANOVA), and unlike ANOVA Kruskal-Wallis does not require the dataset to be normally distributed.

The results of the Kruskal-Wallis test were fairly similar to the ones discussed in the previous section. In fact, the performance areas that showed significant differences between the four groups (DBB, CMR, DB, and IPD) are quality, communication, project change, and business metrics. One possible reason why other metrics that were found significant in the earlier analysis were not on this list is because the sample sizes were divided among more treatments for the current analyses, resulting in smaller sample sizes and causing p-values to be larger. *Table 26* shows the metrics for which the differences between the four groups were found statistically significant using the Kruskal-Wallis test.

In addition to tabular results, *Figures 39* and *40* show the boxplots of the metrics presented in the table. *Figure 39* includes the metrics related to quality and business performance, and offers a visualization of the data. It is clear in the boxplots that IPD projects result in a higher systems quality to the benefit of the owner, and the contractors tend to have a superior financial performance as compared to the other delivery systems. Additionally, IPD projects experience a tenfold reduction in deficiency issues as compared to DBB projects, from around two deficiency issues per million dollars to 0.2 for IPD projects. There is also a noticeable decrease in punchlist items, from 41 punchlist items per million dollars for DBB projects to seven items for IPD projects.

TABLE 26: Summary of Univariate Results for IPD vs. 3 Main PDS

Performance Metric	IPD vs. Other PDS Hypotheses	**p-value**
Change Order Processing	IPD projects experience a faster change order processing time than other projects	0.000
RFI /$M	IPD projects experience a smaller number of RFIs than other projects	0.001
Above are MWW results statistically significant at the 0.01 level		
Punchlist Items	IPD projects have less punchlist items than other projects	0.024
OH&P	IPD projects have more OH&P than other projects	0.026
Systems Quality	IPD projects result in a higher quality than others	0.028
Deficiency Issues	IPD projects result in less deficiency items than other projects	0.033
Above are MWW results statistically significant at the 0.05 level		
% Regu. Changes	IPD projects experience less regulatory changes than other projects	0.084
RFI Processing Time	IPD projects experience a faster RFI processing time than other projects	0.095
Above are MWW results statistically significant at the 0.10 level		

Similarly, *Figure 40* includes boxplots for the metrics related to project change and communication performance. Again, IPD results in a tenfold reduction in the number of RFI as compared to DBB projects, from more than 20 RFI per million dollars to less than two RFI. The RFI processing time is also reduced from 2.5 weeks for DBB to a single week for IPD. Among the different project change variables, changes due to code and regulations are significantly lower for IPD. And more importantly, the change order processing time for all types of changes is reduced from seven weeks for DBB to a single week for IPD. This improvement alone can smooth the process between the different stakeholders and result in a superior project performance.

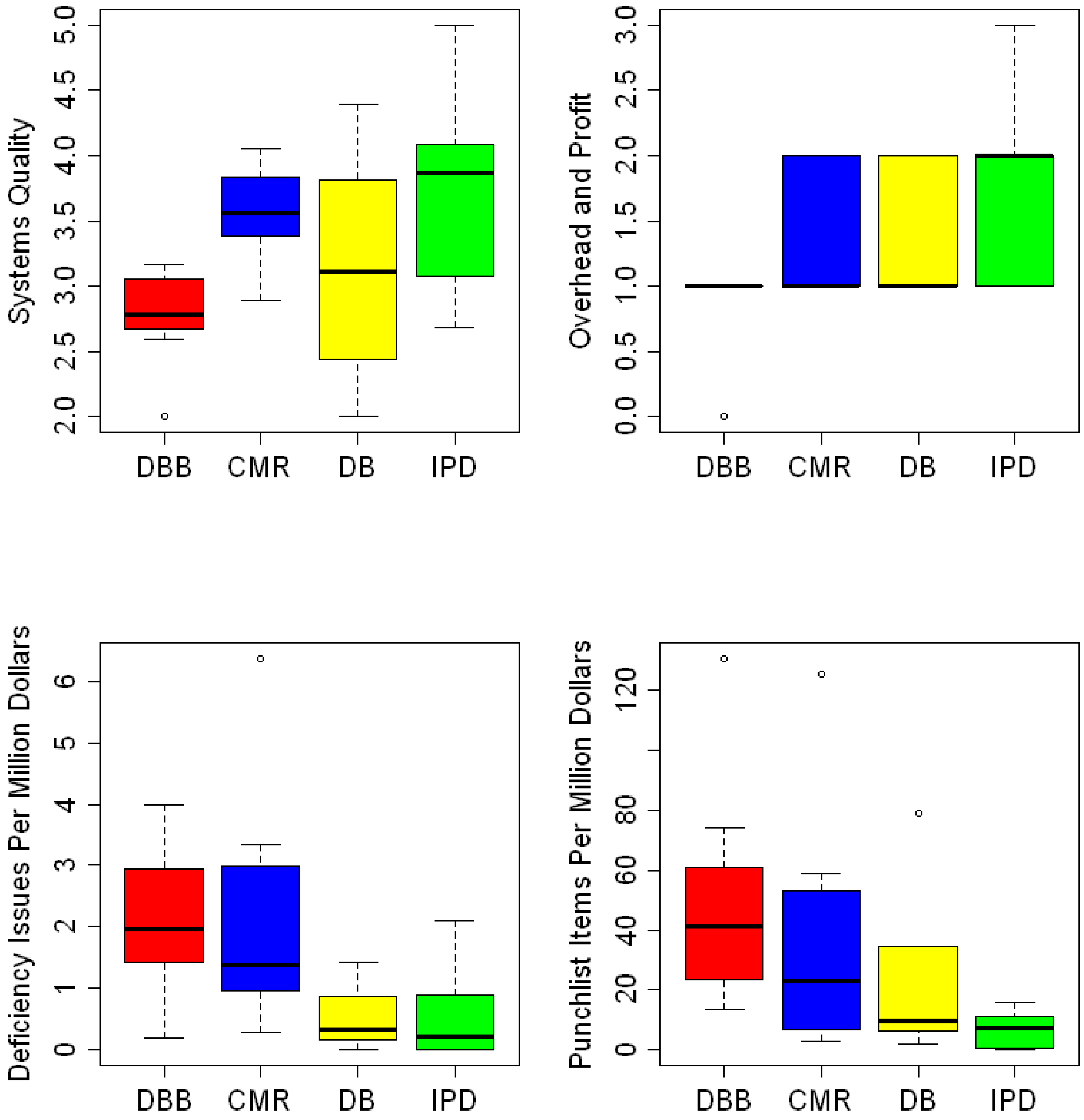


FIGURE 39: Boxplots for PDS Quality and Business Performance

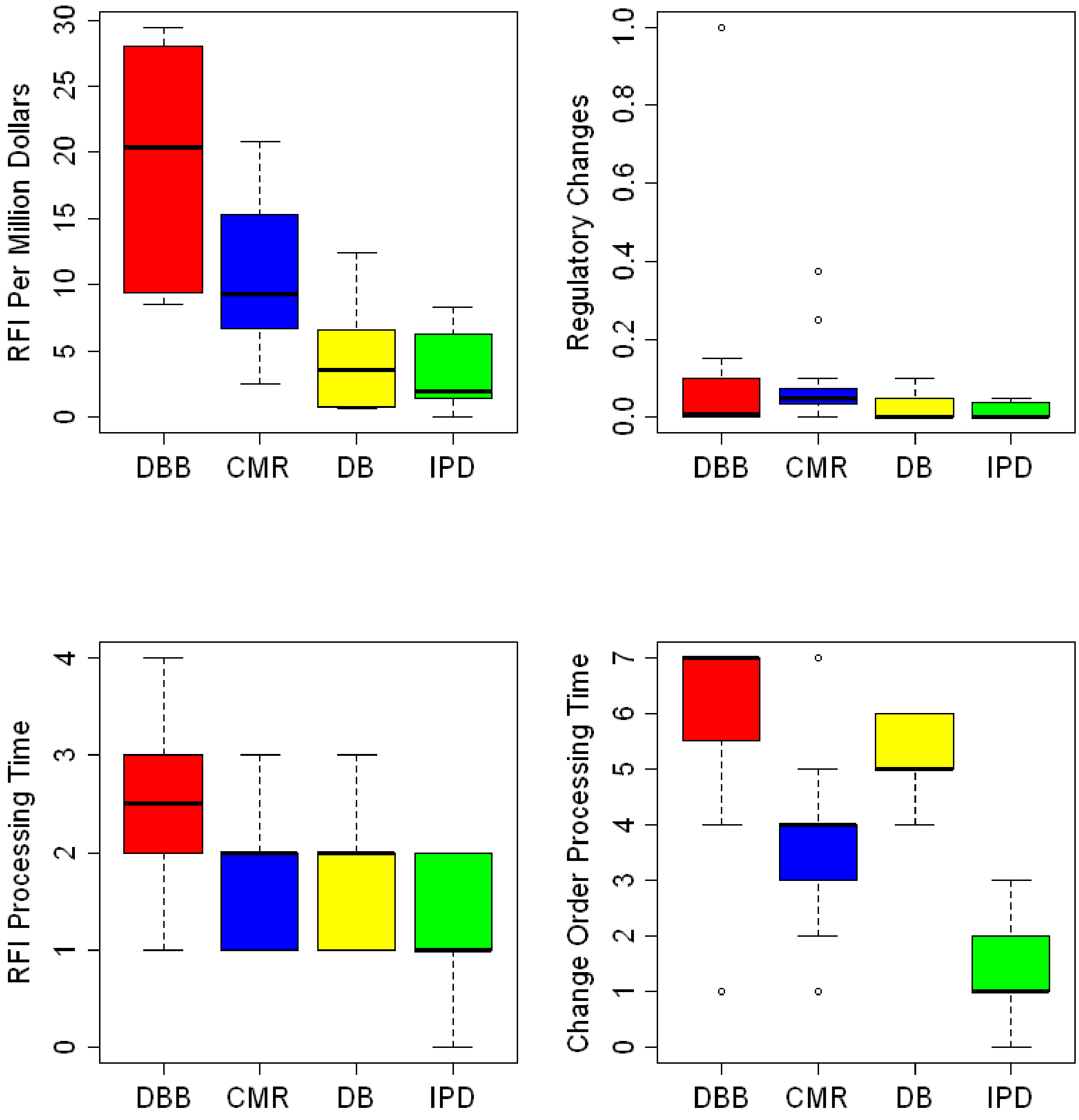


FIGURE 40: Boxplots for PDS Project Change and Communication Performance

While the previous sections compared IPD to non-IPD projects as a whole group, this section investigated the non-IPD projects more closely. The analysis showed results similar to those of the previous analysis in terms of IPD superiority. Moreover, the findings seem to discern a trend among both IPD and non-IPD projects, showing that more integrated delivery systems are experiencing improved performance as compared to less integrated systems. DBB was generally found to be the least performing system, and IPD the highest performing system, while CMR and DBB were typically between those two extremes.

4.5 Conclusions of the Univariate Analysis

In this chapter, several performance metrics were analyzed individually. This concluding section will present a summary of the results and recapitulate the contributions of *Chapter 4*. *Table 25* showed the various performance areas for which the differences between IPD and non-IPD projects were found to be statistically significant. The p-values shown are for the test most appropriate (MWW test or t-test) for each performance metric depending on its data distribution.

Using a very strict threshold where p-values need to be less than 0.01, IPD was proven to have a superior performance in metrics related to quality, communication, and change performance. The quality of the facility is arguably the most important metric that IPD enhances. IPD projects also see less changes and faster processing times.

The number of performance metrics for which IPD projects were significantly different increases to fourteen at the 0.05 level, adding business metrics, recycling metrics, and most notably schedule metrics. IPD projects are experiencing significantly faster delivery times, and contractors are seeing a higher profit and return business with IPD.

The univariate analysis allowed for a major contribution to the project delivery systems literature. It provided the first comprehensive statistical comparison of IPD projects to similar non-IPD projects. One key conclusion here is that IPD delivers higher quality projects, faster, and at no significant cost premium. In fact, one can argue that customers often spend their whole budget and reinvest the savings into the project to get more value out of their facility, resulting in no visible cost savings but major project benefits.

Although the first few cost performance results seemed to confirm findings of a previous study that shows no cost performance differences between IPD and non-IPD projects, comprehensively looking at the remaining performance metrics strongly contradicts the previous literature. Not only does IPD provide schedule and quality improvements, it also offers enhancements on several other key performance metrics.

Finally, there are no known benchmarks for IPD projects. Benchmarks are defined as standard values that can be used as reference point for future projects. The IPD performance data that was collected for this research presents an opportunity to provide the first set of IPD benchmarks. *Figures 39* and *40* presented illustrations of some of the first IPD metrics analyzed in this research. For example, when working on an IPD project, one should expect to see around seven punchlist items and 0.2 deficiency issues per million dollars. Additionally, IPD projects experience less than two RFI per million dollars; and a remarkable one week processing time for both change orders and RFI, which can result in much smoother project as compared to non-IPD projects. These are just a few examples; the results for all the remaining metrics also can be used as IPD benchmarks, which will assist industry professionals gauge their project performance when implementing IPD.

This chapter provided the groundbreaking proof that IPD projects offer an improved performance when compared to non-IPD projects based on 14 metrics. A specific question naturally follows this discovery: since IPD projects offer superior performance based on quality metrics, but offer no differences when it comes to project cost, would it be accurate to say that IPD is a superior delivery system?

Chapter 5 will answer this question. By combining the key metrics into one comprehensive performance metric, *Chapter 5* will present the *Project Quarterback Rating* (PQR). Similar to the quarterback rating in the National Football League, the PQR will be used to measure overall performance for AEC projects, and will be used to present a comparative performance comparison of IPD and non-IPD projects.

Chapter 5. The Project Quarterback Rating (PQR)

Chapter 5 discusses the development, validation, and implementation of the *Project Quarterback Rating* (PQR). The PQR is a comprehensive measure of performance for AEC projects and can be used to gauge overall project performance.

5.1 PQR Motivation

Performance metrics utilized by the sports industry provide benchmarks to compare teams and athletes. For example, the quarterback passer's rating used in the National Football League (NFL), which compares the performance of quarterbacks, is calculated by combining four values for each quarterback: completion percentage, passing yardage, touchdowns, and interceptions (NFL 2012). This combination leads to a single number which then can be used to compare against the ratings of other quarterbacks.

The AEC industry, although four times larger than the sports industry (CSU Fresno 2012), does not have overall performance ratings that are as established. AEC projects are unique, highly complex, and normally require several performance dimensions to achieve project success. As is often the case, it is difficult to determine whether a given AEC project should be considered a success assuming it is completed on time and on budget, but the building systems suffered poor quality and the project experienced major safety issues. Project performance is a complex concept and involves several metrics or outputs that need to be accounted for, some of which do not have standards or cannot be measured easily.

Similar to the quarterback rating in the sports industry, the *Project Quarterback Rating* (PQR) combines the key performance metrics of a project into one number that can facilitate the comparison of projects and provide a basis for quantifying project success. The performance of a project often cannot be measured by one performance metric, such as cost, schedule, or safety. It often involves additional metrics, such as customer relations, profit and return business, and the quality of the end-product or facility. These metrics and others need to be combined into one value that represents the performance of a given project.

To increase the potential for the implementation of this model in the construction industry, the PQR will be developed as a linear function that acts as a weighted average of the several key performance metrics on AEC projects. Use of the model can provide a baseline on which projects can be compared, as well as a venue to discuss of project success.

For this specific study, PQR is used to assist in comparing IPD projects to projects completed using other delivery systems. This comparison has been completed in *Chapter 4* based on individual key performance metrics. However, PQR allows for the combination of all the identified performance metrics and the comparison of a single new metric that illustrates overall project performance.

This chapter introduces a general model that can be applied to AEC projects, regardless of the delivery system used. The project manager or leadership team of a project can assess several performance metrics, which this model combines together to compute the overall performance score or success rating of the project. Like the NFL passer’s rating, the

PQR model can serve as a tool that different stakeholders can use to analyze their projects' performance levels.

5.2 *Mathematical Formulation and Normalization Technique*

The PQR model approach combines seven components, or performance areas, identified by the survey respondents for this research into one comparable score for each project. These seven areas are (1) customer relations, (2) safety, (3) schedule, (4) cost, (5) quality, (6) profit, and (7) communication.

The model computes to each project j a corresponding Project Quarterback Rating PQR_j. The PQR score is based on the seven evaluation criteria with different weights for each. To describe the model that calculates PQR_j, the following notation is used:

$PQR_j = \sum_{i=1}^{I} w_i \, s_{ij}$,

where: j denotes the project number, $1 \leq j \leq J$, $J = 35$,

w_i is the weight of performance area i, $1 \leq i \leq I$, $I = 7$, and

s_{ij} is the composite score of project j for performance area i.

The rationale behind using a linear model lies in its simplicity and the fact that it allows for the addition of several performance metrics. The underlying assumption here is that an overall comprehensive project performance rating PQR_j exists and only depends on the performance areas i. In this model, the performance score PQR_j is calculated as the weighted average of the different performance areas s_{ij}. The performance area scores are combined and normalized before making their way to the weighted average formula introduced above. Moreover, these scores s_{ij} for each of the seven areas also combine many components; for instance, project cost combines the final unit cost and the cost growth as compared to the initial cost estimates. The term X_{ijk} denotes the original scores of performance metrics, with $1 \leq k \leq K_i$ and K_i representing the number of metrics combined in each performance area i. For example, if i represents the cost performance area, then X_{ijk} would represent unit cost and cost growth ($K_i = 2$), which will be combined into s_{ij} that represents the cost performance area, as shown in *Figure 41*.

Figure 41 exhibits three levels: (1) PQR_j is the top level; (2) s_{ij} is the second level that includes the seven performance areas that PQR combines; and (3) X_{ijk} represents all the individual performance metrics listed under each of the seven areas.

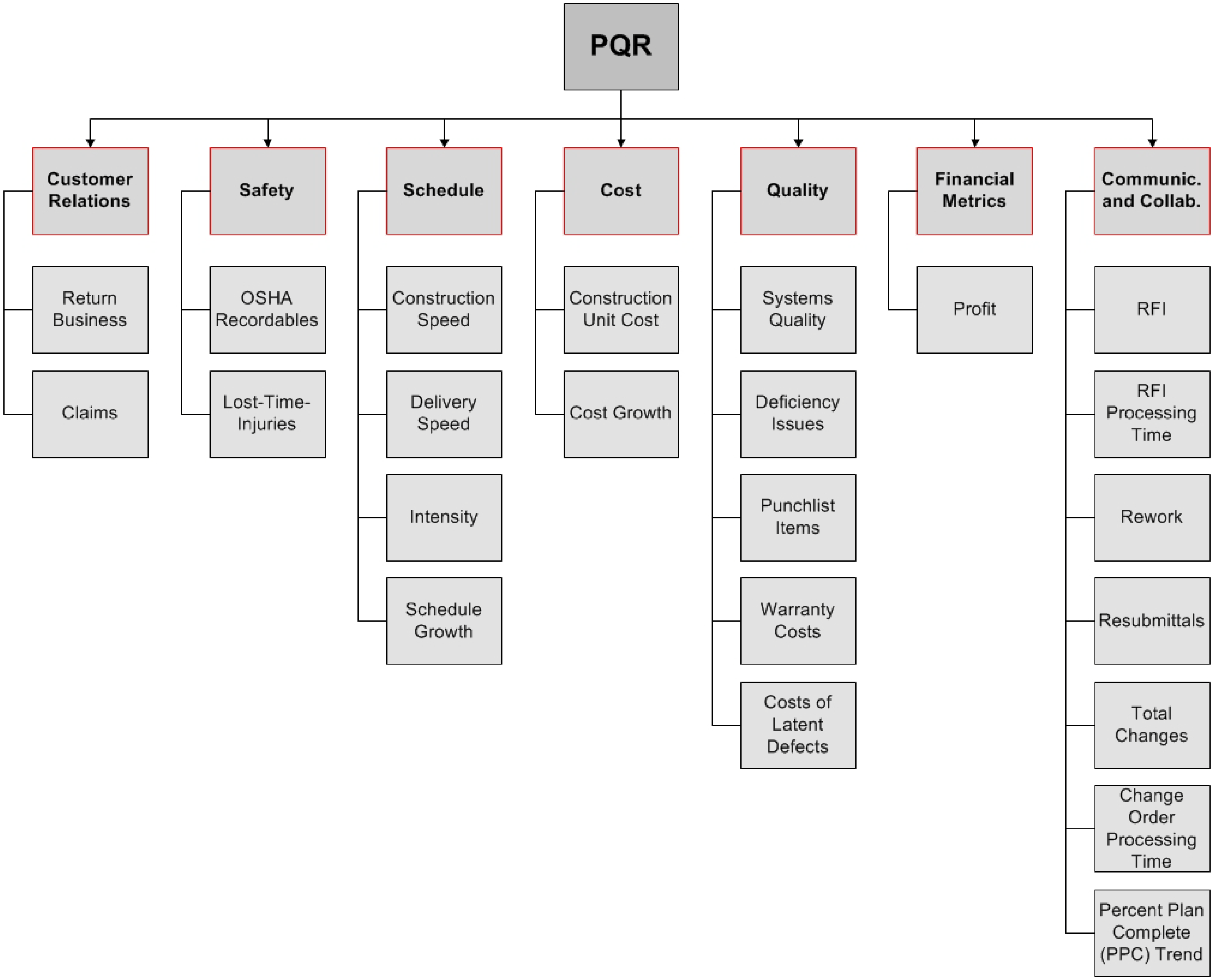


FIGURE 41: PQR Structure

Looking at all these metrics, a problematic issue becomes evident: one cannot add together values for different metrics, such as cost and time. Even in the same performance area, different cost metrics cannot be added (e.g. unit cost in dollars per squarefoot and cost growth in percentage terms). Standardization is the answer to this hurdle and can transform any set of numbers to their equivalent values on the standard normal distribution.

Standardization achieved by subtracting the mean and dividing by the standard deviation, has several advantages. First, positive values indicate above average performance and negative values below average performance, while zero indicates the average project performance regarding the specific metric at hand. This same interpretation also applies to the combined performance area. Second, the combined rating values can be similarly standardized and interpreted in terms of number of standard deviations below or above the average. Basically, the measurement units of the standardized values are the number of standard deviations above or below the average.

To facilitate the reading of the formulas, the following notations are used: the operation $Av_{ik}\left(X_{ijk}\right)=\frac{1}{J}\sum_{j=1}^{J}X_{ijk}$ denotes the average of the array X_{ijk} over the index j, fixing the other indices i and k. The normalization procedure maps each original score X_{ijk} to a normalized version z_{ijk} as follows: fixing a criterion k and a performance area i, the mean score and standard deviation are calculated for each performance metric. These are:

$$\mu_{ik}=Av_{ik}\left(X_{ijk}\right) \qquad \text{and} \qquad \sigma_{ik}=\sqrt{Av_{ik}\left(\left(X_{ijk}-\mu_{ik}\right)^{2}\right)}.$$

The z-scores are then calculated to be $z_{ijk} = \left(X_{ijk} - \mu_{ik}\right)/\sigma_{ik}$. These z-scores hide any variation of the ranges and units of different metrics, giving them equal effect on s_{ij}. The scores for each performance area can be calculated using the equation:

$$s_{ij} = \sum_{k=1}^{K_i} w_{ik}\, z_{ijk} \ .$$

Here w_{ik} denotes the weight of each metric in a specific performance area i. The z-scores z_{ijk} are centered around zero; some values are positive and some are negative. After being used in the weighted average, a similar normalization technique is used to standardize the resulting scores s_{ij}. Since the mean of all s_{ij} is zero, each s_{ij} is directly divided by the standard deviation of all s_{ij}, similar to what was completed previously for the original performance metrics scores. The resulting z-scores of s_{ij} will ensure a straightforward interpretation of the results, similar to what was discussed earlier: a negative score means the project had a lower-than-average performance in the specific area at hand (e.g. cost or safety) and a positive score means the project had an above-average performance, while a score of zero means the project had an average performance. Furthermore, the values above or below average again can be interpreted as numbers of standard deviations relative to the average. For instance, a score of 1.5 means the project was 1.5 standard deviations above the average project performance for the specific performance area.

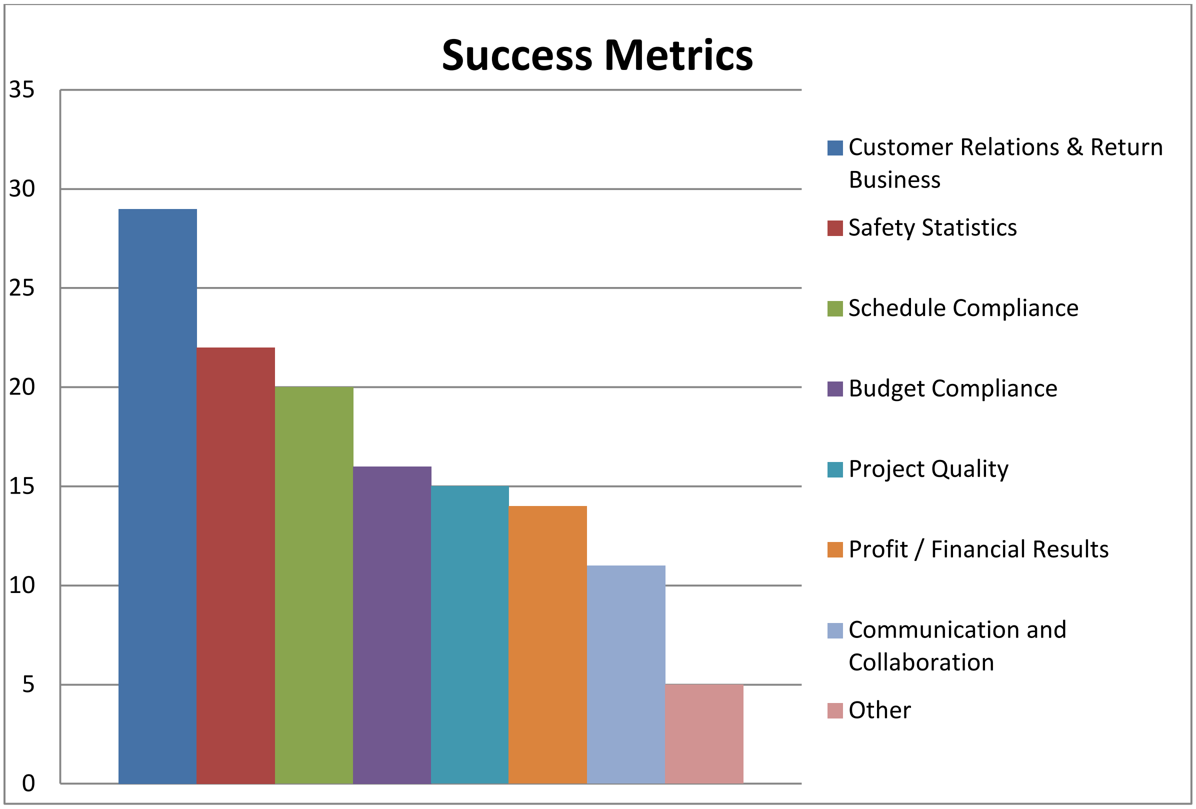
After the transformations, the standardized scores for the seven performance areas then are combined into the PQR formula, $PQR_j = \sum_{i=1}^{I} w_i\, s_{ij}$. The resulting scores undergo one last standardization procedure to warrant the interpretation presented above. The next subsection will populate this mathematical formulation with actual project data and transform it into a function that can be used to assess overall project performance.

5.3 The PQR Formula

Before using project data to develop the PQR model, the weights for each of the seven performance areas must be identified. The weights quantify the level of importance for individual performance areas. One example is whether the cost performance of a project is more important than its safety or schedule performance, and if so by how much. The last section of the data collection survey shown in *Appendix C* was designed specifically for this purpose, and prompted respondents for information regarding the performance metrics their respective companies consider when evaluating project success. The different respondents for each project were asked to identify the five most important metrics when determining project success. There were 34 respondents that addressed this question. The responses were grouped in related performance areas (e.g. safety, cost, and quality), for which the frequencies were calculated and presented in the top part of *Figure 42*. Frequencies then can be converted to percentages by dividing the frequency value for each item by the total sum of all item frequencies: 127. These percentages, which sum up to 100%, are shown at the bottom of *Figure 42*.

Figure 42 shows a total of seven performance areas identified. The metric related to customer relations and return business had the largest frequency, sub-items of which were stated in 29 out of the 34 total responses received for this section of the survey. This finding was especially interesting because it translates into 85% of the respondents for all delivery systems combined, while only 75% of the respondents for IPD projects listed customer relations as a major constributor to project success.

Safety comes in second and schedule performance a close third. The project being on budget was in fourth place. Project quality received the fifth place in the top seven project success metrics, and financial profit was ranked sixth. Lastly, communication and collaboration metrics secured the final spot. If one were to assign weights to these top seven metrics based on their frequencies, the aggregated performance factor PQR would be a linear weighted sum of the performance areas using the weights shown in *Figure 42*.



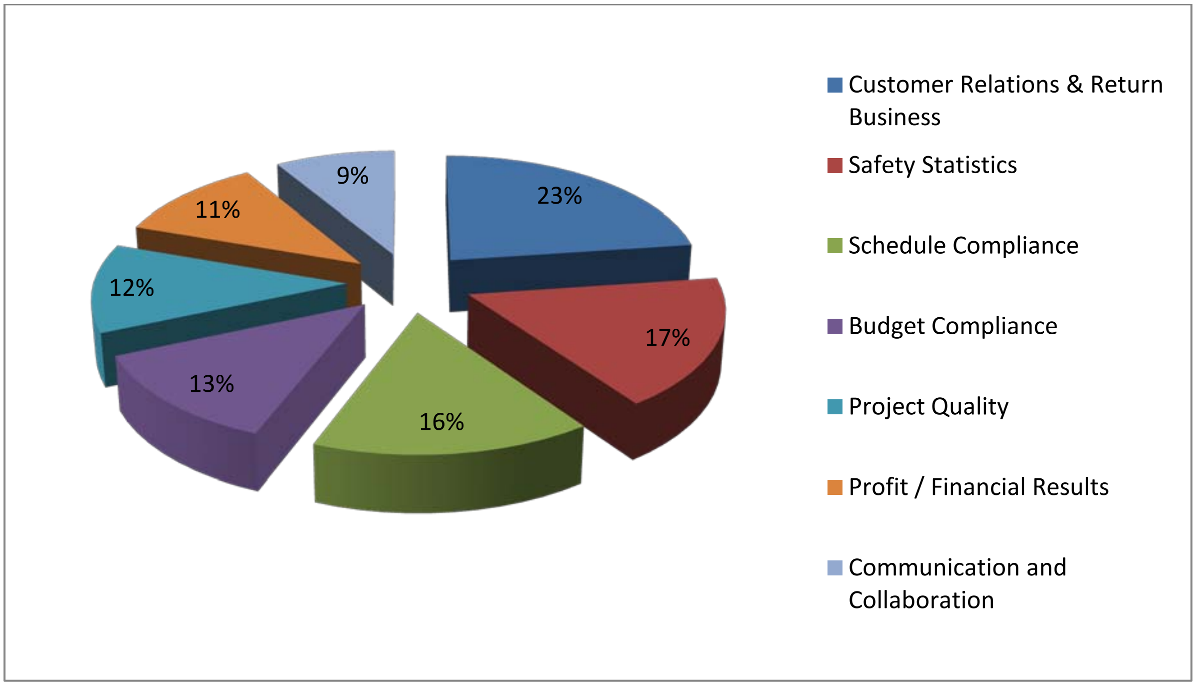


FIGURE 42: Frequencies (top) and Percentages (Bottom) of Performance Areas

These percentages can be used in the PQR formula as follows (*Eq.0*). However, users can adjust the percentages based on specific needs and success factors of their own project.

$$\mathrm{PQR} = \frac{\begin{array}{c} 0.23 * \text{Relations} + 0.17 * \text{Safety} + 0.16 * \text{Schedule} + 0.13 * \text{Budget} \\ +0.12 * \text{Quality} + 0.11 * \text{Profit} + 0.09 * \text{Communication} \end{array}}{0.51} \qquad (Eq.0)$$

The 0.51 value in the denominator is the standard deviation of all project scores, as discussed in the previous section, and ensures the standardization of the PQR function. The mean value for each projects is also subtracted from the nominator of $(Eq.0)$; however, since these terms were already standardized, the mean value is zero. In order to make the above formula simpler to apply, the percentages were divided by the standard deviation upfront, which resulted in new coefficients. These coefficients can be seen in the updated PQR formula as follows (*Eq.1*).

$$\boxed{\mathrm{PQR} = \begin{array}{c} 0.45 * \text{Relations} + 0.34 * \text{Safety} + 0.31 * \text{Schedule} + 0.25 * \text{Budget} \\ +0.23 * \text{Quality} + 0.22 * \text{Profit} + 0.17 * \text{Communication} \end{array} \qquad (Eq.1)}$$

In order to facilitate the use of the PQR formulas, *Appendix G* presents a summary of the equations developed in this chapter. The appendix also includes a numerical example which consists of initial inputs from a sample project, along with their corresponding performance area scores and the final PQR value.

The scores for each of the identified performance areas also need to be standardized individually in order to be used in the PQR formula. Several areas include more than one metric (e.g. OSHA recordables and LTI for safety performance), and therefore these metrics need to be aggregated. A score will be computed for each of the performance areas present in the PQR equation, which will be plugged in to the PQR equation (*Eq.1*) to obtain an overall performance rating for each project.

A user can apply this exact formula to his or her project, and the project's PQR can be interpreted relative to the 35-project database compiled for this study. That is, if the PQR is positive, then the project is above the average performance of the projects in this study. However, other users might have their own large dataset of projects and might want to create their own formula based on their respective dataset. In that case, the users can employ the same development technique discussed here, and the resulting formula could be different depending on their set of projects.

After discussing how the seven performance areas can be linearly combined into the PQR, it is appropriate to discuss how the scores can be computed for each of the seven areas. The following subsections are ordered starting with the top performance area identified by the survey respondents: customer relations.

5.3.1 PQR Part 1/7: Customer Relations

The first, and arguably most important, performance area is customer relations. This is an area that is especially difficult to measure in the AEC industry because it does not have a standard quantifiable metric. For the purpose of the PQR development, two metrics are used as proxies for quantifying customer relations related to a specific project. The first metric is the potential of this project to result in return business. This is the most important customer relations performance metric since one can safely assume that an owner would not hire the same company again if he or she were not satisfied with the company's work product. The second metric is the existence of legal claims between project parties, which consume time and resources from the project stakeholders and do not add value to the customer’s project.

Because of its high level of importance for assessing customer relations, the return business metric was given a weight of 75%, whereas the claims metric was only given the remaining 25% weight. The potential for return business was evaluated on a five-point scale, from very negative (coded to -2), to negative (-1), neutral (0), positive (1), and very positive (2). The claims metric is a binary variable that can only have two values: one for no claims, or zero when claims existed between the project parties. The scoring of the claims variable might be seen as counterintuitive, but it allows for the simple addition of the two metrics. It is important that the ordinal scale of -2 to 2 for return business, and the binary scale of 0 or 1 for claims, be used when applying this equation to other projects to make sure the end result is interpretable.

The mean value for return business was 1.66, and the standard deviation was 0.87. The range for the initial variables extended from -2 to 2, as discussed earlier. The mean for the binary claims variable was 0.91, and the standard deviation was 0.29. Each of the two metrics was standardized individually to obtain z-scores. After this initial standardization, a weighted average of the two new z-scores was computed, and the result was standardized again, hence the 0.87 denominator. The formula for customer relations is devised as follows:

$$\text{Relations} = \frac{0.75 \times \frac{\text{ReturnBus} - 1.66}{0.87} + 0.25 \times \frac{\text{Claims} - 0.91}{0.29}}{0.87}$$

The above equation can be used to calculate customer relations performance scores for AEC projects. For a given project, the resulting score can be used in the first part of the PQR equation (*Eq. 1*) presented earlier in *Section 5.3*. Similar equations will be developed in the following subsections for each of the six remaining performance areas.

5.3.2 PQR Part 2/7: Project Safety Statistics

Safety statistics emerged as the second most important performance area in the survey responses. The existence of Occupational Safety and Health Administration (OSHA) requirements to track safety performance allows this to be a more quantifiable performance area than the previously discussed customer relations. Two performance metrics are combined into the safety performance score: OSHA recordables and lost-time injuries (LTI). Both metrics are normalized per million dollars of construction work.

The mean for OSHA recordables per million dollars was 5.38 and the standard deviation was 5.59; while the range for the initial variables extended from zero to 22.77 recordables. The mean for LTI per million dollars was 1.36 and the standard deviation was 2.41; while the range for the initial variables extended from zero to 8.09.

Each of the two metrics was standardized individually, resulting in two z-scores for each project. The LTI metric was given a weight of 75% and the recordables metric a weight of 25% because LTI is a measure of cases that resulted in lost work days, which are typically more severe than the average recordable injuries and illnesses. Then a weighted average of the two new z-scores is computed, and the result is standardized again. The formula for safety performance is devised as follows:

$$\text{Safety} = \frac{-0.25 \times \dfrac{\text{Recordables} - 5.38}{5.59} - 0.75 \times \dfrac{\text{LTI} - 1.36}{2.41}}{0.87}$$

The negative signs before the two members of the numerator stem from safety statistics being a number that should be minimized, as opposed to maximized as was the case with customer relations. The negative sign will invert the scale and allow for the combination of safety and customer relations metrics by a simple addition. The interpretation will remain in accordance with the previous discussion: positive numbers denote above average performance in units of standard deviations. Unlike the values for the customer relations area, the safety values are actual safety statistics (not coded), and therefore there are no restrictions to use values outside the range of the data collected to build this model. These values are standardized to z-scores, and the normal distribution of the z-scores spreads from minus infinity to infinity, allowing for values outside the original range to be used.

The above equation can be used to calculate an AEC project's score with respect to its safety performance. The resulting score can be used in the second part of the PQR equation (*Eq. 1*) presented earlier in *Section 5.3*. The next subsection deals with schedule performance.

5.3.3 PQR Part 3/7: Project Schedule

The performance area that was ranked third by the survey respondents in terms of overall importance for project success is schedule compliance. Similar to safety statistics, key quantitative schedule metrics can be easily calculated because most project teams keep track of important project dates, such as the construction notice to proceed and substantial completion. Four performance metrics are combined into the schedule performance score: (1)

construction speed in square feet per day, (2) delivery speed also in square feet per day, (3) construction intensity in dollars per day, and (4) construction schedule growth in percentage terms between the initial schedule estimate and the actual schedule.

The mean for construction speed was 331.50 square feet per day and the standard deviation was 259.40; while the range for the initial variables extended from 63.69 to 1348.68. The mean for delivery speed was 210.09 square feet per day and the standard deviation was 169.85; while the range for the initial variables extended from 55.76 to 645.36. The mean for construction intensity was 102,794.94 dollars per day and the standard deviation was 77,658.17; while the range for the initial variables extended from 22,052.34 to 395,124.00. The mean for schedule growth was 12.50 percent and the standard deviation was 66.16; while the range for the initial variables extended from -23.23 to a staggering 329.35 percent.

Each of the four metrics was standardized individually, and each was given a weight of 25%. These weights are consistent with the results of PCA conducted in *Section 4.2.3* for schedule metrics. A weighted average of the four new z-scores is computed, and the result is standardized again. The 2.86 value in the denominator equals the standard deviation 0.72 multiplied by four, since each member of the numerator accounts for one fourth of the total schedule score. The formula for schedule performance is:

$$\text{Schedule} = \frac{\frac{\text{C.S} - 331.50}{259.40} + \frac{\text{D.S} - 210.09}{169.85} + \frac{\text{Intensity} - 102{,}794.94}{77{,}658.17} - \frac{\text{Growth} - 0.125}{0.6616}}{2.86}$$

The formula can be devised with the value in the denominator representing the 25% weight of each metric multiplied by the standard deviation of their weighted average. As discussed earlier for safety statistics, the negative sign before the construction growth value in the numerator will adjust the negative connotation of schedule growth and allow for its z-score to be added to metrics that need to be maximized, like speed.

Similar to the safety equation, there are no restrictions to use schedule values outside the range of the data collected to build this model. This equation can be used to calculate a project's score with respect to its schedule performance, and the resulting score can be used in the third part of the PQR equation (*Eq. 1*).

5.3.4 PQR Part 4/7: Project Cost

Surprisingly, project cost performance was only ranked fourth in importance when assessing overall project success from the contractors' perspective. Similar to safety and schedule, construction cost performance is relatively easy to track because project teams keep useful records of several cost items for different project phases. Two performance metrics are combined into the cost performance score: construction unit cost in dollars per square foot, and construction cost growth from initial estimates to actual costs.

The mean for unit cost was 353.34 dollars per square foot and the standard deviation was 144.67; while the range for the initial variables extended from 176.95 to 788.95. The mean for cost growth was 4.85 percent and the standard deviation was 10.69 percent; while the range for the initial variables extended from -19.57 to 37.72 percent. Each of the two metrics was standardized individually, and each was given a weight of 50%. A weighted average of the two new z-scores was computed, and the result was standardized again. The formula for cost performance is as follows:

$$\mathrm{Cost} = \frac{-0.5 \times \frac{\mathrm{ConstructionUnitCost} - 353.34}{144.67} - 0.5 \times \frac{\mathrm{CostGrowth} - 0.0485}{0.1069}}{0.74}$$

The above equation can be used to calculate a project's score with respect to its cost performance, and there are no restrictions to use cost values outside the range of the data collected to build this model. The resulting score can be used in the fourth part of the PQR equation (*Eq. 1*) presented earlier.

5.3.5 PQR Part 5/7: Project Quality

The performance area following cost performance in terms of overall importance for project success is project quality. Five performance metrics are combined into the quality performance score, some quantitative and others more qualitative in nature but were quantified using ordinal scales. The five project quality performance metrics used are: (1) systems quality on a scale of one to five, as explained in *Chapter 3*, (2) the number of deficiency issues per million dollars, (3) the number of punchlist items per million dollars, (4) the warranty costs, and (5) the latent defect costs. The warranty latent defect costs are both measured on an ordinal scale based on cost percentages relative to total construction costs. For example, if the warranty costs are 0% of the construction cost, the value is coded to zero; however, if they equal between zero and 0.5% of construction costs, the value is coded to one, and 0.6% to 1% is coded to 2. The same coding applies to the cost of latent defects.

The mean for systems quality was 3.40 and the standard deviation was 0.70; while the range for the initial variables extended from two to five. The mean for deficiency issues per million dollars was 1.32 issues and the standard deviation was 1.56; while the range for the initial variables extended from zero to 6.37 issues. The mean for punchlist items per million dollars was 29.48 items and the standard deviation was 33.91; while the range for the initial variables extended from zero to 130.69 items. The mean for warranty costs, on the scale discussed above, was 0.79 and the standard deviation was 0.49; while the range for the initial variables extended from zero to two. Finally, the mean for the cost of latent defects, also on the scale discussed above, was 0.54 and the standard deviation was 0.51; while the range for the initial variables was limited to zero and one.

Each of the five metrics was standardized individually. A weighted average of the five new z-scores was computed, and the result was standardized again. The formula for quality performance is:

$$\mathrm{Quality} = \frac{\frac{\mathrm{Syst.} - 3.4}{1.17} - \frac{\mathrm{Defic.} - 1.32}{15.6} - \frac{\mathrm{Punch.} - 29.48}{339.1} - \frac{\mathrm{Warran.} - 0.79}{4.9} - \frac{\mathrm{Latent} - 0.54}{5.1}}{0.69}$$

In this equation, the systems quality was weighed 60% of the quality score and the remaining metrics weighed 10% each. For simplicity and readability, the weights and standard deviations are compiled together in the denominator of each metric. The equation can be used to calculate a score that represents a project’s quality performance. The resulting score can be used in the fifth part of the PQR equation (*Eq.1*) presented in the beginning of *Section 5.3*.

With respect to the values that can be used when applying this model, there are no restrictions to use values outside the ranges for the number of deficiency issues per million dollars and the number of punchlist items per million dollars. However, it is important that the scoring of the systems quality, warranty costs, and latent defects variables is performed in accordance with the ordinal scales discussed in the first paragraph of this subsection. Respecting these scales ensures the end result is interpretable when applying this equation to other projects.

5.3.6 PQR Part 6/7: Financial Metrics

Another important performance area is the financial profit obtained on a given project. After all, contractors are in the construction business to make profit, which often is the main motivation for companies. This is an area that is difficult to measure in most industries, especially in the AEC industry where profit margins are low and where competitive bidding is still the norm in a major portion of the industry. Since most companies would not reveal their net profit on their projects, this metric had to be disguised by including job overhead. Therefore, the only metric used here is job overhead and profit (OH&P).

OH&P as a percentage of project cost was evaluated on an ordinal five-point scale, from negative OH&P (coded to zero), to less than 5% (coded to 1), between 5% and 10% (coded to 2), between 11% and 15% (coded to 3), and more than 15% which was never attained. It is important that the same range of zero to three be used when applying this equation to other projects to make sure the result is interpretable.

The mean for OH&P was 1.45 (on the scale of 0 to 3 discussed earlier), and the standard deviation was 0.67. The values for OH&P were standardized once. Unlike the previously discussed performance areas, an additional standardization does not need to be completed because the financial performance area is only made up of one metric and therefore does not involve obtaining the weighted averages of multiple z-scores. The formula for financial performance is devised as follows:

$$\text{Financial Perf.} = \frac{\text{OH}\&P - 1.45}{0.67}$$

This simple equation can be used to calculate relative financial performance scores of projects. The resulting score can be used in the sixth part of the PQR equation (*Eq.1*) presented in the beginning of *Section 5.3*.

5.3.7 PQR Part 7/7: Communication and Collaboration

The seventh and final performance area identified in the survey results as a key component of overall project success is communication and collaboration among the project team. Seven performance metrics are combined to provide a proxy for the communication and collaboration performance score: (1) the number of RFI per million dollars, (2) the RFI processing time measured in weeks, (3) the extent of rework, (4) the number of resubmittals per million dollars, (5) the absolute value of the total percentage of change, (6) the change order processing time, and (7) a representation of the Percent Plan Complete (PPC) trend. PPC is defined in *Appendix B* as a measure of work flow reliability, calculated by dividing the number of actual task completions by the number of planned task completions.

Data for some of the above metrics was quantitative, while ordinal data for other metrics was collected when exact quantitative values were difficult to obtain. For example, if the percentage of rework on the project was 0%, this value was coded to zero; if it was between zero and 1%, it was coded to one; 1% to 2% was coded to two; 2% to 3% coded to three; and values above 3% were coded to four. Regarding the total percent of change on the project, absolute values were used (i.e. negative values were converted to positive values). As for the PPC trend, only three values were possible: -1 for a decreasing trend, 0 for a stable PPC, and 1 for an increasing trend. The original numeric values for the remaining metrics were left intact: number of RFI and their processing time, number of resubmittals, and change order processing time.

The mean for the number of RFI per million dollars was 9.38 and the standard deviation was 8.03; while the range for the initial variables extended from zero to 29.44 RFI. The mean time that was needed to process each of these RFI was 1.71 weeks and the standard deviation was 0.91; while the range for the initial variables extended from zero to four weeks. The mean for resubmittals per million dollars was 1.76 and the standard deviation was 3.53; while the range for the initial variables extended from zero to 16 resubmittals. The mean for the rework ordinal variable, on the scale discussed above, was 1.21 and the standard deviation was 0.64; while the range for the initial variables extended from zero to four. The mean for the PPC trend, also on the scale discussed above, was 0.19 and the standard deviation was 0.48; while the range for the initial variables was limited to the values of negative one, zero, and one. The mean for total percent change in absolute value was 8.87 percent and the standard deviation was 8.77; while the range for the initial variables extended from zero to 35 percent. The mean for change order processing time was 3.61 weeks and the standard deviation was 2.26; while the range for the initial variables extended from zero to seven weeks.

The metrics were standardized individually, and then multiplied by their respective weights. Half of the communication and collaboration score relied on the processing times for both change orders and RFI (i.e. 25% for RFI processing time and 25% for change order processing time), since processing times are good indicators for effective communication among team members and quantitatively reflect how well different stakeholders are working together to achieve a fast process. The remaining five metrics each received an equal weight of 10%.

After weighted average of the new z-scores was computed, the resulting values were standardized again. The formula for communication and collaboration performance is as follows:

$$\mathrm{CC} = \frac{\frac{\mathrm{P}-0.19}{4.8} - \frac{\mathrm{RFI}-9.38}{80.3} - \frac{\mathrm{T}-1.71}{3.62} - \frac{\mathrm{S}-1.76}{35.3} - \frac{\mathrm{W}-1.21}{6.4} - \frac{|\mathrm{C}|-8.87}{87.7} - \frac{\mathrm{T}'-3.61}{9.02}}{0.62}$$

In this equation, P is used as a symbol for PPC, T for RFI processing time, S for resubmittals, W for rework, C for percent of total changes, and T’ for change order processing time. For simplicity and readability, the weights and standard deviations are compiled together in the denominator of each metric.

With respect to the values that can be used when applying this model, there are no restrictions to use values outside the ranges for the number of RFI per million dollars, the

RFI processing time (in weeks), the number of resubmittals per million dollars, the absolute value of the total percentage of change, and the change order processing time (in weeks). However, it is important that the scoring of the rework and PPC trend variables is performed in accordance with the ordinal scales discussed earlier in this subsection, to ensure the end result is interpretable when applying this equation to other projects.

The above equation can be used to calculate score for a project with respect to its quality performance. This resulting score can be used in the last part of the PQR equation (*Eq.1*) presented in the beginning of *Section 5.3*. This seventh component concludes the PQR model development.

The different weights used for performance metrics within each performance area were modified as part of a sensitivity analysis, and minor variations were observed for the resulting PQR values; however, the overall ranking of the projects in terms of PQR was not affected. As discussed earlier in this section, in order to facilitate the use of the PQR formulas, *Appendix G* presents a summary of the equations developed in this chapter. The appendix also includes a numerical example which consists of initial inputs from a sample project, along with their corresponding performance area scores and the final PQR value.

5.4 PQR Validation Using Factor Analysis

Like with any newly developed model, testing and validation are needed to confirm that the PQR function is adequate. Therefore, before PQR can be used to compare IPD to non-IPD projects, two independent statistical techniques will be used to validate the PQR function: factor analysis and multidimensional scaling. The two independent overall performance techniques will be used separately and their results will be compared to the PQR results for confirmation.

Factor analysis is a statistical technique that is used to model and quantify latent variables that cannot be measured (e.g. intelligence). There are several observable variables that can be included in the measurement of intelligence, such as logical thinking, mathematical ability, and linguistic ability. Factor analysis attempts to find the intelligence latent variable based on the data from the measureable variables, with no preconceptions (Cox 2005).

AEC project success, similar to intelligence, is a latent variable that cannot be measured directly. Performance data for metrics used to measure success were collected in this research. In this section, factor analysis is used to model project success based on these performance metrics.

One commonly used method for factor analysis is through principal components. The method first finds PCs and then rotates them so they line up more closely with some of the original variables. In PCA, the interpretation of PCs is often difficult because many coefficients tend to be more or less equal. Factor analysis increases the values of some of the coefficients, and makes others negligible, facilitating the interpretation of the results. A more detailed explanation of factor analysis can be found in Bartholomew and Knott (1999).

Factor analysis can validate PQR by exploring the performance metrics and looking at project success from a different perspective. This analysis provides two benefits. First, it allows for a 3D visualization of performance that facilitates the understanding of the multidimensional problem at hand, while offering a new way to look at AEC project performance. Second, the 3D results can be combined into one dimension that, then in turn, can be compared to the PQR to validate its significance. Since PQR is based on survey responses to questions on project success factors, the independent factor analysis unsupervised technique provides a means to validate the PQR results.

Several iterations are typically required to identify the adequate number of factors needed. In other words, when reducing the dimension of the dataset, one needs to determine the satisfactory number of factors that best represent the original metrics. The first iteration started with all the performance metrics for which data was available, and attempted to reduce the dimension of the dataset to one factor. The second step changed the number of attempted factors to two, the third step attempted three factors, and so on. At each stage, the loadings or coefficients of all performance metrics were observed, and the metrics that obtained negligible loadings were removed from the analysis because they cluttered the list without significantly affecting the results. The final run was reached when all the remaining performance metrics had significant loadings. The first column of *Table 27* shows the outstanding performance metrics used in the final analysis.

The analysis resulted in three factors representing around 46% of the variance observed in the original dataset. The loadings for the three factors are shown in *Table 27*. One can see that Factor 3 consists solely of unit cost. Therefore, Factor 3 will be renamed the Cost Factor. The three largest loading values for Factor 2 are construction speed, delivery speed, and intensity. All three are schedule metrics. Therefore, Factor 2 will be renamed the Schedule Factor. This factor also includes small loadings for labor factor and recordables.

TABLE 27: Factor Loadings

	Factor Loadings		
Performance Metrics	**Factor 1**	**Factor 2**	**Factor 3**
Quality	**-0.52**		
RFI	0.70		
RFI Processing Time	0.73		
Change Order Processing Time	0.60		
Resubmittals	0.74		
Deficiency Issues	0.56		
OH&P	**-0.58**		
Return Business	**-0.77**		
Construction Speed		**0.95**	
Delivery Speed		**0.96**	
Intensity		**0.94**	
Labor Factor		0.56	
Unit Cost			**0.77**
Cost Growth	0.45		
OSHA Recordables		0.41	
PPC Trend	**-0.40**		
Labor Extra	0.37		
|Changes|	0.47		
Punchlist Items	0.42		
SS Loadings	*4.01*	*2.38*	*1.38*
Proportion of Variance	*0.25*	*0.15*	*0.09*
Cumulative Variance	*0.25*	*0.40*	*0.49*

Factor 1, which accounts for most of the variance, consists of quality metrics as well as non-traditional metrics that are used in addition to the more standard cost and schedule metrics identified earlier. These non-traditional metrics are mostly related to the business and communication performance areas, such as OH&P and return business, RFI, and resubmittals. Therefore, Factor 1 will be renamed Quality, Business, and Communication Factor. It is interesting that the negative loadings in Factor 1, highlighted in the table, all correspond to performance metrics that need to be maximized in a project. Through a change of sign, Factor 1 emphasizes the contrast between these metrics and others that typically need to be minimized.

The hypothesis that these three factors are sufficient to represent the original metrics was tested, and revealed a Chi-squared statistic of 133.79 on 117 degrees of freedom. This translates into a p-value of 0.137. The p-value is non-significant, which means that the three factors are sufficient to summarize all the performance variables.

The factors can be interpreted more easily by transforming the variables in order for them to increase homogeneously. Therefore, Factor 1 was transformed by taking the negative of all its values, so that an increase in quality results in an increase in Factor 1. A similar transformation was used for Factor 3 so that a decrease in cost results in an increase in Factor 3 or cost performance. After this minor transformation, all three factors' scores were more intuitive and increase in the same direction.

Once the factor analysis is completed, scores of each of the three dimensions can be calculated for each project in the database. These scores then can be plotted and color-coded based on their delivery system in order to observe the behavior of IPD projects as compared to non-IPD projects. In the following figures, each number represents one of the 35 projects in the database, Using specific numbers instead of plotting regular points facilitates the interpretation of the plot by allowing the reader to follow the same anonymous project across several graphs.

Figure 43 is a plot of Factors 1 and 2, with IPD projects shown in green, non-IPD projects in red, and IPD-ish projects in black. In general, IPD projects were on the right side of the plot, denoting the superiority of IPD projects with respect to Factor 1 or the quality, business, and communication metrics. This superiority has been statistically proven in the univariate analysis of *Chapter 4* based on the analysis of individual metrics. The factor analysis confirms this finding when the metrics are combined. Furthermore, one can see that the IPD projects are not clustered in any specific way relative to the vertical axis, representing no great differences in schedule metrics.

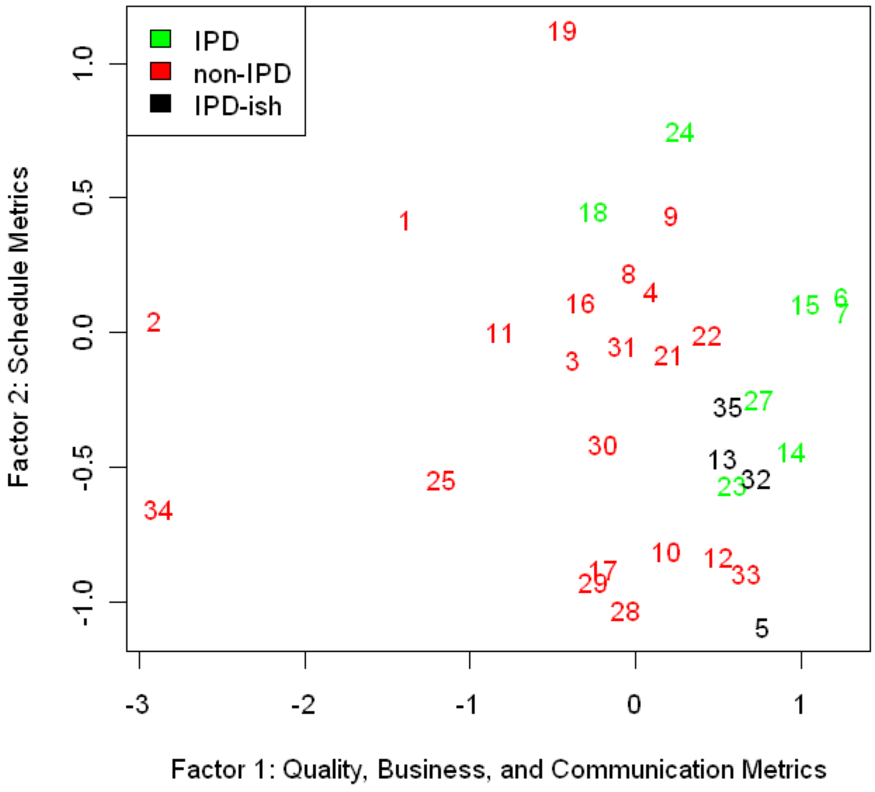


FIGURE 43: Factor Analysis – Factors 1 and 2

Figure 44 is a plot of Factors 1 and 3. The cost metrics can be interpreted similarly to the schedule metrics; however, it is interesting that some IPD projects were located in the upper right corner of the plot, showing higher performance in both Factor 1 and Factor 3.

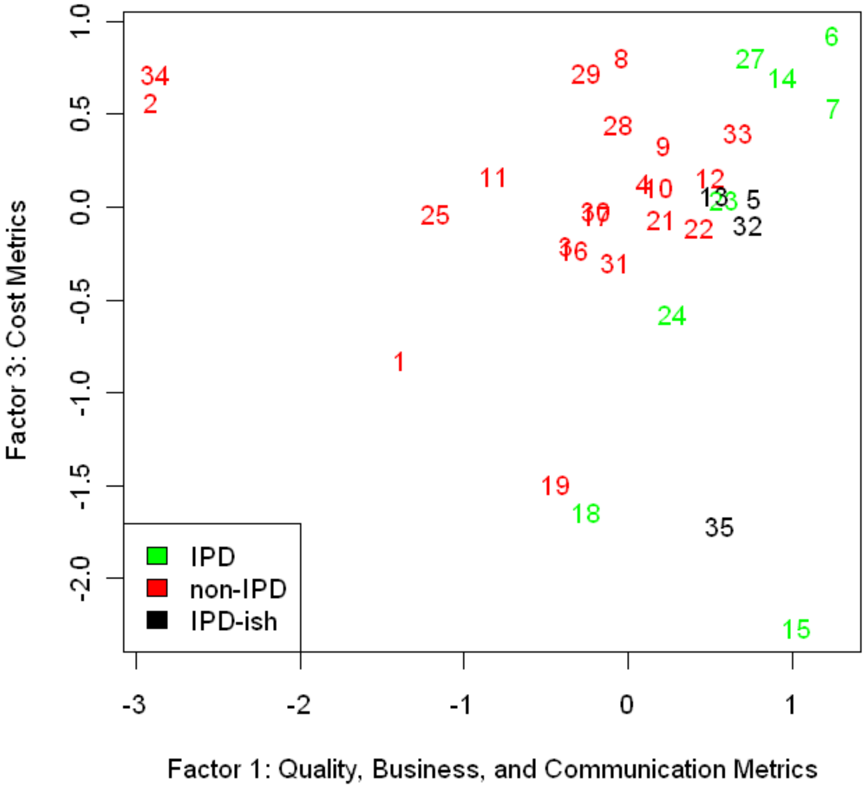


FIGURE 44: Factor Analysis – Factors 1 and 3

In order to visualize the last combination of the three factors, *Figure 45* is a plot of Factors 2 and 3. This plot does not show major differences for IPD projects, which performed similarly to the remaining projects in terms of both cost and schedule. The cost performance plot is interesting because some IPD projects were the highest performers, while other IPD projects were the lowest performers, as shown by the green numbers on the vertical extremes.

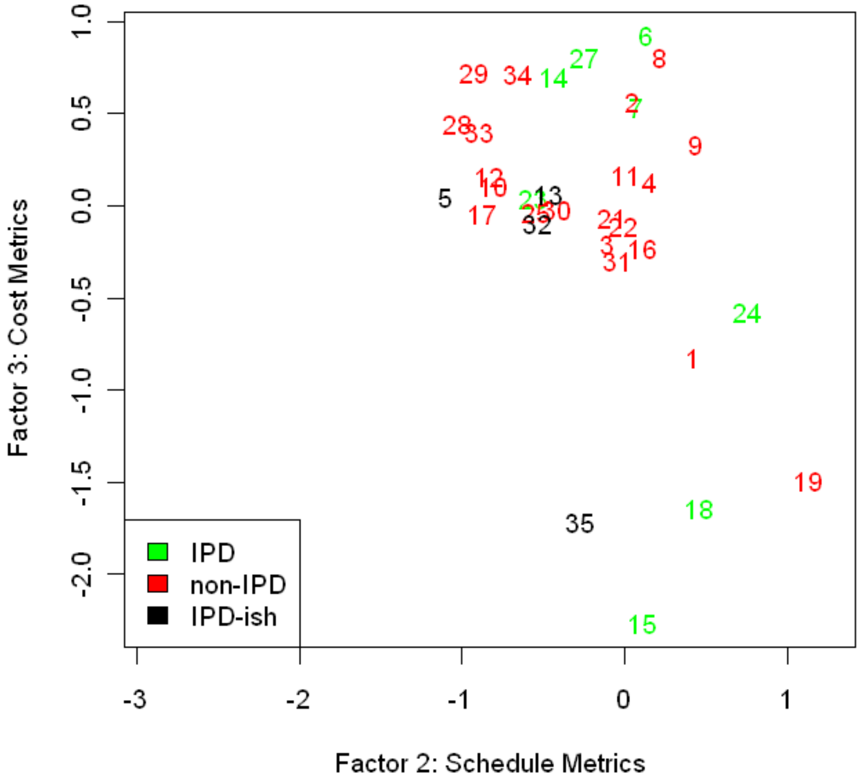


FIGURE 45: Factor Analysis – Factors 2 and 3

Since all the key performance metrics were summarized by only three factors, these three factors can be shown together in a 3-dimensional plot. *Figure 46* displays a 3D plot, which was created especially for this purpose. The upper and lower parts of the figure show two different angles when rotating the plot. The plane shown in the plot is actually a regression plane; however, it is not being used as such, but instead provides a better visualization of the 3D plot in this two-dimensional document.

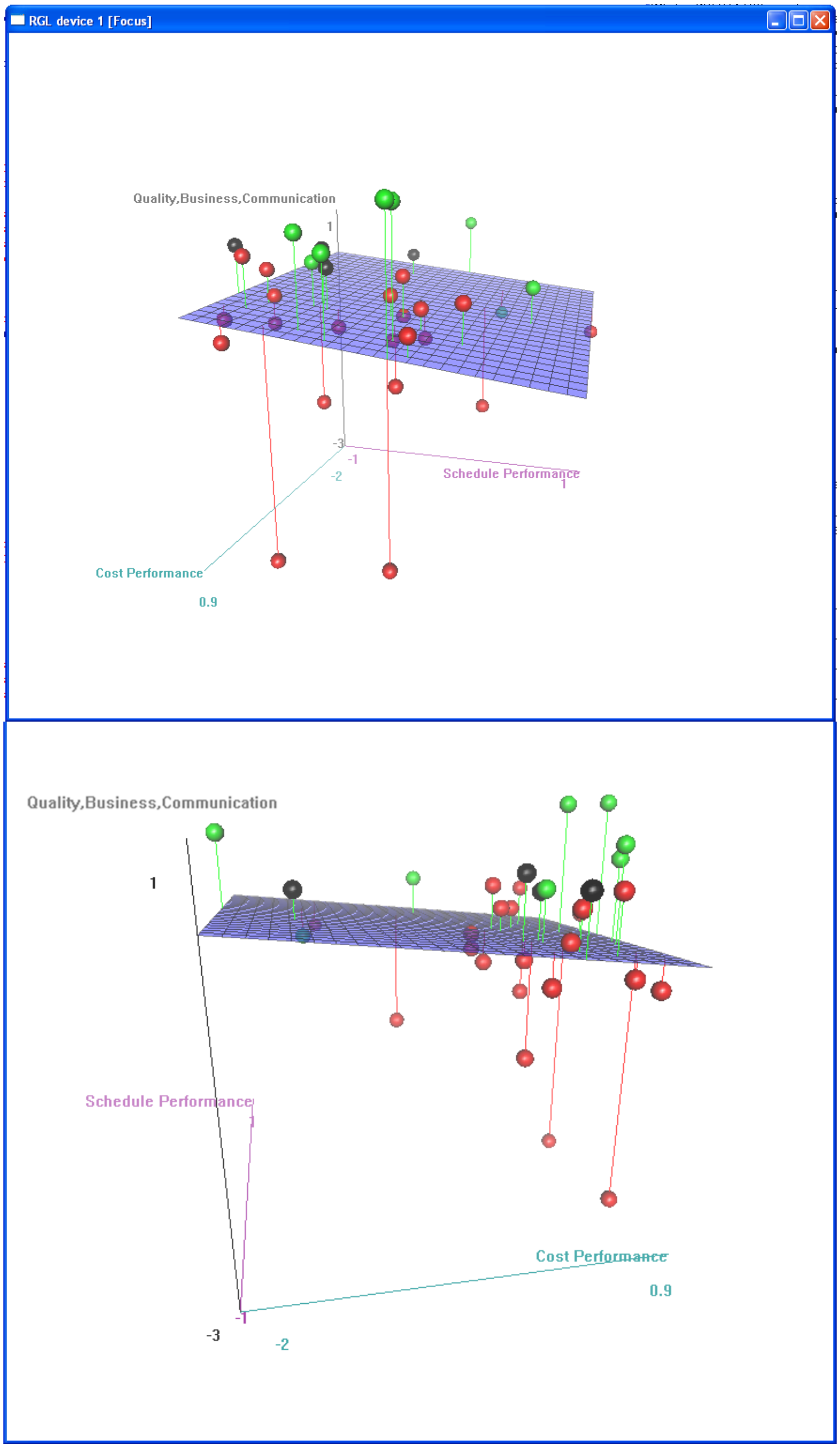


FIGURE 46: 3D Plots of the Factor Analysis Scores

Non-IPD projects, separated by delivery method, such as DBB, DB, and CMR, were also analyzed to determine if they are behaving homogeneously or differently based on the three factors. As both plots in the upper portion of *Figure 47* show, the most visible differences are between IPD projects in green and DBB projects in red, which scored visibly lower on Factor 1. These factors also were combined in a 3D plot, as shown in the lower right corner of *Figure 47*.

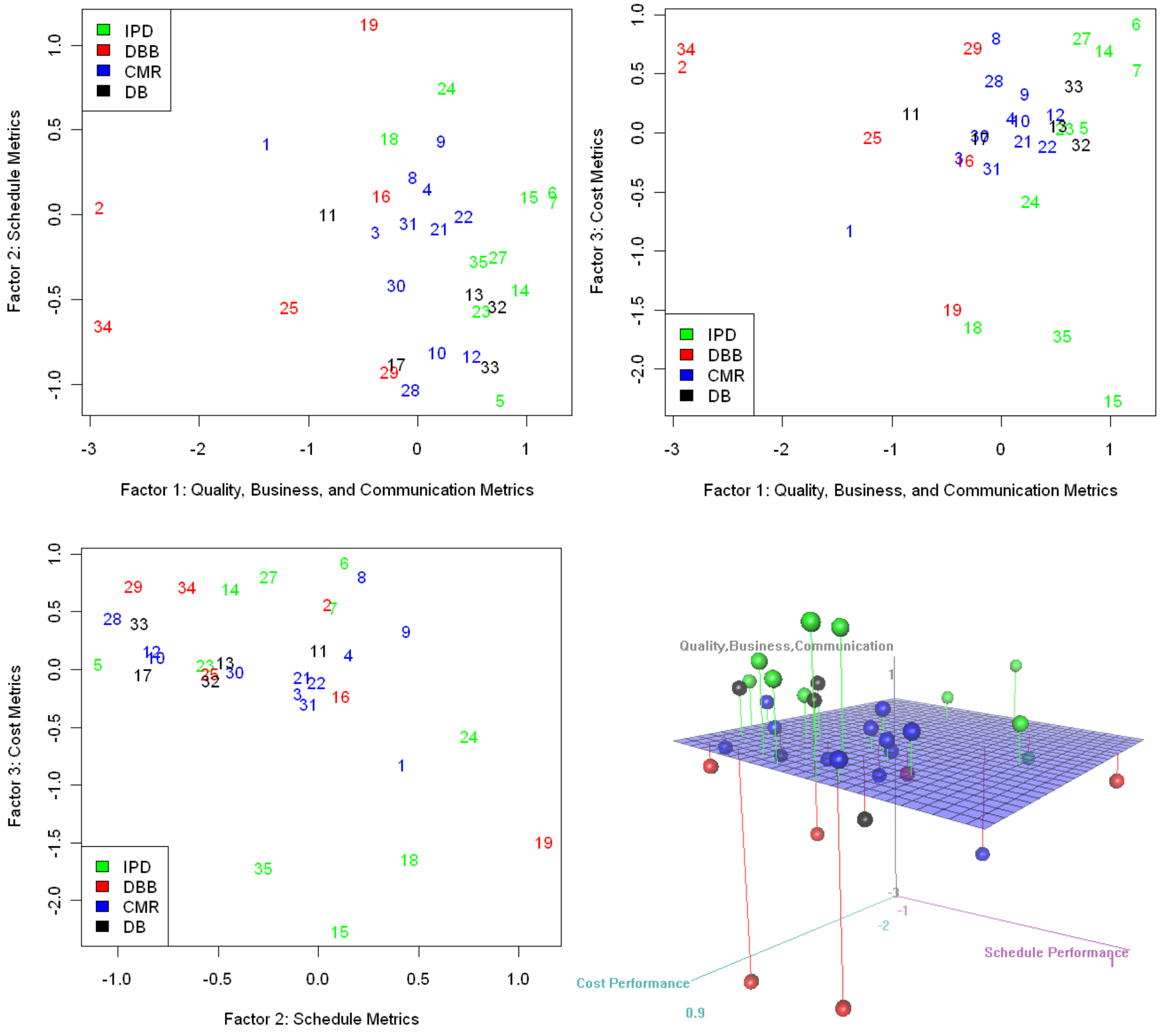


FIGURE 47: Factor Scores for Separate PDS

The factor analysis proves that the problem at hand is multidimensional and can be represented in a minimum of three separate dimensions. These three dimensions are best visualized by a 3D plot. However, for the purpose of this subsection, these three dimensions will be further combined into one dimension that approximates overall performance. The norm of the three dimensions is calculated to compute a unified metric for performance. This metric is compared to PQR to verify and validate it.

Before calculating their norm, the scores for the three factors have to be transformed to positive values. For each factor, the minimum over all projects is subtracted from the z-score of each project, making the minimum value zero. These new values then are divided by their maximum, which gives transformed factor scores that range from zero to one. These scaled scores are used to calculate the norm by taking the square root of the sum of their squares. The factors norm for each project, which can be higher than one, will be compared to the PQR.

Since the units for the PQR are z-scores, scaling of the PQR is now needed in order to compare z-scores to norms. Similar to the factor scores, PQR scores are converted to a range of zero to one by subtracting the minimum and then dividing by the maximum. The upper part of *Figure 48* displays how the norm of the three factors (in red) almost mirrors the scaled PQR scores (in blue).

The bottom part of *Figure 48* shows the same scores after the values for the norm were divided by their maximum value in order to limit the maximum to one. The projects in this plot are ordered in increasing values for the scaled norms, and the linear regression of the PQR scores is almost identical to the plot of norm values.

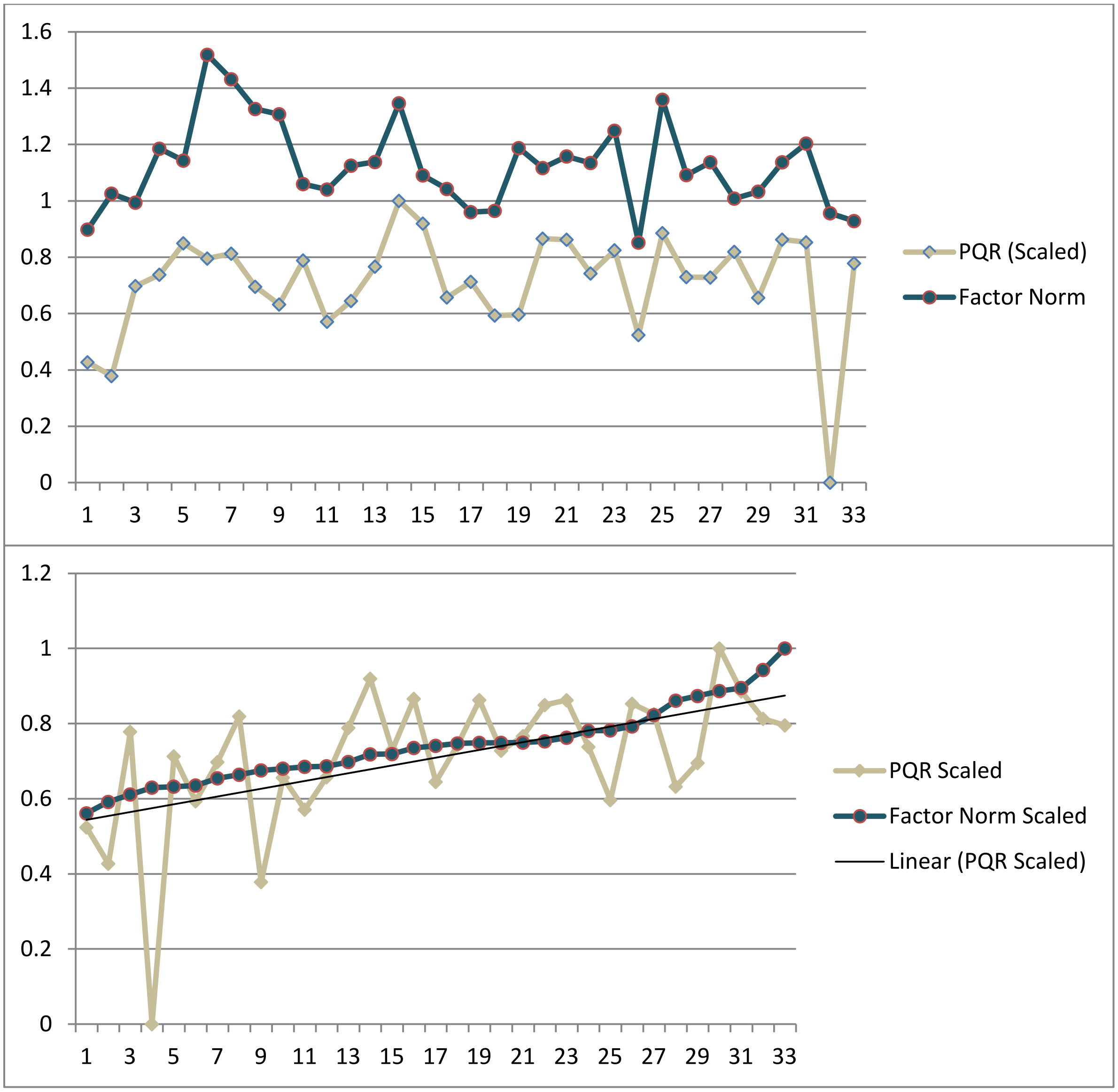


FIGURE 48: Comparing PQR and the Norm of the Three Factors

The factor analysis is an unsupervised multivariate data analysis technique that is independent form the survey weights used to develop PQR. The similarity of the two results confirms that PQR is an adequate performance model, which can be used to compare overall performance for AEC projects. In the next section, multidimensional scaling also will be used as a third method to independently confirm the validity of PQR.

5.5 PQR Validation Using Multidimensional Scaling (MDS)

Factor analysis was used in the previous section to validate the newly developed PQR. Another method used for PQR validation is multidimensional scaling.

When modeling the project quarterback rating, it is difficult to collect data for overall project success. Each individual has a different baseline for success according to his or her individual experiences. In fact, when one is asked how successful a specific project was, he or she is likely to think about this specific project as compared to their previous project experiences. This experience is different for every individual, and therefore there is no definitive baseline by which to compare.

This section attempts to correct this issue by providing a quantifiable baseline that enables a more accurate comparison of project success. Individuals who provided data for more than one project were asked to compare their projects in pairs. This method clarifies the baseline of each comparison. Eleven projects collected from the same company were compared to each other, resulting in a total of 55 pairs of projects compared. A panel of company executives and project managers that were involved in these projects met and discussed the matrix in order to compile comparative data for this section of the research study.

Once the comparative ratings were available, classical multidimensional scaling (MDS) can force the order of all projects in one dimension. MDS is used to represent an observed proximity matrix that can measure similarity or dissimilarity. For example, MDS can use the airline intercity distances for several pairs of cities, and then reconstruct a map of these cities based on the paired dimensions. This technique was used on the pairwise comparison data to output a rating that orders all projects in the same dimension of overall project success. A more detailed explanation of MDS can be found in Cox and Cox (2000).

For the purpose of this study, the MDS algorithm begins with a similarity matrix and allocates positions for the 11 individual projects based on one or more dimensions. MDS uses absolute differences so the absolute values of the differences in overall project performance were used as inputs.

Initially, MDS was performed to force all projects into one dimension: the rating of projects in terms of overall performance. The first row in *Table 28* shows how project B was on one extreme with an overall score of *-2.6*, while all the remaining projects were around the same value of *0.3*. In fact, this analysis led to discussions with the industry respondents familiar with project B, and they identified it as the least performing project, by far, out of this sample.

Because the performance of one project received an extremely low rating as compared to other projects, it could potentially skew the overall rating. Therefore, the one-dimensional MDS analysis was re-run without the input of project B. The results are shown in the second row of *Table 28*. This new result makes perfect sense in terms of overall performance when checking the pairwise comparisons of the remaining projects. However, since removing the value for project B alters the final results of the original MDS analysis, additional analysis is needed to assess the new findings.

TABLE 28: MDS Variations

1D	B	G	J	K	F	D	E	A	H	C	I
	-2.62	-0.02	0.25	0.25	0.26	0.27	0.29	0.30	0.31	0.34	0.37
1D w/out project B	B	G	J	K	C	F	H	I	A	E	D
		-1.10	-0.81	-0.81	-0.36	-0.15	0.10	0.53	0.63	0.96	0.99
2D	B	G	J	K	C	F	H	I	A	E	D
		-1.09	-0.83	-0.83	-0.38	-0.17	0.08	0.51	0.62	0.95	0.98
Percentage difference		0.6%	2.2%	2.2%	6.1%	8.7%	20.3%	4.2%	2.5%	1.4%	1.1%

Thefore, a two-dimensional analysis was performed to present a different picture of these comparative results. The projects were now forced into two dimensions. In the 2D analysis, shown in *Figure 49*, one can see that project B is so different than the other projects to a point where the first dimension mostly expresses project B versus other projects. The original comparative scores are adequately expressed by the second dimension of the 2D results (labeled as *Coordinate 2* in the figure). The value for project B was deleted in the third row of *Table 28* because the difference between project B and the remaining projects is expressed through the other dimension, as stated earlier.

As expected, the 2D results came out very similar to the one-dimensional results that did not consider project B. In fact, there was less than 5% difference in values, on average, and the same order of projects was maintained. Both 1D and 2D techniques were used to showcase the complexity of the problem, which makes any simple solution difficult.

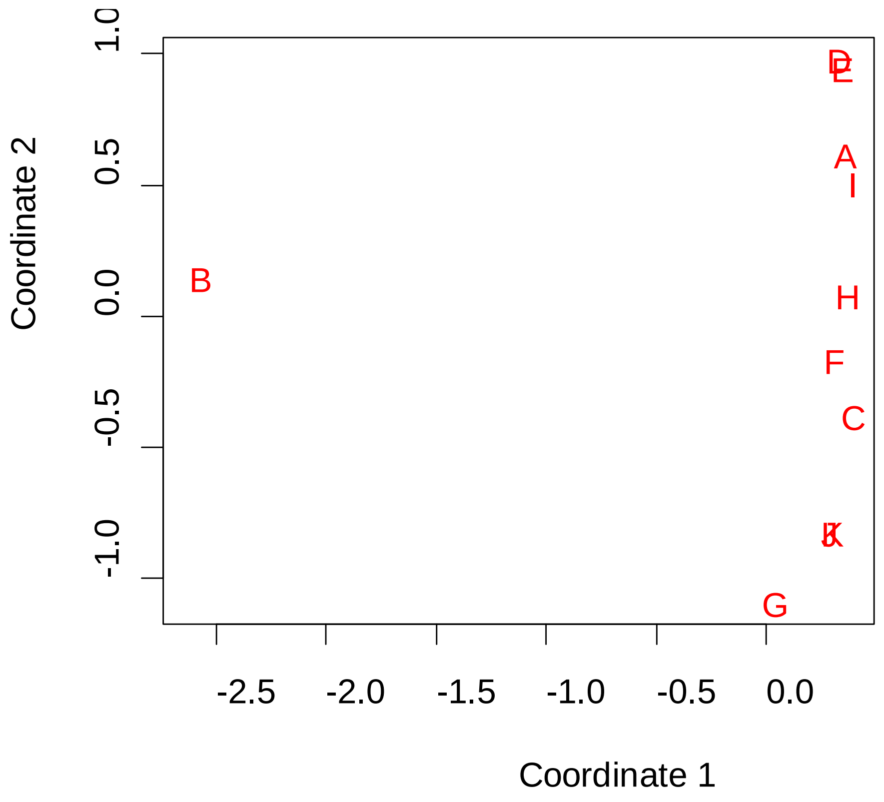


FIGURE 49: MDS 2D

The MDS analysis presents insightful results and provides a baseline to validate PQR. The key outcome is a dimension that represents overall project performance for several projects. This MDS score, only available for a subset of 11 projects in the database, will be compared to the PQR of this subset. *Figure 50* demonstrates how PQR compares to the MDS scores that represent the actual overall project performance for these 11 projects. The horizontal axis represents the 11 projects in increasing PQR values (shown in red). The MDS scores also increase with the PQR, further validating that PQR is an adequate model for overall project performance.

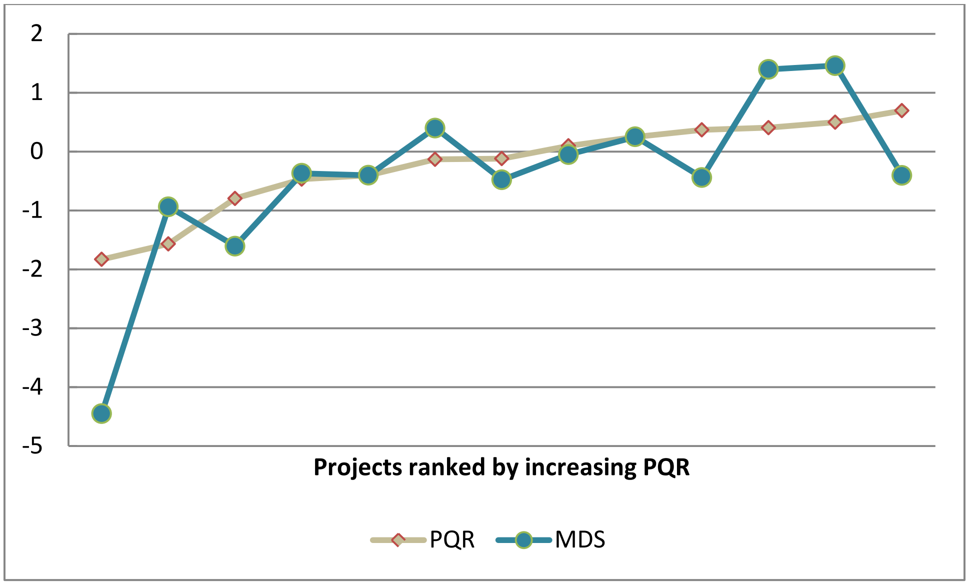


FIGURE 50: Comparing PQR to MDS Results

Data collection for the MDS analysis was independent from the data collection and development of the PQR. The MDS analysis was based on comparative ratings of pairs of projects, while the PQR was based on a survey of top performance metrics that contribute to overall project success. As was discussed earlier, the factor analysis also was completed independently of both MDS and PQR development, since it relies solely on an unsupervised exploration of the raw performance metrics data. In this chapter, three different methods with three different origins all gave similar results.

While it is important that PQR gives a true representation of overall project performance, it is also important to recognize that PQR is only meant to be an approximation of this overall performance. One reason for this disclaimer is that different industry experts have different views and opinions regarding what are the top metrics to consider, as shown in the survey results. Additionally and perhaps more importantly, different projects tend to have different priorities in terms of performance; i.e. sometimes cost can be the driver if the budget cannot be increased, other times schedule is key when an official opening date has already been announced to the public. Therefore, the PQR model does not need to be, and cannot be an exact model of overall project performance. Rather, it consists of an adequate combination of several key metrics to provide a quantitative approximation of overall project success.

Now that the PQR model has been validated using independent techniques, it can be utilized to compare IPD to non-IPD projects overall performance. The following section complements *Chapter 4* by going one step beyond the testing of individual performance metrics and areas, and test overall project performance.

5.6 Testing IPD Projects Overall Performance Using PQR

Chapter 4 compared IPD to non-IPD performance based on individual metrics, and the results demonstrated that IPD has a higher performance based on several metrics, and yet no performance differences on other metrics. The development of the PQR offers an opportunity to compare IPD and non-IPD projects' overall performance using the single new metric.

First, PQR was computed for all 35 projects in the database by using the equations presented in this chapter. IPD projects were then separated from non-IPD projects. The average PQR for IPD projects was 0.54, or more than half a standard deviation above the average project in the research database. The average PQR for non-IPD projects was -0.31, or about a third of a standard deviation lower than the average project in the database.

As was shown for individual performance metrics, *Figure 51* includes boxplots comparing the PQR for IPD and non-IPD projects. The difference between the two samples is clear, and what is also interesting here is that IPD-ish projects are behaving like IPD projects when it comes to overall project performance.

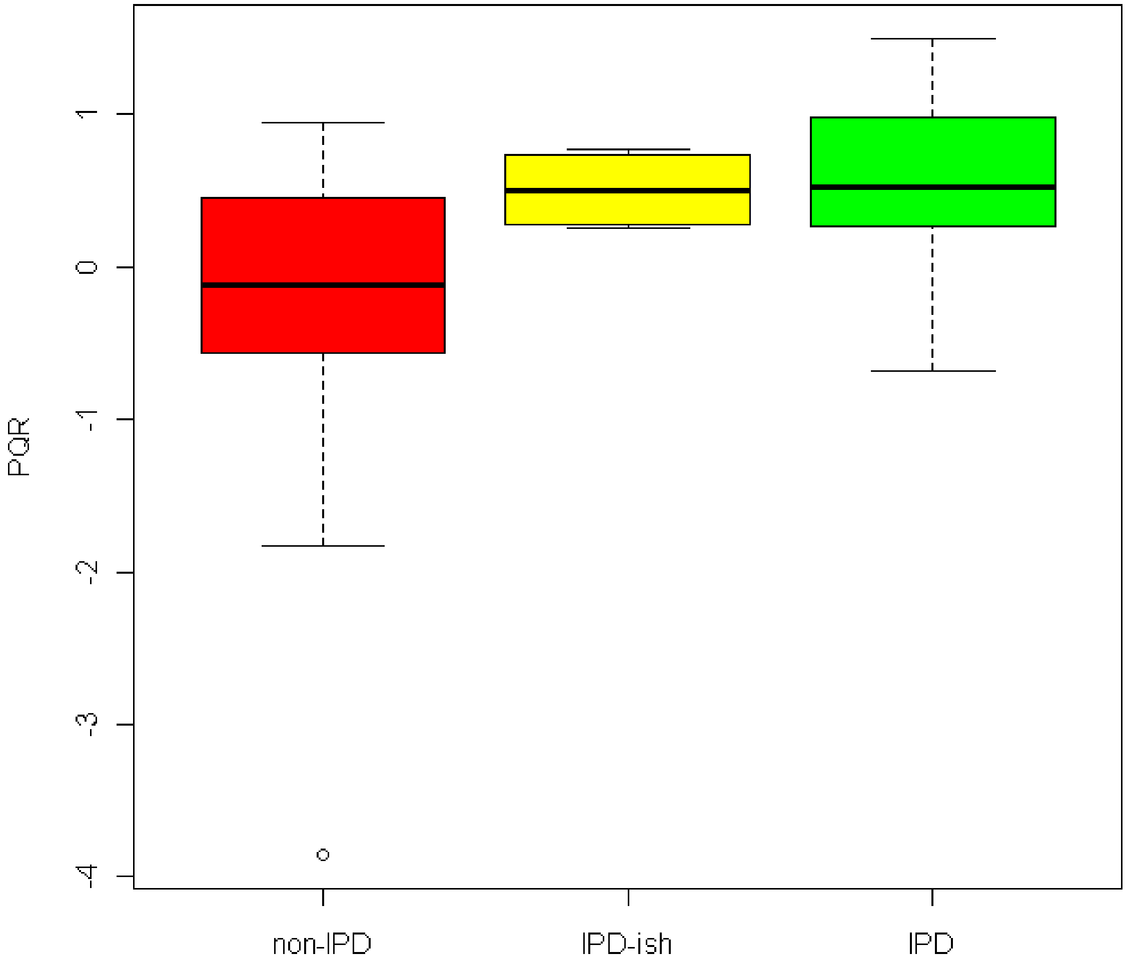


FIGURE 51: PQR Boxplots

Then MWW tests were performed on the PQR data to determine if these differences in performance are statistically significant. The tests resulted in a p-value of 0.015, which is considered significant at the 0.05 level. This result statistically confirms IPD projects have a superior overall performance as compared to non-IPD projects.

Furthermore, t-tests were conducted to complement the MWW results. *Figure 52* shows Q-Q plots that are fairly linear, meaning the PQR data can be assumed normally distributed. This assumption strengthens the results of the t-test. The one-sided test showed a p-value of 0.009, which is considered very significant even at the strict 0.01 significance level and proves that IPD projects have a superior performance.

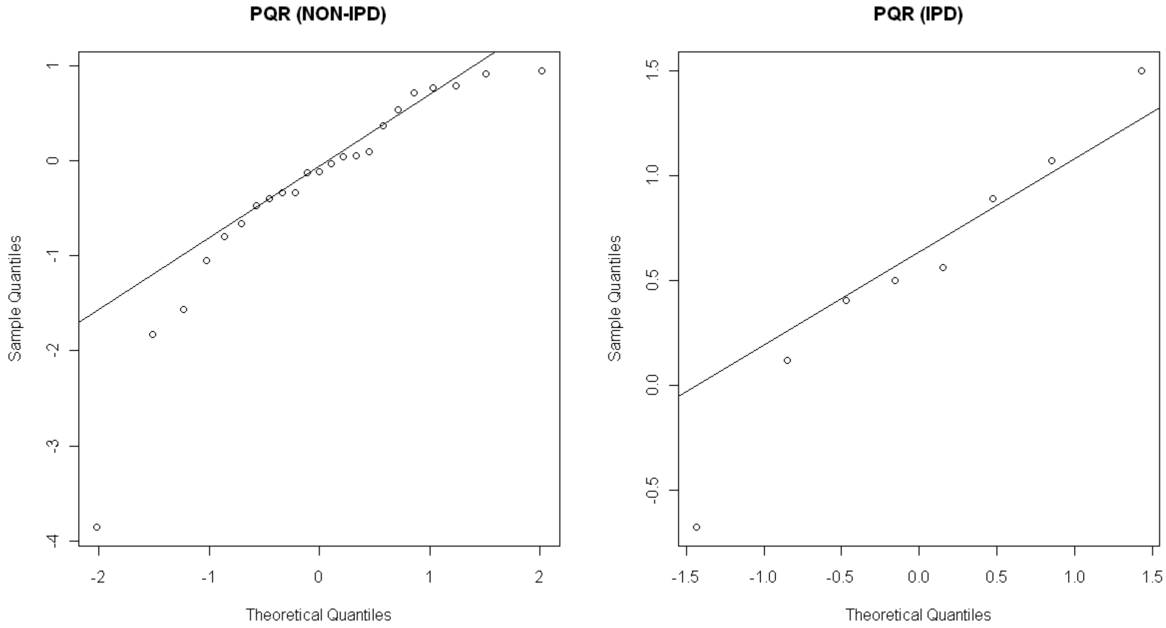


FIGURE 52: PQR Q-Q Plots

This analysis offers additional proof that IPD projects do not only show superior performance based on individual performance metrics, but also on overall project performance as gauged by the newly developed PQR that integrates key metrics and provides a comprehensive performance assessment.

Next, one final analysis splits the non-IPD projects into DBB, CMR and DB. *Figure 53* shows the increase in overall project performance when moving from DBB to CMR to DB and finally to IPD.

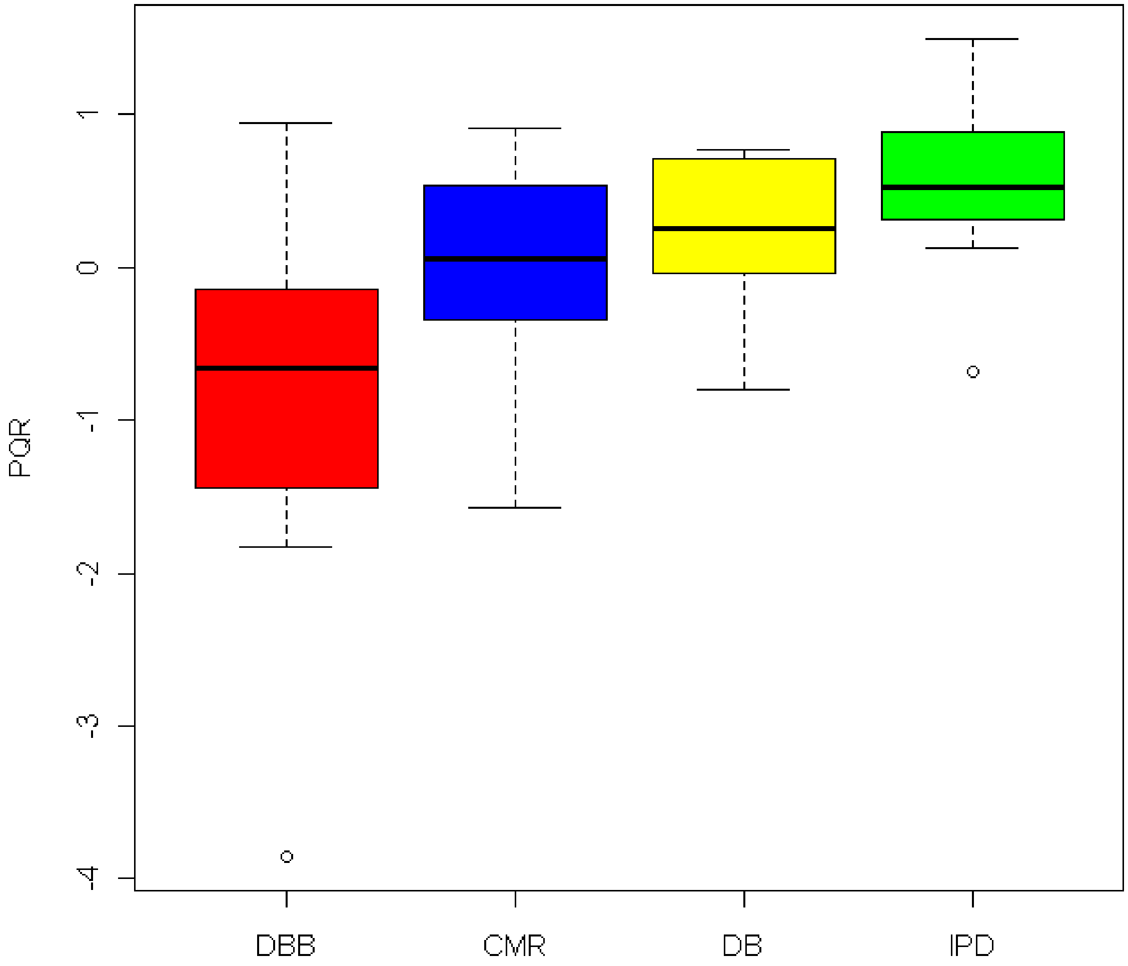


FIGURE 53: Boxplots for PDS PQR

This visual representation is supplemented by the noteworthy results of the Kruskal-Wallis test that show a p-value of 0.008 denoting the differences in overall performance are statistically significant at the 0.05 level and the 0.01 significance level. Additionally, a similar ANOVA test resulted in a p-value of 0.002, which is even more statistically significant. This last result undoubtably proves that more integrated delivery systems result in increased overall project performance for complex institutional projects.

After developing, validating, and using PQR to prove that IPD projects have a superior overall performance, the next step is to model PQR with the individual delivery characteristics in order to understand what are the most important variables that have the strongest effect on overall project performance. This chapter provided the groundbreaking statistical proof that IPD projects have an improved performance when compared to non-IPD projects. Two questions naturally follow this discovery:

1. What IPD delivery characteristics are driving these performance benefits?
2. Can IPD project performance be modeled using these key project delivery characteristics?

Chapter 6 will answer these two questions. The chapter presents several multivariate analyses that investigate how the *3T* delivery areas introduced in *Chapter 3* – Terms, Tone, and Tools – affect overall project performance. PQR will be modeled using the various delivery characteristics of AEC projects.

Chapter 6. Modeling IPD Performance using Multivariate Data Analysis

Previously, *Chapter 4* discussed the univariate analysis, which proved that using IPD results in increased project performance based on 14 metrics. *Chapter 5* discussed the development of the Project Quarterback Rating, which integrates all key performance metrics, and demonstrated that IPD projects result in superior overall performance. *Chapter 6* builds on what was discussed earlier and serves one major purpose: it presents explanatory models to elucidate how various IPD delivery characteristics are affecting overall project performance. In other words, this chapter investigates and identifies the actual characteristics that are responsible for the superior IPD performance.

Chapter 6 is divided into four sections. The first three sections cover three different methods used to build performance models for AEC projects. First, stepwise regressions are used as a variable selection technique. Second, principal component analysis is used on all delivery characteristics as a dimension reduction technique, making possible the identification and interpretation of new uncorrelated variables that explain key delivery dimensions. Third, the principal components of each individual *3T* domain are recovered, which are in turn used to model performance. The 3T principal components are used to visualize a 5D performance model. Finally, the fourth section builds on the first three and tests the identified characteristics versus project performance on an individual basis.

6.1 Variable Selection with Stepwise Regressions

Stepwise regressions were used to model overall project performance. This procedure tests all the delivery variables against PQR and includes only the significant variables in the final model. The result is a model of performance based on key delivery inputs.

Major approaches for variable selection are: (1) forward selection, (2) backward elimination, and (3) stepwise regressions that combine the forward and backward approaches. As the name suggests, forward selection starts with an empty model that does not include any variables. This approach then tests each individual variable, and the variables that are statistically significant are included in the model. The reverse procedure is applied for backward elimination, which starts with a model that includes all potential variables. The procedure would then test each individual variable, and the variables that are not statistically significant would be removed from the model. Combining these two approaches results in stepwise regressions, which add or remove variables from the model at each stage.

In this section, the combined stepwise regression method is applied to model PQR using all the individual delivery characteristics. The model started empty with no variables included originally, and then added and deleted variables that were most correlated with PQR. There were five resulting variables in the final model, which resulted in an R^2 value of 0.71: fiscal transparency, new experience with the project team, lump sum compensation, use of BIM, and use of innovative tools and techniques. The scoring of each of these variables was discussed earlier in *Section 3.5*. This model is presented as follows:

PQR1 = -1.8333 + 0.0896 fiscal transparency + 0.6036 new experience with the project team
- 0.5534 lump sum compensation + 0.0191 BIM - 0.1572 innov. tools and techniques

Altering the model by removing the relatively broad characteristic representing the use of innovative tools and techniques, which is a combined delivery characteristic that includes prefabrication, mockups, point cloud technologies, project training sessions, constructability analyses and safety awareness trainings, caused a new variable to be included: the Last Planner System (LPS). The modified model had a lower R^2 of 0.66:

PQR2 = -2.0429 + 0.0704 fiscal transparency + 0.5481 new experience with the project team
- 0.4885 lump sum compensation + 0.0160 BIM - 0.0156 LPS

An interesting modification would be including of the project delivery system as the initial variable to begin the model-building process, since it was proven, in the previous chapter, to have an effect on overall project performance. This alteration did not change the model in a significant manner, and the output remained the same. Another modification would be removing the variables with the lowest p-values, which in this case were PDS, fiscal transparency, and LPS. The resulting model, solely made up of the new experience with the project team, lump sum compensation, and the use of BIM, still had an R^2 value of 0.63.

PQR3 = -1.9175 + 0.6440 new experience with the project team
-0.6538 lump sum compensation + 0.0166 BIM

It is interesting that the three key variables represent the 3Ts originally introduced in this study: contractual terms, social tone, and tools and techniques. All of these models originally started with an initial iteration consisting of an empty model that included no variables. Similar models, for which the initial iterations included all potential variables, were also investigated.

The first model included 14 input variables, some of which were statistically significant: contractor procurement, lump sum compensation, OCIP, new experience with the team, fiscal transparency, use of JIT, use of BIM and use of innovative tools and techniques. All of these significant variables were used to build another model that started with these variables and was allowed to delete or reinsert any of these variables at any stage in the model building process. The final model turned out to be exactly the same as the PQR1 model built previously with five resulting variables and an R^2 value of 0.71:

$$\begin{aligned} \text{PQR1} = {} & -1.8333 + 0.0896 \text{ fiscal transparency} + 0.6036 \text{ new experience with the project team} \\ & - 0.5534 \text{ lump sum compensation} + 0.0191 \text{ BIM} - 0.1572 \text{ innov. tools and techniques} \end{aligned}$$

Figure 54 shows a comparison of the three models (*PQR1*, *PQR2*, and *PQR3*) to the actual PQR, with the first model, *PQR1*, showing the highest R^2 value. The horizontal axis in the figure represents all 35 projects ordered based on increasing PQR scores. One interesting observation is that the top two PQR scores correspond to IPD projects.

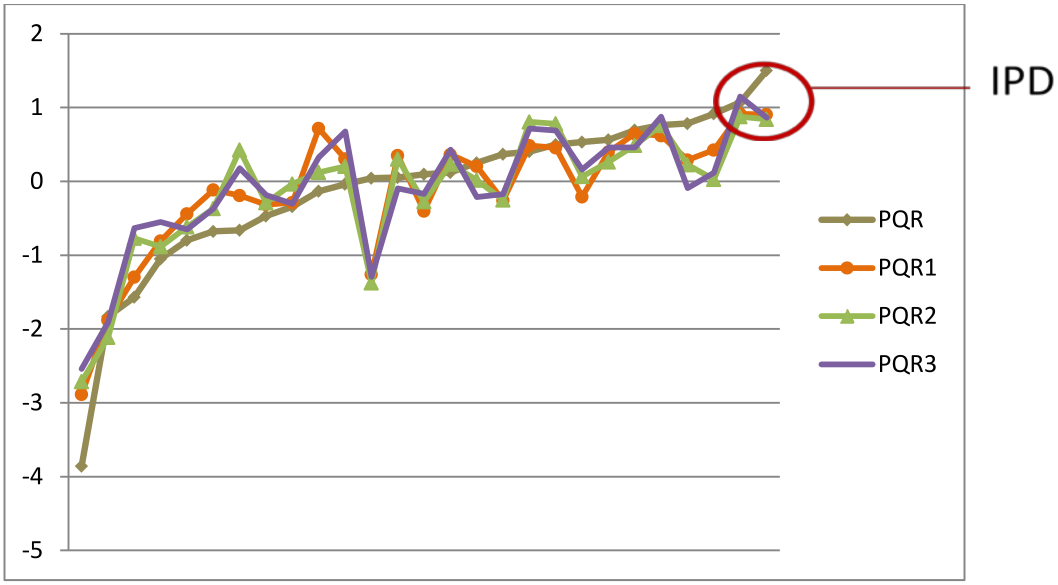


FIGURE 54: PQR Stepwise Regression Models

Since the *PQR1* model is based on regressions, the ranges of the inputs need to be respected when the model is used for a new project. *Section 3.5* discusses in detail the coding of the several variables used in this study. The working range of the variables used in this model are zero to nine for fiscal transparency, two to five for the new experience of the CM/GC, zero to three for lump sum compensation, zero to 52.23 for BIM, and 0.83 to 8.17 for the use of innovative tools and techniques, on the scales discussed in *Section 3.5*.

Stepwise regression models can replicate PQR behavior, but cannot fully explain project performance using delivery characteristics because of multicollinearity. When some independent variables are correlated, in this case delivery characteristics, the model does not allow for adequate interpretation of individual variables. To showcase correlations between the input variables, *Figure 55* shows a heatmap of correlations for all key delivery variables in addition to PQR. Lighter colors signify stronger correlations between variables, and therefore the diagonal is entirely white. Multicollinearity does not reduce the predictive power of the model; it only affects the interpretation of individual variables. The correlations between the performance metrics can be found in *Appendix F*.

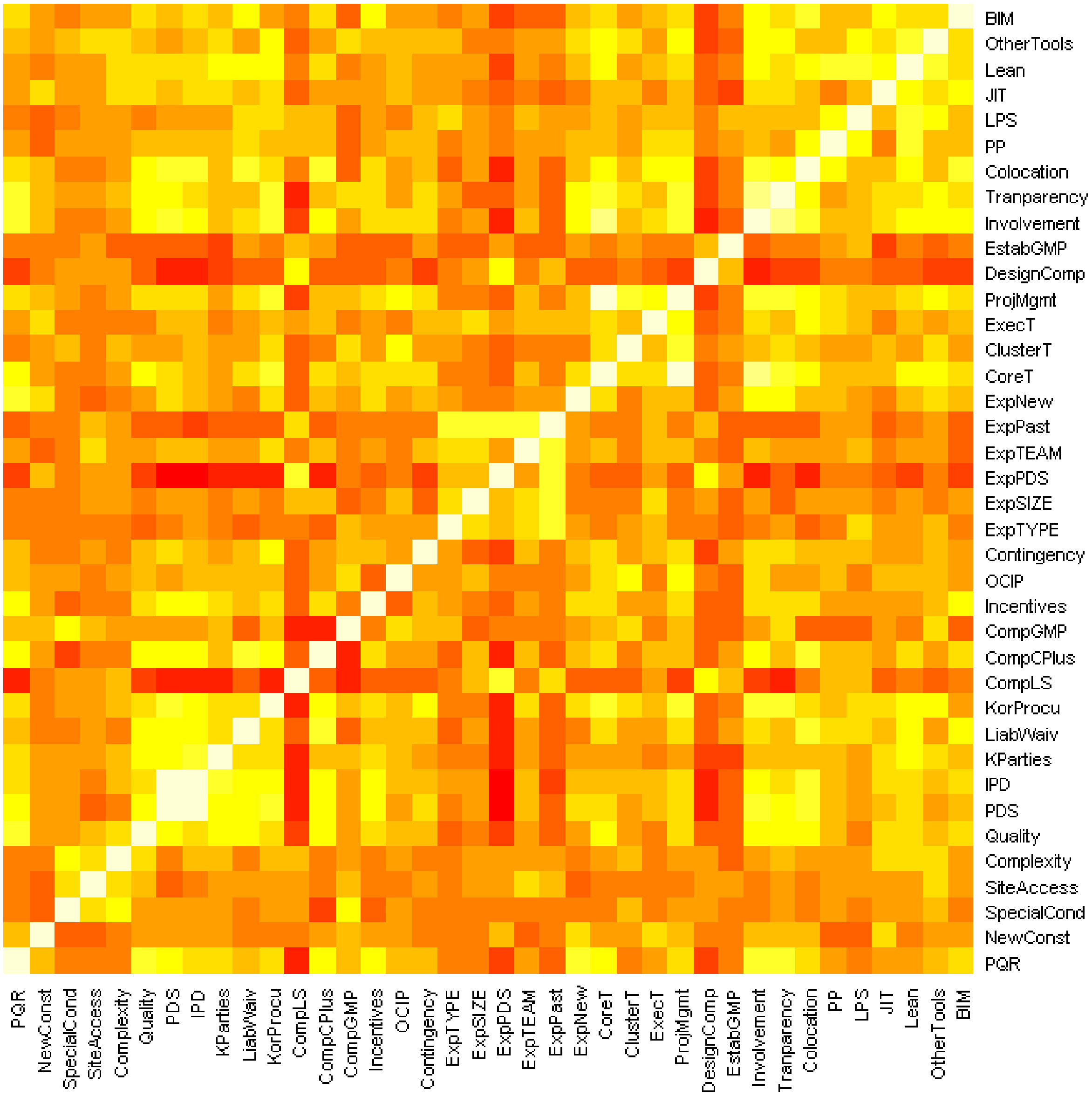


FIGURE 55: Heatmap of Correlations for Delivery Characteristics

Another related concern with stepwise regressions is that the model might omit variables that are important to project performance, but happen to be correlated with variables already included in the model. This issue limits the use of the model to interpret the effect of individual variables, and therefore, the model cannot be viewed as an explanatory model. The final issue is the large number of input variables, or delivery characteristics, which complicate the interpretation of the model and could also affect the prospects of it being used in the AEC industry.

The solution lies in complementing the previous models with another explanatory analysis. As such, principal component analysis satisfies the concerns expressed about stepwise regressions.

6.2 Principal Component Analysis (PCA) of All Inputs

As introduced in *Section 4.1*, Principal Component Analysis (PCA) is a dimension reduction technique for quantitative data. PCA linearly combines the original variables into new variables that are uncorrelated with each other, such that a few of these new variables will explain most of the variation in the original data. When using PCA, multicollinearity is not an issue since the new variables, the principal components, are not correlated. Additionally, the dimension of the dataset is reduced without compromising the information available in the original variables.

In this section, PCA is applied on all the input delivery characteristics, disregarding the output performance metric for the moment. The aim is to create new input variables that summarizes delivery characteristics and can be used later to model performance. Twenty-two combined and individual delivery characteristics were used as original input variables for the analysis. The first six new resulting variables explained about 73% of the variation in the original variables, and the first 11 new variables explained more than 90% of the original variation. *Figure 56* is a scree plot of the variances of each principal component. The scree plot highlights how the first principal component explains a major portion of the variance of the entire dataset. The plot also shows that subsequent components explain decreasing portions of variance.

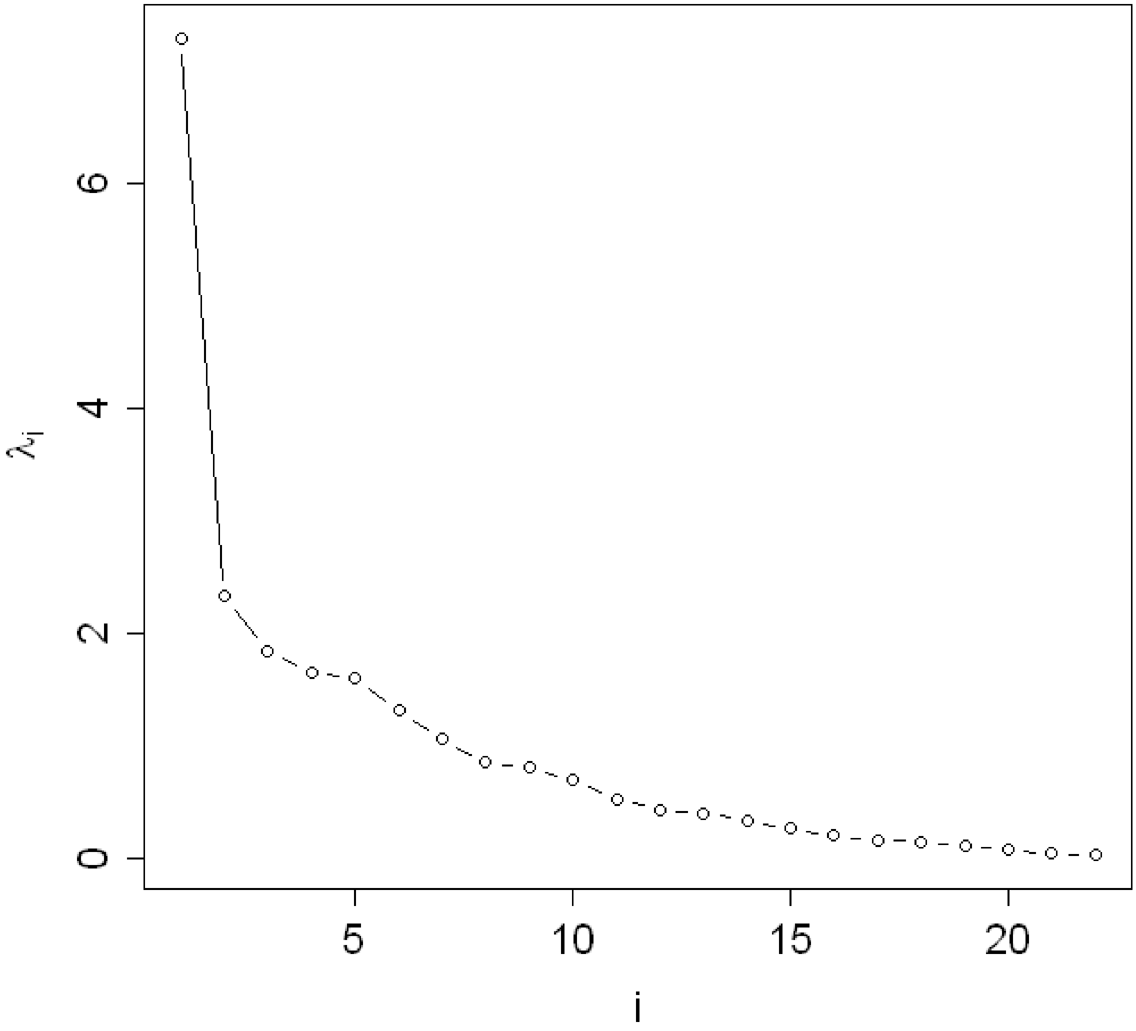


FIGURE 56: Scree Plot of the the PC variances

Table 29 provides more details about the first few components. For each PC, the table presents its standard deviation, its individual proportion of the variance, and its cumulative proportion of the variance when combined with the previous components.

TABLE 29: Components of Combined Delivery Characteristics

Component	Standard Deviation	Proportion of Variance	Cumulative Proportion
1	2.6279	0.3313	0.3313
2	1.5129	0.1055	0.4367
3	1.3380	0.0836	0.5204
4	1.2786	0.0779	0.6010
5	1.1813	0.0665	0.6674
6	1.1250	0.0603	0.7277
…	…	…	…
22	0.1409	0.0009	1.0000

Typically, the first principal component is of considerable importance because it is the single variable that explains the most variation in the dataset. Therefore, the first principal component resulting from this analysis will be discussed. As expected, $PC1_{all}$ is a linear combination of original delivery characteristics. Its equation is as follows:

$PC1_{all}$ = - 0.287 PDS - 0.190 K Parties - 0.218 LiabWaivers - 0.276 Procurement
+ 0.216 Lump Sum Compensation - 0.169 Incentives - 0.159 Contingency
+ 0.263 Experience PDS - 0.157 Experience New - 0.247 Mgmt Structure
+ 0.245 Design Complete - 0.273 Transparency - 0.312 Involvement
- 0.251 CoLocation - 0.132 PP - 0.124 LPS - 0.158 JIT
- 0.245 Added Lean Tools - 0.201 Innovative Tools - 0.191 BIM

Looking closely at the different coefficients, most are negative, with only three exceptions. The corresponding three variables are:

- Experience PDS: a project team very experienced with the particular project delivery system used on the project;
- % Design Complete: a high percent of the design having been already completed at the time when the contractors join the project team; and
- Lump Sum Compensation.

It is interesting that all three variables represent an “anti-IPD” set of delivery characteristics. First, and by definition, in IPD projects the contractors join the project team at a very early stage, often when the design has not even started. Therefore a high percent of design complete would signify that the project cannot be IPD. Second, IPD projects do not typically use lump sum compensation, as discussed in *Section 3.6*. Third, no project team can be overly experienced with IPD because IPD is a recent delivery system. Typical project teams are generally experienced with the traditional DBB system a great deal more than the emerging IPD system.

On the other hand, the characteristics with negative coefficients illustrate IPD projects. The first is the type of project delivery system, followed by a large number of parties signing the same contract. Other IPD characteristics include a strong project management structure as discussed in *Chapter 3*, high fiscal transparency, physical collocation, and a high degree of unconventional participation of all stakeholders in all project phases.

This analysis shows that $PC1_{all}$ contrasts IPD delivery characteristics with non-IPD delivery characteristics. It is interesting that this first principle component reduced the dataset to one very critical dimension by combining and contrasting key delivery characteristics. *Figure 57* shows a representation of the loadings for $PC1_{all}$ through a biplot. The arrows referring to the three anti-IPD variables are pointing to the right, whereas the strong IPD characteristics are pointing to the left. The opposite direction means these variables are negatively correlated, and the length of the arrows shows the levels of correlation between them.

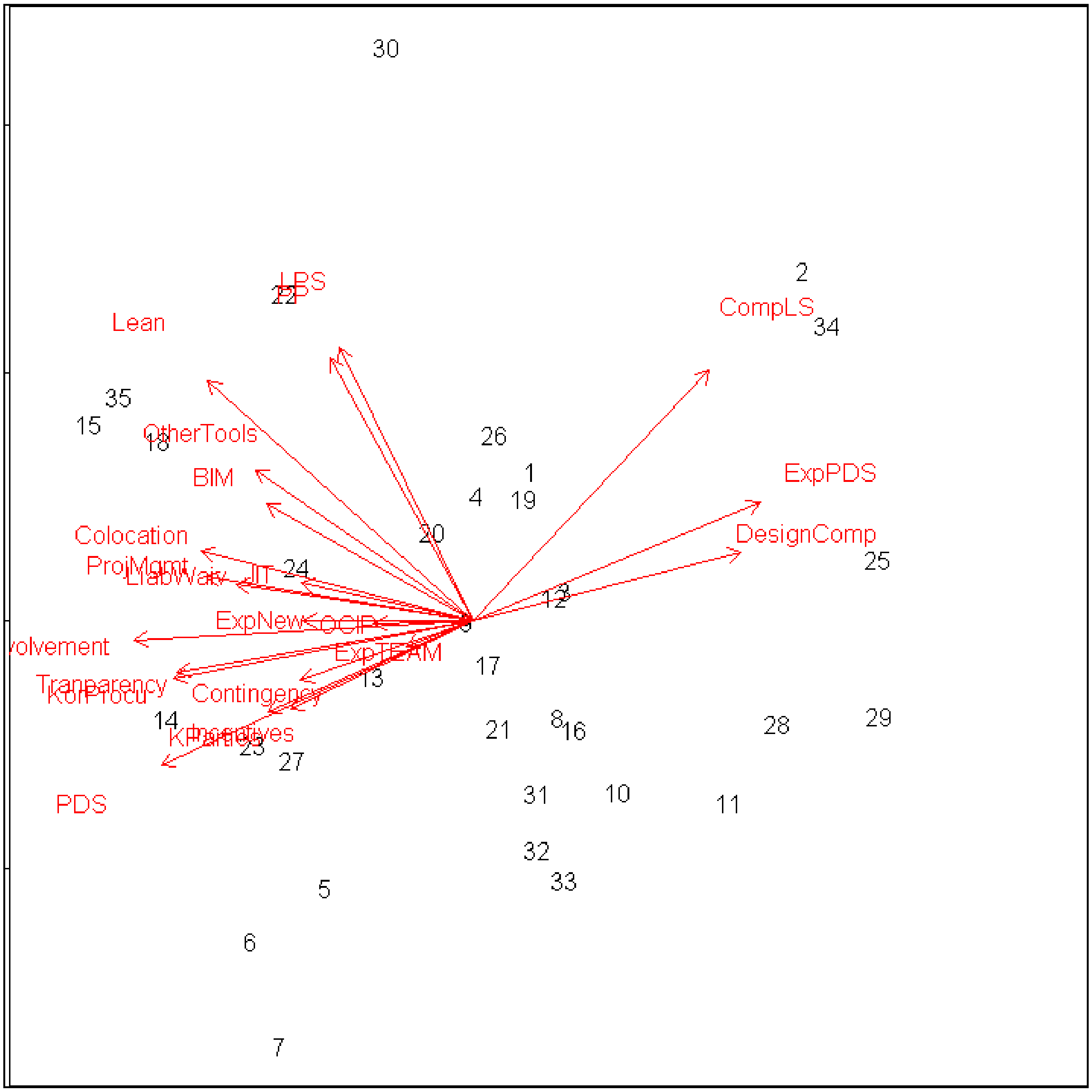


FIGURE 57: Biplot Based on PCA of All Inputs

It is interesting that the variables representing lean construction tools and techniques, especially LPS and PP, are close to orthogonal to the IPD-ness axis, which means there is a fairly weak correlation between IPD and lean construction tools. The biplot also includes the projects used for this analysis. For confidentiality purposes, the project names cannot be shown so the projects are denoted by numbers.

Figure 58 elaborates on the biplot findings by color-coding the IPD projects in green, the non-IPD projects in red, and the few IPD-ish projects in black. This plot validates the interpretation of $PC1_{all}$ and clearly shows all the IPD projects plotted on the left side of the figure, and all non-IPD projects plotted on the right side of the figure.

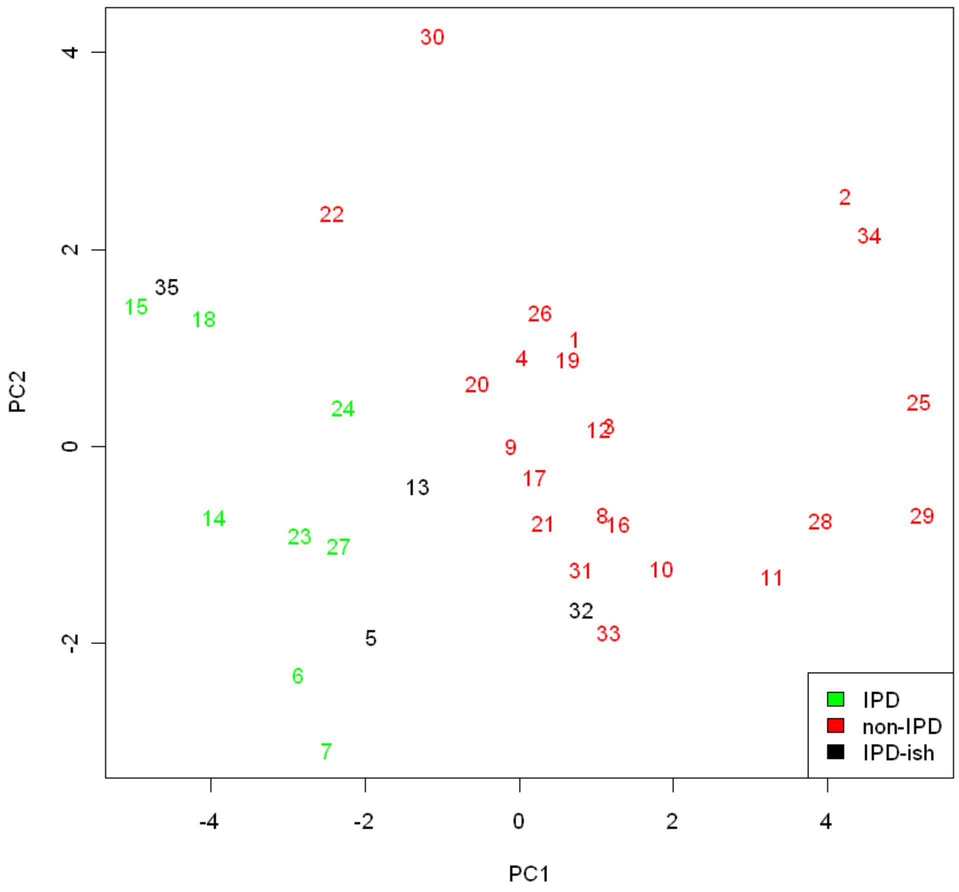


FIGURE 58: $PC1_{all}$ Highlights IPD vs. Non-IPD Characteristics

The remaining components ($PC2_{all}$, $PC3_{all}$, etc…) all have different loadings, meaning they combine different lists of variables with varying coefficients. Not all principal components are interpretable, and PCA does not guarantee that the results will be meaningful. However, in this case, $PC1_{all}$ adequately reduced the dimension of the dataset to one interpretable dimension that explains the variation related to IPD versus non-IPD delivery characteristics.

PCA can be used, as discussed earlier, to provide few new uncorrelated input variables, which allow for the variables in the model and their coefficients to be interpreted. A new model for PQR is developed by using all 22 principal components of the delivery characteristics. A linear regression of the 22 principal components versus PQR gives R^2=0.65 with $PC1_{all}$, $PC2_{all}$, and $PC5_{all}$ showing high statistical significance (at the 0.01 level). Therefore, another similar model is developed only using these three components, resulting in a simpler model with a lower R^2 value of 0.50:

$$\boxed{PQR4 = -0.21\ PC1_{all} - 0.20\ PC2_{all} + 0.26\ PC5_{all}}$$

Now that the PCs are uncorrelated, the model can be interpreted easier. For instance, lower values of $PC1_{all}$ result in high PQR. A previous discussion of $PC1_{all}$ concluded that lower values of $PC1_{all}$ represent IPD projects, while high values represent non-IPD projects. This analysis provides yet another proof of superior project performance for IPD.

$PC2_{all}$ is a new component, and contrasts three variables with all the remaining variables: the project delivery system, the number of parties signing the same contract, and the use of incentives in the contract. Similar to the variables in $PC1_{all}$, these three variables have negative coefficients while the coefficients are positive for all other variables. An analogous interpretation can be made for the effect of $PC2_{all}$ on PQR, providing proof that IPD projects with multiparty contracts and incentives outperform other projects.

$PC5_{all}$ increases with an increased use of BIM. It includes many variables but the BIM variable is given the largest loading. This result proves that the use of BIM causes superior project performance.

This explanatory model provides new findings by identifying specific IPD techniques and additional delivery characteristics that increase project performance, most notably: the use of a multiparty contract with incentives, the use of BIM, the low percent design complete when the contractors join the project, and avoiding lump sum compensation when possible. In this section, PCA was performed on all delivery characteristics combined. As discussed in *Chapter 3*, these delivery characteristics can be separated in three distinct groups: Tone, Terms, and Tools. The following section will discuss PCA as applied to the 3Ts as separate groups to provide further investigation of individual delivery characteristics that can boost project performance.

6.3 PCA Applied on the 3Ts

The previous two sections introduced and discussed both stepwise regressions and PCA for all delivery characteristics combined. The first dealt with correlated input variables, while the second looked at new uncorrelated principal components of the inputs. This section falls in between these two extremes discussed earlier. By looking more closely at each cluster of delivery characteristics, dimension-reduction can be performed for each of the 3T delivery areas separately to identify new variables or components that explain major variations in each of the 3Ts. The identified 3T PCs are uncorrelated within the same delivery area, but correlated between the different 3T areas, and therefore, can be used to model PQR.

6.3.1 PCA Applied on Contractual Terms

Terms make up the first area of the 3Ts and represent the organizational delivery characteristics. There are 11 original variables or delivery characteristics included in this area. Similar to the analysis in *Section 6.2, Figure 59* shows a scree plot of the variances for the *Terms* PCs, and *Table 30* provides more details about the first few components. The first three principal components represent approximately two-thirds of the variation of all the *Terms* delivery characteristics.

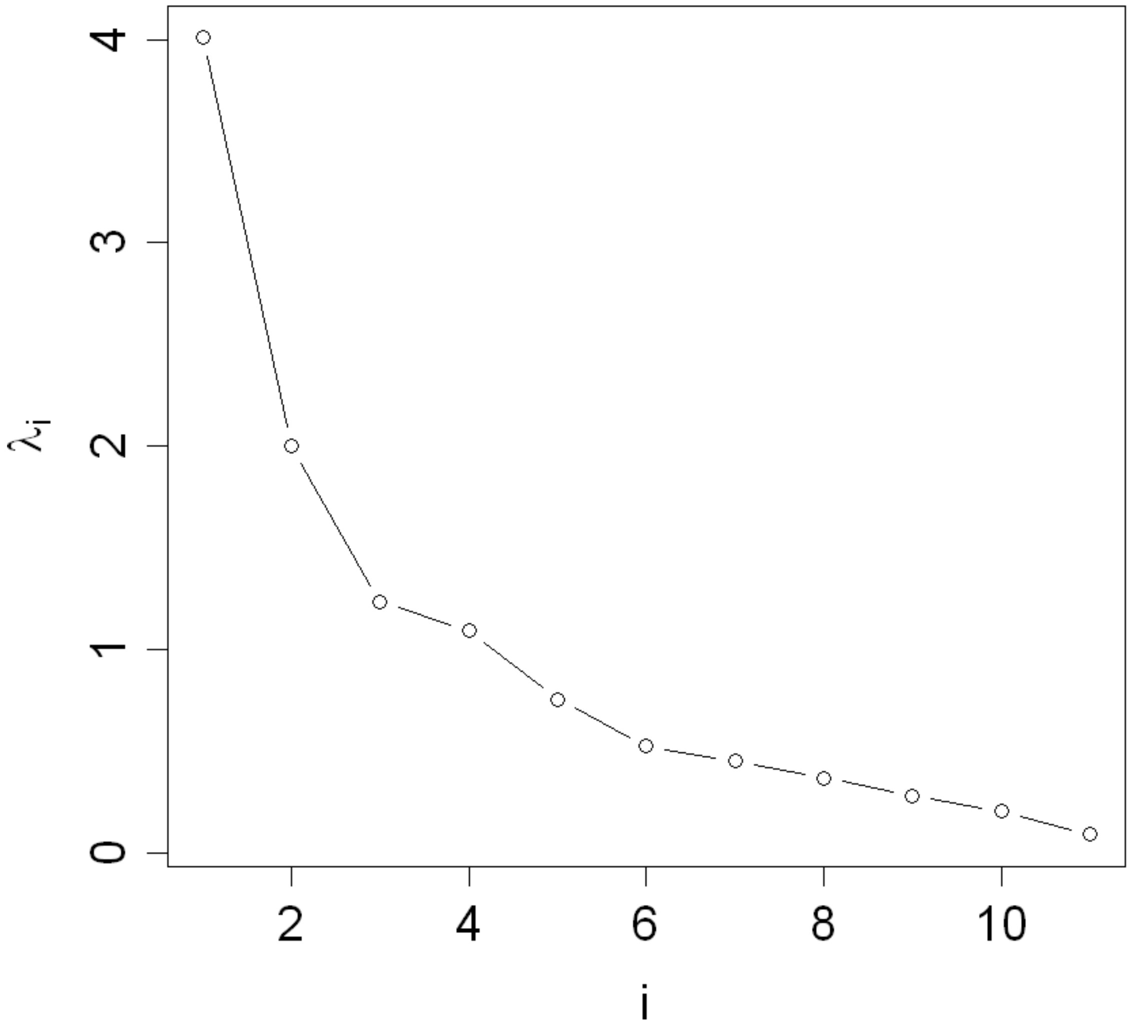


FIGURE 59: Scree Plot for Terms Components

TABLE 30: Variances of Terms Components

Component	Standard Deviation	Proportion of Variance	Cumulative Proportion
1	2.003	0.365	0.365
2	1.414	0.182	0.546
3	1.110	0.112	0.658
4	1.044	0.099	0.758
…			
11	0.300	0.008	1.000

The makeup of each component then is investigated to interpret the few key dimensions embodied in this area. Interestingly, *Table 31* shows how the first principal component for the terms delivery characteristics $PC1_{terms}$ is very similar to $PC1_{all}$ of all delivery characteristics combined, discussed previously in *Section 6.2*. $PC1_{terms}$ again represents the IPD project delivery system and its affiliated characteristics, and contrasts them with the lump sum compensation of contractors and high percentage of design complete when these contractors join the project. The second principal component mainly contrasts Cost Plus compensation and GMP compensation. The third component mostly represents an owner-controlled insurance program (OCIP), while the fourth component highlights the number of parties signing the same contract, as well as the existence of incentives in the contract, contrasting these two items with the amount of contingency included in the contract to account for unknown project risk.

TABLE 31: Terms Component Loadings

Variable	Components				
	1	2	3	4	…
PDS	-0.443				
KParties	-0.315		-0.118	**0.567**	
LiabWaiv	-0.314	0.288	-0.294	-0.147	
Selection	-0.368			-0.260	
Comp_LS	**0.383**	0.280		-0.178	
Comp_Cplus	-0.287	**0.484**	-0.145	-0.105	
Comp_GMP		**-0.663**		0.135	
Incentives	-0.248	0.213	0.395	**0.400**	
OCIP	-0.112	-0.243	**-0.706**	-0.198	
Contingency	-0.230	-0.168	0.445	**-0.559**	
DesignComp.	**0.330**	0.138		-0.127	

The identified *Terms* principal components are used to model PQR. The only statistically significant component (at the 0.01 level) was $PC1_{terms}$. PQR is then modeled against $PC1_{terms}$, resulting in a R^2 value of 0.32. $PC1_{terms}$ has a negative coefficient, which demonstrates that lower $PC1_{terms}$ values (associated with more "IPD-ness") increase PQR.

$$PQR5 = -0.29\ PC1_{terms}$$

6.3.2 PCA Applied on Social Tone

Tone, which represents social delivery characteristics, is the second delivery area. Similar to the previous area, there also are 11 original variables or delivery characteristics included. The scree plot of the *Tone* PCs shown in *Figure 60*, and the variance values shown in *Table 32*, highlight that again the first three components represent approximately two-thirds of the variation of all the *Tone* delivery characteristics.

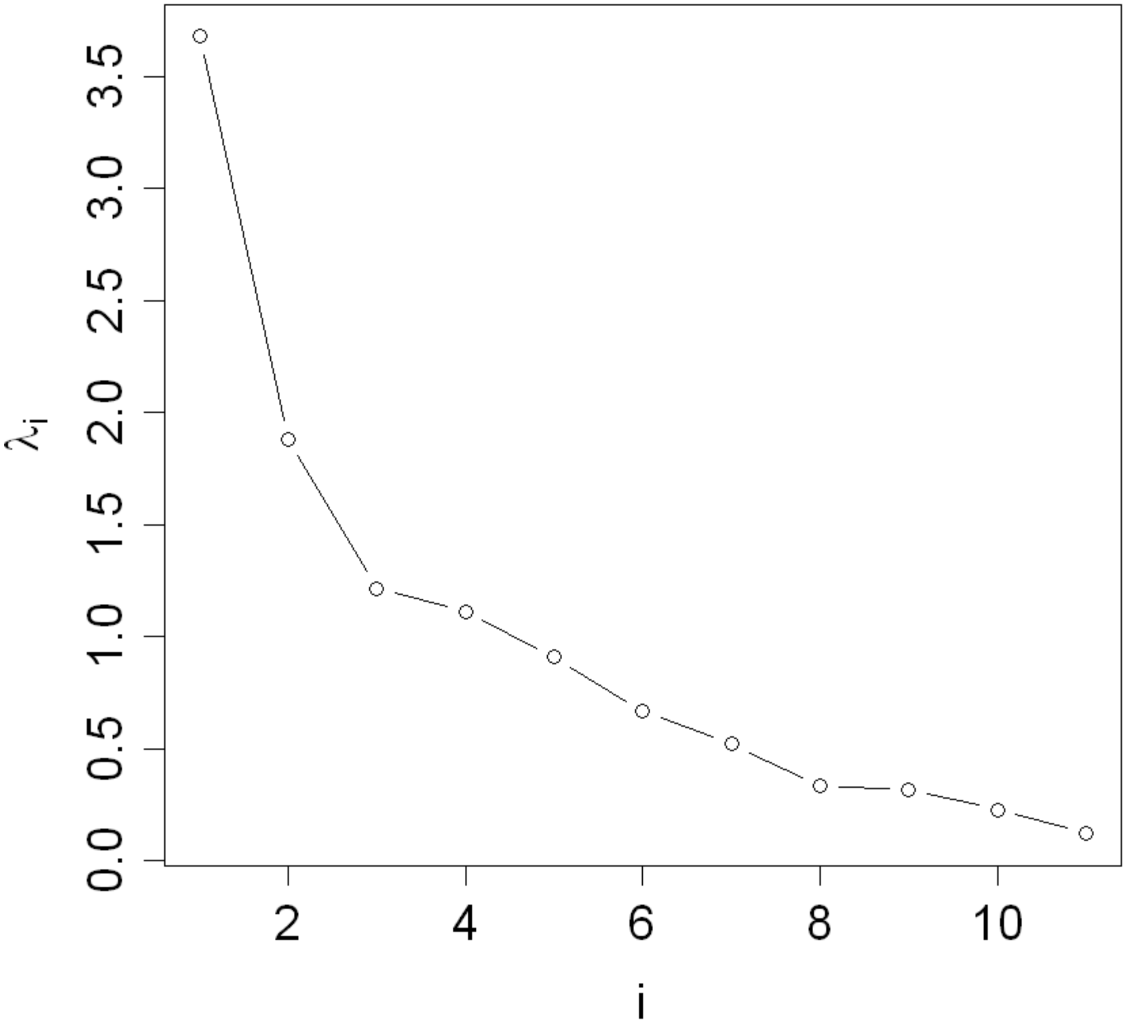


FIGURE 60: Scree Plot for Terms Components

TABLE 32: Variances of Tone Components

Component	Standard Deviation	Proportion of Variance	Cumulative Proportion
1	1.918	0.335	0.335
2	1.371	0.171	0.506
3	1.103	0.111	0.616
4	1.055	0.101	0.717
…			
11	0.350	0.011	1.000

The makeup of these *Tone* components is illustrated in *Table 33*. PQR is then modeled using the *Tone* principal components, two of which were statistically significant: $PC1_{tone}$ and $PC3_{tone}$. When these two significant components are used to model PQR, they alone result in a R^2 value of 0.51.

$$PQR6 = 0.31\ PC1_{tone} – 0.36\ PC3_{tone}$$

TABLE 33: Tone Component Loadings

Variable	Components				
	1	2	3	4	…
Exp. Type		**-0.500**	-0.253	-0.350	
Exp. Size		**-0.544**	0.278		
Exp. PDS	-0.292	-0.229	-0.312	-0.236	
Exp. Team		-0.459		0.251	
Exp. New	0.277		**-0.639**	0.110	
Core Team	**0.415**			-0.159	
Cluster Team	0.205		0.308	**-0.754**	
Exec. Team	0.249	-0.407	0.249		
Involvement	**0.474**			0.154	
Transparency	**0.421**		-0.298	-0.201	
Co-Location	0.380		0.310	0.286	

$PC1_{tone}$ represents the core team as a key part of the project management structure, as well as the active involvement and participation of all major stakeholders along the various project phases, and a high level of fiscal transparency between these stakeholders. Involvement of stakeholders combines several variables that gauge numerous matters, such as the owners’ participation in the construction process, the designers’ support during construction, and the contractors’ participation in the design stages. $PC3_{tone}$, which reflects the new experience of the stakeholders as a team, gauges the chemistry between the CM/GC and the subcontractors on one hand, and the CM/GC and the designers and owners on the other hand.

$PC1_{tone}$ has a positive coefficient while $PC3_{tone}$ has a negative coefficient, meaning higher $PC1_{tone}$ values and lower $PC3_{tone}$ values increase PQR. These results show that a core leadership team and fiscal transparency among the project stakeholders increase overall project performance. The results also provide quantitative proof that a good relationship between the key stakeholders, and their nonconventional participation in all project phases, also increase project performance.

6.3.3 PCA Applied on Tools and Techniques

Tools are the third leg of the delivery characteristics tripod, and represent functional areas, such as technologies and processes used during the project. The several tools are combined into six variables on which PCA is performed. The scree plot of the *Tools* components, shown in *Figure 61*, and the values for variance, shown in *Table 34*, demonstrate that two components represent around 70% of the variation of all the *Tools* delivery characteristics.

As performed for the other two areas, PQR was modeled using the *Tools* principal components. No components were statistically significant at the 0.05 level; however, the p-value of 0.064 for $PC3_{tools}$ was significant at the less stringent 0.10 level, and results in a low R^2 value of 0.07 when used to model PQR. The makeup of the *Tools* principal components is illustrated in *Table 35* and shows BIM is the major factor in $PC3_{tools}$. This result suggests that the use of BIM improves overall project performance; however, because of the weak statistical significance this claim still needs more testing and validation, which will be performed in the next section of this chapter.

$$PQR7 = 0.37\ PC3_{tools}$$

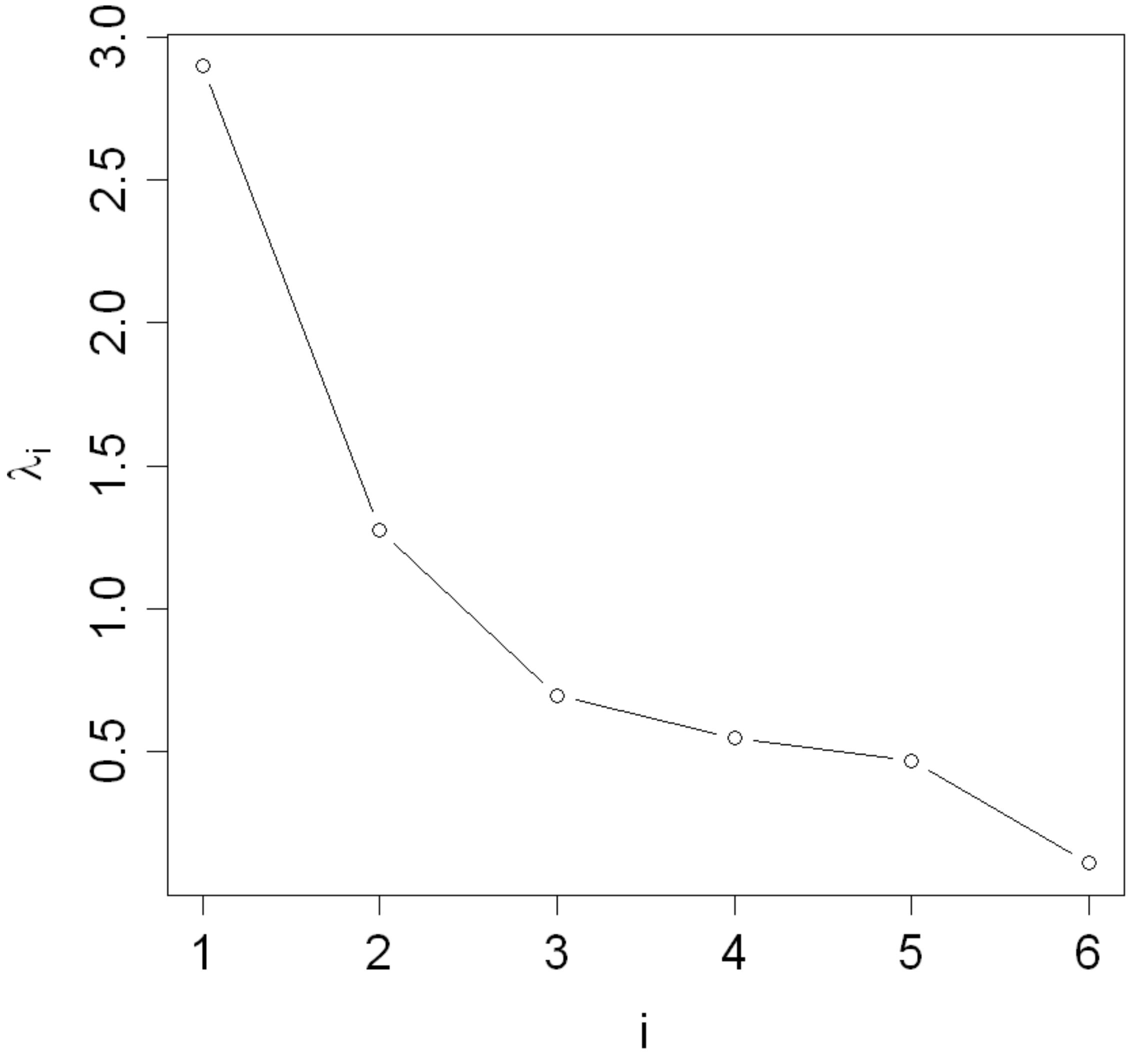


FIGURE 61: Scree Plot for Tools Components

TABLE 34: Variances of Tools Components

Component	Standard Deviation	Proportion of Variance	Cumulative Proportion
1	1.703	0.483	0.483
2	1.130	0.213	0.696
3	0.834	0.116	0.812
4	0.740	0.091	0.903
5	0.685	0.078	0.981
6	0.334	0.019	1.000

TABLE 35: Tools Component Loadings

Variable	Component				
	1	2	3	4	…
Pull Planning	-0.360	0.550	0.422	-0.147	
LPS	-0.423	0.333	-0.343		
JIT	-0.316	-0.627	-0.231	-0.501	
Lean Constr. Tools	-0.539	0.115	-0.128	-0.294	
Innovative Tools	-0.429	-0.161	-0.278	0.785	
BIM	-0.341	-0.391	**0.747**	0.150	

In this section, the project delivery characteristics were categorized with respect to the 3Ts as introduced in *Chapter 3*, then PCA was completed on each delivery area separately to identify the main drivers of overall project performance. The results of the *Terms* characteristics showed once again that IPD projects and their contractual characteristics improve performance as compared to lump sum projects with a high percent of design complete. The results of the *Tone* characteristics showed that a core leadership team and the use of fiscal transparency, a good relationship between the stakeholders, and their involvement throughout all project phases, also increase project performance. Finally, the results of the *Tools* characteristics suggest that BIM increases performance – a result that still needs to be validated.

Additional project information that is not related to the 3T delivery characteristics can provide an important background when studying a project. Some of these project information items include: the percentage of new construction as opposed to additions or renovation, special conditions, such as weather abnormality or labor unavailability for the project, the access to the site (from unlimited to severely restricted), and the complexity of project systems. A PCA was performed for these additional characteristics, and the resulting components were used to model PQR. None of the four components were statistically significant, and therefore this line of investigation was abandoned.

6.3.4 The Spinning 5D Performance Model

In the previous subsections, principal components from each of the 3T delivery characteristics individually explained portions of overall project performance. These components will now be used together to model PQR.

First, PQR was modeled against *all* the components identified earlier for all three delivery areas. The most statistically significant components, at the 0.000 level, were $PC1_{terms}$, $PC3_{tone}$, and $PC3_{tools}$. These components were used to model PQR and resulted in a R^2 value of 0.67, which means the identified three components alone explain more than two thirds of the variation in overall project performance.

$$PQR8 = -0.30\ PC1_{terms} - 0.48\ PC3_{tone} + 0.48\ PC3_{tools}$$

This result validates the 3T hypothesis where all three dimensions are vital for overall project success. The use of IPD (*terms*) and BIM (*tools*), combined with good relationships and chemistry between the key project stakeholders (*tone*), result in increased overall project performance.

Figure 62 illustrates how the PQR8 model (based on only three principal components) compares to the real PQR for all 35 projects in the database. The model, with a R^2 value of 0.67, adequately represents the actual PQR values.

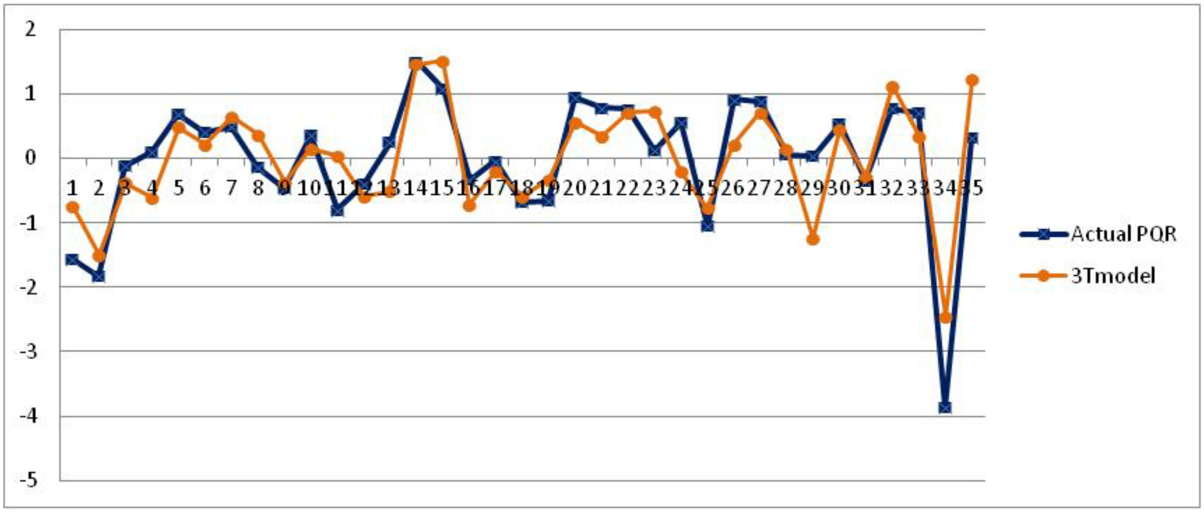


FIGURE 62: Actual PQR vs. PQR Modeled with 3T's

A 3-dimensional model, shown in *Figure 63,* illustrates the key principal components across the three dimensions identified earlier. Color-coding of the PDS adds a fourth dimension to the plot, clearly showing the IPD projects in green scoring high on the *Terms* dimension. One also can see that IPD projects do not necessarily have higher scores in the other two dimensions, *Tone* and *Tools*. The upper and lower parts of the figure show two different angles when rotating the plot. The plane shown in the plot is a regression plane; however, it is not being used as such, but instead provides a better visualization of the 3D plot in this two-dimensional document.

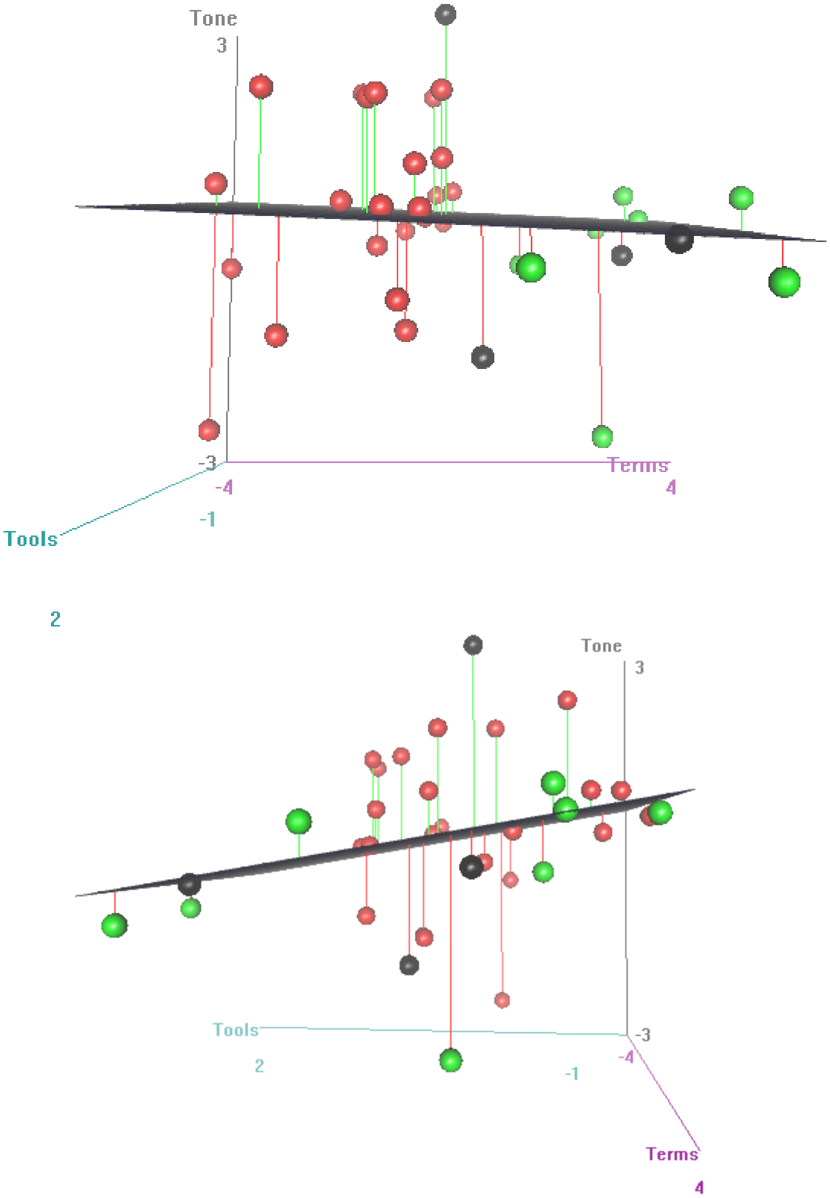


FIGURE 63: 4D Representation of Projects along 3T Components

Figure 64 adds the fifth and arguably most interesting dimension by accounting for the overall performance of the plotted project through their PQR scores. The size of the spheres is representative of the relative PQR of the corresponding projects. The figure shows the projects farthest from the origin of the plot tend to have the largest spheres.

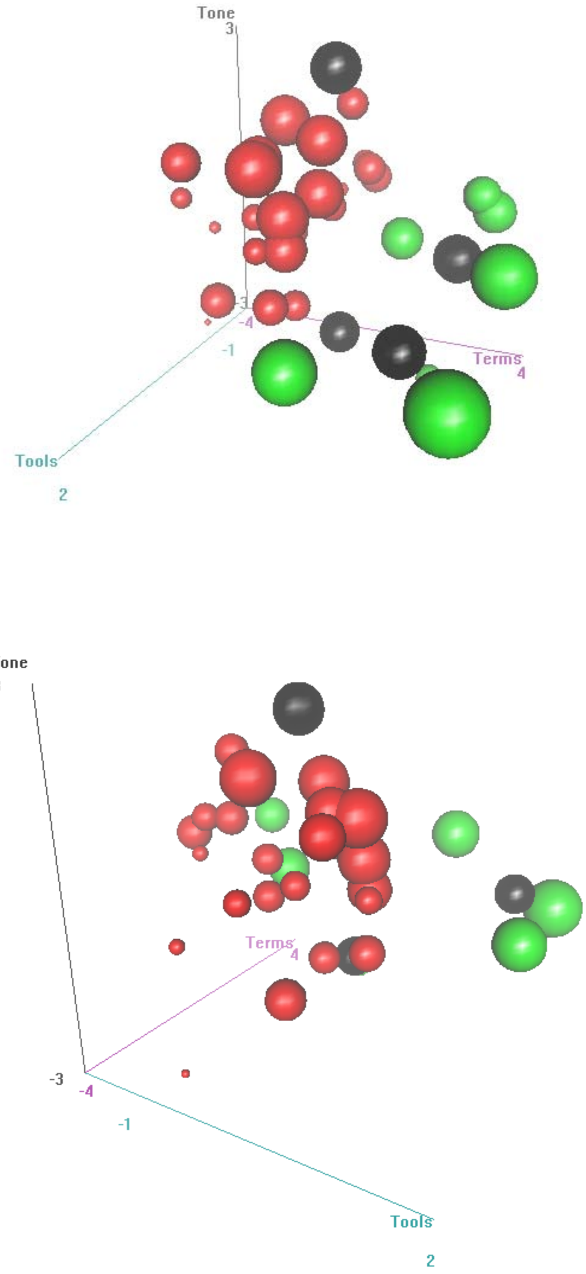


FIGURE 64: 5D Representation of Projects along 3T Components

3D visualization can be difficult on a two-dimensional sheet of paper, and therefore these same results also are presented in 2D. The dependent variable is the PQR and the independent variable is the distance of a given sphere from the origin. The 3T scores are translated to start at the point (0,0,0) and then the norm of each vector is calculated. *Figure 65* shows these scaled distances on the horizontal axis, and their corresponding PQR on the vertical axis. There is a clear positive correlation between the distance of each project from the origin, or its 3T score, and its overall performance PQR.

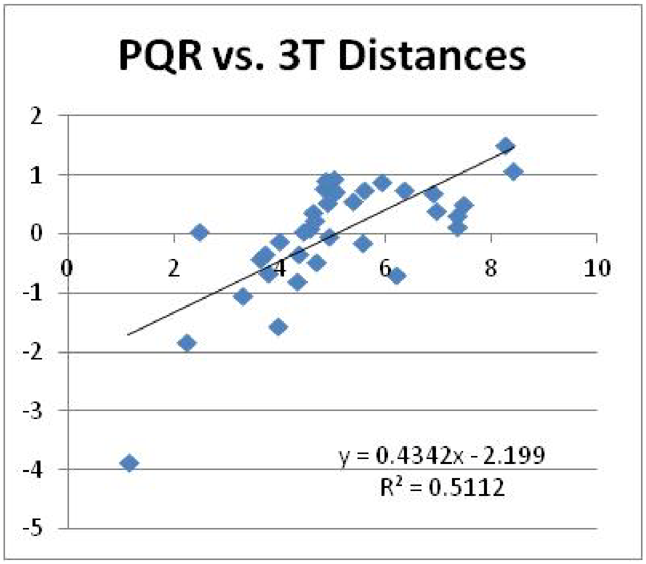


FIGURE 65: Compressing the 3D plot in 2D

6.4 Investigating Individual Delivery Characteristics

The delivery characteristics that are impacting performance as identified in this chapter will now be tested individually versus PQR. The first characteristic is the percent of design completed prior to the involvement of contractors. The upper part of *Figure 66* shows a plot of all 35 projects with the percent of design completed on the x-axis and the PQR on the y-axis. The regression line is trending downwards, meaning a higher percent of design complete results in lower overall project performance, but shows a low R^2 value of 0.13. The bottom portion of the figure shows how the unconventional involvement of stakeholders affects PQR. The regression, with a R^2 value of 0.36, shows that more involvement of key stakeholders in all phases of the project increases performance.

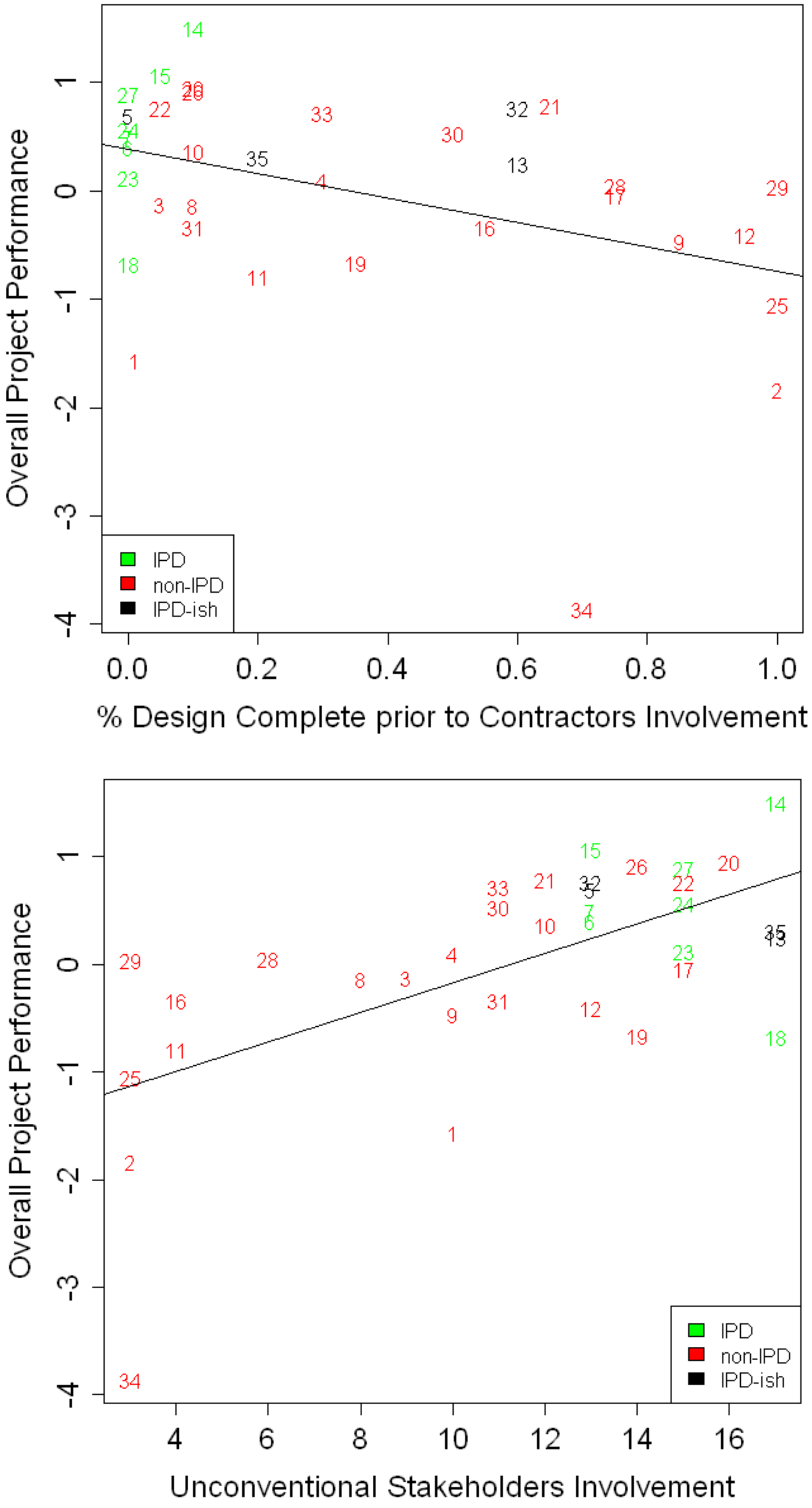


FIGURE 66: PQR versus % Design Complete (top) and PQR versus Stakeholders' Involvement (bottom)

In addition to the key contractors' early involvement in the design and preplanning stages of the project, this analysis proves that the owners' participation in the construction process and the designers' support during construction also have a direct effect on performance. This evidence points to more integrated approaches to project delivery as a potential booster of project performance.

Another delivery characteristic related to stakeholders' involvement is the existence of a project leadership team, the *Core Group*. The *Core Group* score increases with the number of representatives on the leadership team, their meetings frequency, and their authority and responsibilities. As *Figure 67* shows, PQR is increasing with the *Core Group* score, but the linear regression has a low R^2 value of 0.15.

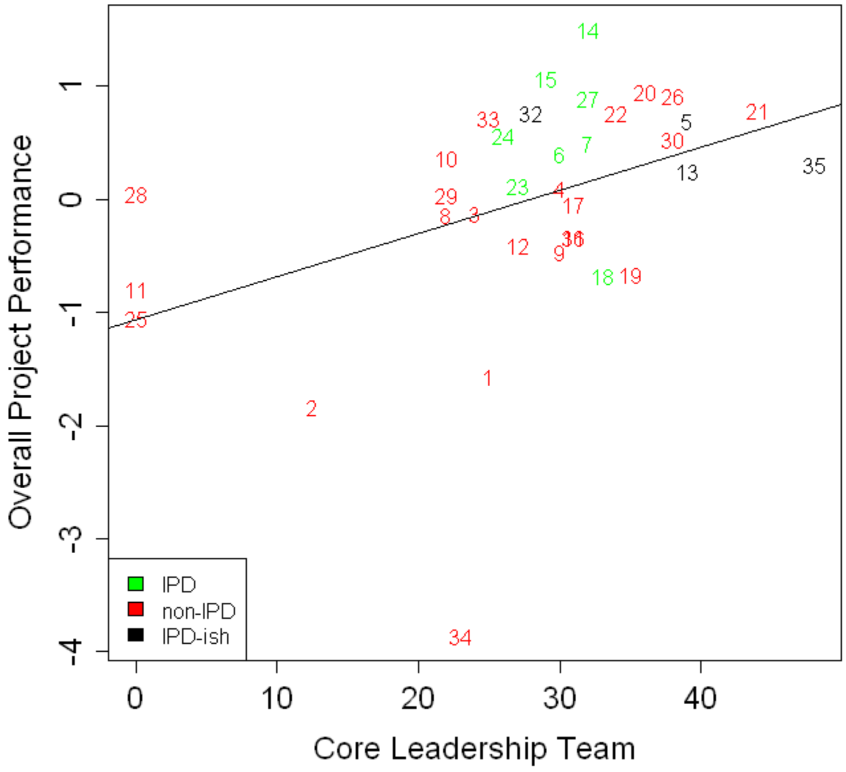


FIGURE 67: PQR versus Core Group

Besides the existence of a structure for project management, which potentially inhibits involvement of stakeholders in all project phases, another closely related characteristic is how stakeholders rate their current experience in terms of satisfaction with the project team. General contractors were asked to rate their experience working as a team with the designers and owners, on one hand, and with the subcontractors on the other hand. This variable, which is a proxy for chemistry among project stakeholders, showed a R^2 value of 0.39 when modeling PQR regardless of the project delivery system used, as shown in the upper part of *Figure 68*.

The bottom part of *Figure 68* shows how the use of BIM processes and tools affect PQR. The line also exhibits a positive slope, signifying BIM increases performance; however, with a R^2 value of 0.12, the relationship is not as strong as with any of the previous variables. Therefore, it appears as though the tools and processes increase overall project performance, but contractual characteristics, and especially those that impact the social characteristics of the project, are even more effective at increasing overall project performance.

The last part of this section tests the effects of fiscal transparency and financial compensation of stakeholders on overall project performance. The upper part of *Figure 69* shows that lump sum compensation generally results in decreases in PQR, and therefore is not an ideal method to compensate stakeholders. The regression showed a R^2 value of 0.31, which is interestingly high considering that this is only one small aspect of project delivery characteristics. Other methods, such as cost plus, GMP, and negotiated fees, can be used instead.

The bottom part of *Figure 69*, with a R^2 value of 0.42, shows that the use of fiscal transparency between project stakeholders generally leads to an improved overall performance. This technique was used to its fullest extend in all IPD and IPD-ish projects in the database, where all project costs were fully transparent to all key stakeholders.

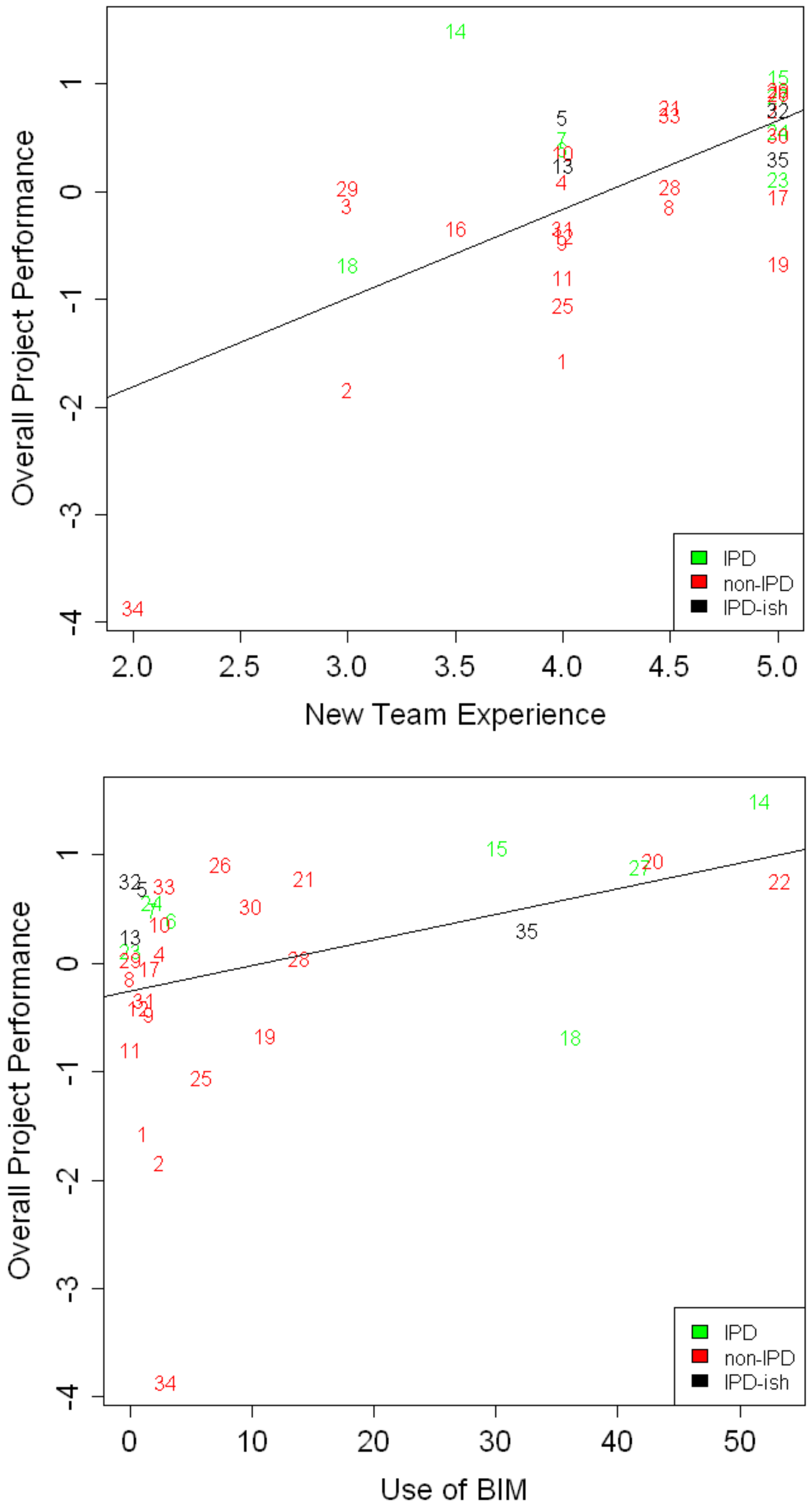


FIGURE 68: PQR versus New Team Experience (top) and PQR versus BIM (bottom)

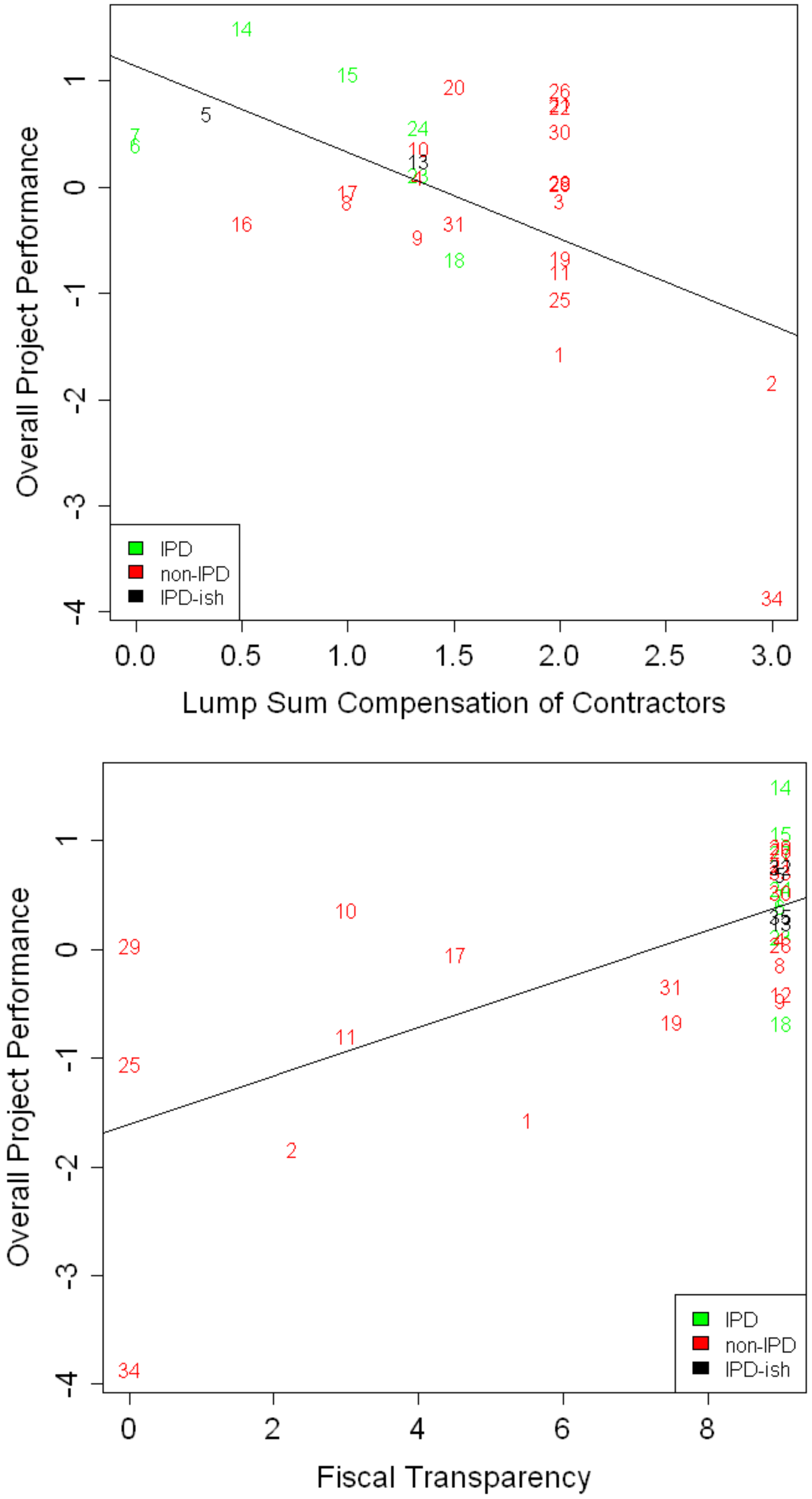


FIGURE 69: PQR versus Lump Sum Compensation (top) and PQR versus Fiscal Transparency (bottom)

Some identified key characteristics had binary responses, and therefore MWW and t-tests were used to examine their individual effect on performance. The three characteristics investigated are: (1) the use of an Owner Controlled Insurance Program (OCIP), (2) the signing of liability waivers between project participants, and (3) the existence of incentives in the contract. For each of these characteristics both MWW and t-tests were conducted. The p-values of both tests were comparable.

The hypothesis that an owner-controlled insurance program results in an increased overall performance is not supported by the tests, which showed a p-value of 0.108. The hypothesis that liability waivers result in increased performance is not supported at the 0.05 level, but its p-value of 0.077 can be considered small enough for the less stringent 0.10 level. Finally, the third hypothesis that the use of incentive clauses increase performance is strongly supported by the tests, which resulted in a p-value of 0.000. This last finding provides quantitative answers that encourage project owners to share savings and unused contingency dollars with project stakeholders.

This section provided a solid addition to the findings of this chapter by individually testing the results of the multivariate analyses. Highlights of the results include: key social characteristic, such as the stakeholders’ unconventional participation in project phases and the current experience and chemistry of the project team. Additional key characteristics are lump sum compensation and the use of incentive clauses. These delivery characteristics, and others that were identified in the previous sections, were tested against overall project performance, and then the results and their implications were discussed.

Chapter 6 offered several contributions to the project delivery literature. In addition to the contributions of the previous chapters, this chapter highlighted individual characteristics that can be used with or without IPD to increase project performance. The use of these delivery techniques can offer great advantages, especially in areas where the use of IPD is restricted by law, such as public sector projects.

Chapter 7. Conclusion

7.1 Summary of Research Methods

In this study, IPD performance was studied quantitatively and comprehensively. Univariate analyses were performed to determine which individual performance metrics are positively affected by IPD. The project quarterback rating was develoepd to integrate key performance metrics and was later used to highlight IPD improvements on overall project performance. Multivariate analyses established the effect of a number of IPD-related delivery characteristics on overall project performance.

The results of this methodology led to distinct contributions to the body of knowledge. The following sections will provide a summary of these contributions along with the key results of this research, a discussion of the limitations of this study and the barriers to widespread IPD implementation, as well as a call for future research.

7.2 Summary of Results and Contributions

The six specific objectives set out at the beginning of this dissertation were met and provide the following contributions to the AEC industry:

1) <u>IPD Definition</u>: In conformance with the definition of a project delivery system, a standardized definition of *Integrated Project Delivery* is provided based on the data collected. IPD is defined as *a project delivery system distinguished by a multiparty agreement and the very early involvement of the key participants, ideally at 0% design but definitely before 10% design complete*. This definition

can now be used as a standard in future studies to provide a common basis of comparison.

2) IPD Superior Performance: The performance of IPD projects is superior when compared to that of other delivery types based on 14 different metrics categorized in six performance areas: project quality, schedule, changes, recycling, communication, and profit. Part of this contribution is highlighting key performance metrics other than cost and time, which also have large effects on overall project performance. IPD projects exhibit a statistically significant increase in quality metrics, as measured by the quality of major building systems, the number of deficiency issues, the number and cost of punchlist items, and the cost of warranty. These quality improvements are not associated with any significant cost premiums. Regarding schedule performance, IPD projects also are delivered faster, experiencing statistically significant increases in delivery speed as compared to other delivery systems. This is a major contribution because faster delivery results in earlier beneficial use of the facility, which could translate into major cashflow revenues for the customer. Another key improvement of IPD projects relates to change performance, most notably significant reductions in design-related changes and regulatory changes, and a reduction of change order processing times to an astonishing one week; the median for non-IPD projects is four weeks. Communication was also considerably improved with IPD projects, which experienced significantly less resubmittals and RFIs, as well as faster processing times for the latter. In addition to redusing process waste in terms of

deficiency issues, changes, resubmittals, and processing times, IPD projects also generate less material waste. Finally, not just owners are ending up with a higher quality facility; contractors as well seem to be making more profit on IPD projects, which appear to offer a win-win situation when applied properly on high-complexity projects. These findings affirm that IPD significantly reduces project waste and offers potential benefits for complex institutional projects, and is worth the use, research and investment. Stakeholders of such projects should consider using IPD to take advantage of these benefits.

3) PQR Development: A comprehensive project success rating was developed and validated, combining key performance metrics identified by the industry respondents. The *Project Quarterback Rating* (PQR) complemented the previous contribution by allowing for comparisons of overall performance for IPD and non-IPD projects. IPD projects again saw a statistically significant increase in performance based on this newly-developed metric. A key finding was achieved when non-IPD projects were split in DBB, CMR, and DB samples, and the more integrated systems showed significantly higher performance.

4) Indentification of Key Delivery Characteristics: delivery characteristics responsible for the superior overall performance of IPD projects were identified through multivariate analyses. Several delivery characteristics, grouped under contractual terms, social tone, and tools and techniques (3Ts), were investigated to understand characteristics that ultimately lead to superior performance. The *Terms* results demonstrate once again that IPD projects and their contractual

characteristics, including a low percent of design complete and contractual incentives, improve performance as compared to lump sum projects with a high percent of design complete. *Tone* results show that a core leadership team, fiscal transparency, chemistry between project stakeholders, and their active and constructive involvement throughout all project phases, all increase performance. Finally, *Tools* results show that the extensive use of BIM also increases project performance. All the above delivery characteristics identified can be used to improve performance of IPD and non-IPD projects alike.

5) <u>Models Development</u>: The development of performance models assists the AEC industry in understanding and assessing project performance. The performance models are based on principal components of the project delivery characteristics (first combined, then separated along the 3Ts). Additionally, the combination of explanatory models into a 5D representation highlights the effects of the 3T delivery characteristics on overall project performance.

6) <u>Recommendations</u>: All the previous analyses and findings can be used to provide recommendations and benchmarks for the emerging IPD system. Key stakeholders should be actively involved in all the phases of a project as opposed to only being involved based on traditional requirements. Specifically, contractors need to be involved as early as possible in the design phase. Team members also need to be conscious of how their social interactions impacts project performance. Therefore, it is imperative to maintain healthy working relationships among stakeholders. Lump sum compensation should be avoided when possible because

it is associated with poor project performance. Conversely, it is recommended that stakeholders strive for utilization of incentive clauses and multiparty contracts because these are statistically coupled with superior performance.

7.3 Research Limitations

Although these findings provide strong support for IPD, they do not imply that IPD is the magic solution to all project delivery issues. As discussed in *Section 1.5*, it is difficult to collect data from all key stakeholders of AEC projects, and therefore the CM/GC were targeted since they typically have access to most of the project data needed for the research effort. Therefore, the research was conducted from a CM/GC point of view.

IPD is a recent development and there are not many IPD projects that are completed and available for study. Additionally, some CM/GC consider IPD to be their trade secret and are reluctant to share information regarding IPD performance so they can maintain their competitive advantage. As a result, the dataset used for this research included 35 projects, which is smaller than some of the studies discussed in the *Literature Review*. When extrapolating the findings of this research, the relatively small dataset (relative to the total number of comparable projects delivered in the same timeframe) should be taken into account. This said, it should also be known that this is the largest dataset of IPD projects available to date.

Furthermore, most projects delivered using IPD are large complex institutional projects, such as healthcare facilities and university research laboratories, which benefit highly from innovations in multi-trade settings. Consequently, this study was mostly limited to these types of projects in order to perform fair comparisons of projects with comparable levels of complexity. Therefore, although one can learn a great deal from the findings of this research, it is not recommended to draw conclusions for other types of facilities that were not included in the dataset, such as schools or residential construction.

7.4 Barriers to IPD Implementation

Even though IPD and other delivery characteristics have been proven to quantitatively boost facility performance, challenges remain in implementing these findings throughout the industry. The professional *Cognitive Impairments* and institutional *Adaptive Challenges* previously discussed in *Section 3.1.1* are major barriers to industry-wide implementation of innovative ideas and techniques. Another potentially problematic area includes the issues of risk allocation and insurance in the IPD no-claim no-fault environment, especially when liability waivers are used. In fact, the insurance market for IPD has not yet matured (ENR 2010) and IPD has yet to be tested in courts.

Once the above hurdles are overcome, there will remain at least one major barrier before the industry-wide adoption of IPD: the issue of fairness when using public funds. The public sector has always lagged behind the private sector in terms of adopting new collaborative delivery techniques. One example is the use of design-build (DB) on transportation projects. Even after several years of DB implementation in public projects, protests are still common when the lowest bidder is not chosen, due to the potential for misuse of taxpayers’ money. IPD projects, like DB projects, are not procured based on the lowest bidder because the design is incomplete when contractors join the team. Previous research shows that having a fair and open team selection process is the most important factor for public project success (El Asmar 2009). Additionally, mathematical models have been developed to reassure taxpayers that the selection methodology is fair in public projects (El Asmar et al. 2010 and 2012). The use of such methods can help IPD overcome one of the hurdles that previously prevented fast DB implementation in public projects.

7.5 Future Research

This study provides the first ever quantitative assessment of IPD performance. Several follow-up research studies can be conducted to further confirm the findings or build upon the results of this study.

Given that the data collected for this study is limited to a CM/GC perspective, similar studies could target owners, designers, and subcontractors in order to get a more comprehensive picture of project performance from several stakeholders' perspectives. When targeting other stakeholders, this study and the developed comprehensive survey in available in *Appendix C* can be used as a framework for data collection and analysis to replicate the research for different stakeholders. Additionally, facility types other than complex institutional projects can be targeted once a reasonable amount are delivered using IPD and available for data collection.

When more IPD projects are completed and larger datasets gathered, a similar study can be conducted in order to confirm this study's findings and draw broader conclusions regarding IPD performance. An interesting outcome of such an effort is to provide another snapshot at IPD performance to discover whether this performance is changing over time when stakeholders become more experienced with the new delivery system.

Addressing these prospective research areas will culminate in a better understanding of IPD, its performance, and its related characteristics. This understanding can help further increase future AEC project performance.

7.6 Final Remarks

IPD offers a superior performance as compared to other delivery systems. This study offer key contributions to the delivery systems literature, most notably by presenting a standardized definition of IPD and establishing statistically significant findings that IPD projects are superior in 14 different performance metrics. Another key contribution is combining key performance metrics in the *Project Quarterback Rating* (PQR) to complement the previous contribution by showing that IPD projects experience a statistically significant increase in overall project performance. One more key contribution is the development of explanatory performance models, which established Terms, Tone, and Tools (3Ts) as key clusters of delivery characteristics that significantly impact AEC project performance.

These major contributions allowed for the development of IPD benchmarks in terms of IPD’s effects on individual performance metrics, as well as recommendations regarding which delivery characteristics are proven to increase performance. The research shows the emerging IPD system and several of its associated delivery characteristics offer superior performance and reductions in waste for complex projects, and therefore are worth the investment of resources by project teams. Major project stakeholders see several benefits related to IPD implementation. This emerging delivery system is very promising and demonstrates that the right path to solving some of the AEC industry woes starts with more industry integration. Many more efforts are needed to solve the major hurdles and inefficiencies in the industry; IPD is only the beginning.

Appendix A – BIBLIOGRAPHY

1. AIA - California Council. (2007). Integrated Project Delivery: A Guide. Sacramento, CA: AIA California Council.
2. AIA - California Council. (2007). Integrated Project Delivery: A Working Definition. Sacramento, CA: AIA California Council.
3. AIA - California Council. (2009). Experiences in Collaboration: On the Path to IPD. Sacramento, CA: AIA California Council.
4. AIA - California Council. (2010). Integrated Project Delivery: Case Studies. Sacramento, CA: AIA - California Council.
5. American Institute of Architects. (2011). “IPD case studies.” AIA, AIA Minnesota, School of Architecture – University of Minnesota.
6. Babbie, E. (2010). *The Practice of Social Research*, 12th edition. Wadsworth Publishing, Cengage Learning. Belmont, CA.
7. Ballard, G. (2008). "The lean project delivery system: an update." *Lean Construction Journal*, 4, 1–19.
8. Bartholomew, D., and Knott, M. (1999). *Latent Variable Models and Factor Analysis*, 2nd Edition. London: Edward Arnold.
9. Bennett, J., Pothecary, E., and Robinson, G. (1996). "Designing and Building a World-class Industry: The University of Reading Design and Build Forum Report." Centre for Strategic Studies in Construction, University of Reading, Reading.
10. Branca, A. J., and Linde, C. W. (1987). *Cost effective design/build construction*. RS Means Co., McGraw-Hill, New York.
11. Bureau of Labor Statistics (BLS 2012). “How to Compute a Firm’s Incidence Rate for Safety Management.” Injuries, Illnesses and Fatalities, <http://www.bls.gov/iif/osheval.htm> (October 20, 2011).
12. California State University, Fresno. Craig School of Business. "Sports Marketing Option.” Accessed at <http://www.craig.csufresno.edu/departments/mrkt/Sports-mktg-snapshot.asp> (July 11, 2012).

13. Chan, A. P. C., Scott, D., and Lam, E. W. M. (2002). "Framework of Success Criteria for Design/Build Projects." *J. Constr. Eng. Mgmt*, 18(3), 120-128.
14. Chan, A. P. C., Scott, D., and Chan, A. P. L. (2004). "Factors affecting the success of a construction project." *J. Constr. Eng. Mgmt*. 130(1), 153-155.
15. Cho, S., Ballard, G., Azari, R., Kim, Y. (2010). “Structuring ideal project delivery system." *Proceedings of IPPC4, 2010.*
16. Cho, S., and Ballard, G. (2011). “Last Planner and Integrated Project Delivery.” *Lean Construction Journal*. Lean and Integrated Project Delivery special issue. 67-78.
17. CURT (2004). “Collaboration, Integrated Information and the Project Lifecycle in Building Design, Construction and Operation.” WP-1202.
18. CURT (2007). “Construction Strategy: CURT’s Path toward LEAN Project Delivery.” WP-1004A.
19. Construction Industry Institute (2001). “Quantifying the cumulative impact of change orders for electrical and mechanical contractors.” Research Report 158-11.
20. Construction Industry Institute (2005). “Lean principles in construction.” Research Summary 191-1.
21. Construction Industry Institute (2007). “Roadmap for lean implementation at the project level.” Ballard, G., Kim, Y., Jang, J., Liu, M. Research Summary 234-11.
22. Construction Industry Institute (2011). “Starting from scratch: a new project delivery paradigm.” Ballard, G., Kim, Y., Myers, C., Rogers, G., Strickland, J., Taylor, W., and Voll, F. Report 271-1.
23. Cox, T. F., and Cox, M. A. A. (2000). *Multidimensional Scaling,* 2nd Edition. Boca Raton: Chapman & Hall/CRC.
24. Cox, T. F. (2005). *An Introduction to Multivariate Data Analysis*. Hodder Arnold, London, UK.
25. Debella, D. C., and Ries, R. (2006). "Construction delivery systems: A comparative analysis of their performance within school districts." *J. Constr. Eng. Mgmt*. 132, 1131.
26. Diekmann, J. E. et al. (2004). “Application of Lean Manufacturing Principles to Construction.” CII Research Report 191-11.CII, Austin, TX.

27. Eckblad, S., Rubel, Z. and Bedrick, J. (2007). “Integrated Project Delivery: What, Why and How.” American Institute of Architects (AIA) Convention. May, 2, 2007. AIA 150. Growing Beyond Green.
28. El Asmar, M. (2009). “*Quantitative methods to support decision-making for design-build team selection in highway state agencies*.” MS Thesis, University of Wisconsin-Madison.
29. El Asmar, M., Lotfallah, W., Whited, G., and Hanna, A.S. (2010). “Quantitative Methods for Design-Build Team Selection.” *J. Constr. Eng. Mgmt.* 136 (8), 904-912.
30. El Asmar, M., Lotfallah, W. B., Loh, W.-Y., and Hanna, A. S. (2012). “Uncertainty Reduction in Multi-Evaluator Decision Making.” *Journal of Computing in Civil Engineering*, 26(1), 105.
31. Elkington, J. (1998). Cannibals with forks: The triple bottom line of 21st century business. New Society Publishers, Gabriola Island, BC, Canada
32. Engineering News Record (2010). Post, N. “Integrated-Project-Delivery Boosters Ignore Many Flashing Red Lights.” May 5, 2010.
33. Engineering News Record (2011). Post, N. “Pioneers push paradigm shift.” September 19, 2011.
34. Engineering News Record (2012). Construction Cost Index History. <http://enr.construction.com/economics/historical_indices/construction_cost_index_history.asp> (May 31, 2012).
35. Forbes, L. & Ahmed, S. (2011). *Modern construction: Lean Project Delivery and integrated practices*, Taylor & Francis Group, Boca Raton, FL.
36. Hanna, A.S., and Gunduz, M. (2004). “Impact of Change Orders on Small Labor-Intensive Projects.” *J. Constr. Eng. Mgmt*. 130 (5), 726-733.
37. Hanna, A.S., Camlic, R., Peterson, P.A., and Lee, M. (2004). “Cumulative Effect of Project Changes for Electrical and Mechanical Construction.” *J. Constr. Eng. Mgmt.* 130 (6), 732-771.
38. Hanna, A.S., and Swanson, J. (2007). “Risk Allocation by Law – Cumulative Impact of Change Orders.” *Journal of Professional Issues in Engineering Education and Practice*, 133 (1), 60-66.

39. Hanna, A.S. (2010). Construction Labor Productivity Management and Methods Improvement. ISBN-13: 978-0-9829042-0-6.
40. Hanna, A.S. (2011). CEE 498 Class Notes. Construction Project Management.
41. Horman, M. J., and Kenley, R. (2005). "Quantifying levels of wasted time in construction with meta-analysis." *J. Constr. Eng. Mgmt*. 131 (1), 52-61.
42. Ibbs, C. W., Kwak, Y. H., Ng, T., and Odabasi, A. M. (2003). "Project delivery systems and project change: quantitative analysis." *J. Constr. Eng. Mgmt*., 129, 382.
43. Ireland, V. (1984). “Virtually Meaningless Distinctions between Nominally Different Procurement Forms”. CIB W65 Proceedings of the 4th International Symposium on Organisation and Management of Construction, Vol. 1, University of Waterloo, Waterloo, Ontario, pp. 203-212.
44. Jolliffe, I. T. (2002). *Principal Component Analysis*. New York: Springer-Verlag.
45. Kim, Y. & Dossick, C. (2011). “What makes the delivery of a project integrated? A case study of Children’s Hospital, Bellevue, WA.” *Lean Construction Journal*, Lean and Integrated Project Delivery Special Issue, 53-66.
46. Konchar, M., and Sanvido, V. (1998). "Comparison of U.S. project delivery systems." *J. Constr. Eng. Mgmt*., 124, 435.
47. Korkmaz, S., Riley, D., and Horman, M. (2010). "Piloting Evaluation Metrics for Sustainable High-Performance Building Project Delivery." *J. Constr. Eng. Mgmt*. 136, 877.
48. Lean Construction Institute. (2011). “What is Lean Construction?” <http://www.leanconstruction.org/whatis.htm> (Nov. 30, 2011).
49. Lehmann, E. L. (2006). “Nonparametrics: Statistical Methods Based on Ranks,” Springer. ISBN 10: 0-387-35212-0.
50. Lichtig, W. A. (2005). "Sutter Health: Developing a Contracting Model to Support Lean Project Delivery." *Lean Construction Journal*, 2(1), 105-112.
51. Lichtig, W. (2007). “Creating a Relational Contract to Support Lean Project Delivery.” Presentation to the Lean Construction Institute Relational Contracting Meeting. Chicago, IL. June 14-15.
52. Liker, J. K. (2004). *The Toyota Way*. McGraw-Hill, New York.

53. Love, P. (2002). "Influence of project type and procurement methods on rework costs in building construction projects." *J. Constr. Eng. Mgmt*. 128 (1), 18-29.
54. Matthews, O., and Howell, G. A. (2005). "Integrated Project Delivery: An Example of Relational Contracting." *Lean Construction Journal*, 2(1), 46-61.
55. Menassa, C., Mangasarian, S., Asmar, M. E., and Kirar, C. (2012). "Energy Consumption Evaluation of United States Navy Leed Certified Buildings." *J. Perform. Constr. Facil.* 1(1), 144.
56. Menches, C. L., and Hanna, A. S. (2006). "Quantitative Measurement of Successful Performance from the Project Manager's Perspective." *J. Constr. Eng. Mgmt.*, 132, 1284.
57. Molenaar, K.R. (1995). "Appropriate Project Characteristics for Public Sector Design-Build Projects." M.S. thesis, University of Colorado.
58. Molenaar, K. R., Songer, A. D., and Barash, M. (1999). "Public-Sector Design/Build Evolution and Performance." *J. Constr. Eng. Mgmt.*, 15(2). 54-62.
59. Molenaar, K. R., and Navarro, D. (2011). "Toward Use of Key Performance Indicators in Highway Design and Construction." Transportation Research Board 90th Annual Meeting.
60. Mossman, A., Ballard, G., and Pasquire (2010) "Integrated Project Delivery – Innovation in Integrated Design and Delivery." Draft for the Architectural Engineering and Design Management.
61. NASFA, COAA, APPA, AGC, AIA. (2010). Kenig, M., Allison, M., Black, B., Burdi, L., Colella, C., Davis, H. "Integrated Project Delivery for Public and Private Owners."
62. National Cooperative Highway Research Program (NCHRP). (2007). "Comparing state DOT's construction project cost and schedule performance: 28 best practices from nine states." Project 20-24, Task 37A. Joe Crossett & Lauren Hines, TransTech Management, Inc. Washington, DC.
63. National Football League. (2012). NFL Quarterback Rating Formula. <http://www.nfl.com/help/quarterbackratingformula>.

64. Nelson, J. S., Hanna, A. S., Russell, J. S., and Hendrickson, M. L. (2008). “Leadership in the global construction industry: the challenge of the brutal facts.” Proceedings of Leadership and Management in Construction LEAD October 16-19, 2008, South Lake Tahoe, CA.
65. Oberlander, G.D. and Zeitoun, A.A. (1993). "Early Warning Signs of Project Changes, the Construction Industry Institute, The University of Texas at Austin, Austin, T.X, Source Document 91.
66. Office of Government Commerce (2007), “Achieving Excellence in Construction Procurement Guide”, vol. 5, 6, <www.ogc.gov.uk>. (April 27, 2011).
67. Oyetunji, A. & Anderson, S. (2006) "Relative effectiveness of project delivery and contract strategies." *Journal of Construction Engineering and Management*, 132 (1).
68. Pocock, J. B. (1996). The Relationship between Alternative Project Approaches, Integration, and Performance. PhD dissertation, University of Illinois at Urbana-Champaign.
69. Pocock, James B., Chang T. Hyun, Liang Y. Liu, and Michael K. Kim. (1996). “Relationship between Project Interaction and Performance Indicators”. *Journal of Construction Engineering and Management*, 122(2), 165-176.
70. Rankin, J., Fayek, A. R., Meade, G., Haas, C., and Manseau, A. (2008). "Initial metrics and pilot program results for measuring the performance of the Canadian construction industry." *Canadian Journal of Civil Engineering*, 35(9), 894-907.
71. Riley, D., Sanvido, V., Horman, M., McLaughlin, M., and Kerr, D. (2005). "Lean and Green: The Role of Design-Build Mechanical Competencies in the Design and Construction of Green Buildings." Construction Research Congress 2005: Broadening Perspectives - Proceedings of the Congress, April 5, 2005 - April 7, 2005.
72. Rojas, E. M., and Kell, I. (2008). "Comparative analysis of project delivery systems cost performance in Pacific Northwest public schools." *J. Constr. Eng. Mgmt.*, 134, 387.
73. RS Means (2009*). Building Construction Cost Data*, 67th Annual Edition. Kingston, MA.

74. Sands, M. S. (2010). "Standards and Measures—Whole-building Metrics Driving Innovation and High Performance." *Lean Construction Journal*, 1-16.
75. Sanvido, V., and Konchar, M. (1998) Project Delivery Systems: CM at Risk, DB, DBB. CII RT133.
76. Smith, R., Mossman, A., Emmitt, S. (2011). “Lean and integrated project delivery.” *Lean Construction Journal* 2011, 1-16.
77. Songer, A. D., and Molenaar, K. R. (1996). "Selecting design-build: public and private sector owner attitudes." *J. Constr. Eng. Mgmt*., 12(6), 47-53.
78. Teicholz, P., Goodrum, P. M., and Haas, C. T. (2001). "US Construction Labor Productivity Trends, 1970–1998." *J. Constr. Eng. Mgmt*., 127 (5).
79. Teicholz, P. (2004). “Labor Productivity Declines in the Construction Industry: Causes and Remedies.” *AECbytes Viewpoint, #4.*
80. Thomas, S. R., Macken, C. L., Chung, T. H., and Kim, I. (2002). "Measuring the Impacts of the Delivery System on Project Performance – Design-Build and Design-Bid-Build." NIST GCR 02, 840.
81. Thomsen, C. (2010). “Managing Integrated Project Delivery”. Construction Management Association of America (CMAA).
82. Tradeline (2007). Allen, J. “One big idea for construction delivery: Risk realignment.” Published September 19, 2007.
83. Tradeline (2009). Allen, J. “*Lean management is a game-changer for facility design and construction*.” Published July 14, 2009.
84. UK Office of Government Commerce (2007). “The integrated project team. Team working and partnering.”
85. U.S. Census Bureau (2012) “*Construction Spending Total construction*.” Accessed at <http://www.census.gov/construction/c30/c30index.html>.
86. U.S. Department of Commerce (2012). Bureau of Economic Analysis. Accessed at <http://www.bea.gov/iTable/iTable.cfm?ReqID=9&step=1>.
87. Womack, J. P., and Jones, D. T. (2003). *Lean Thinking*, Simon & Shuster, New York.

Appendix B – DEFINITIONS AND GLOSSARY OF TERMS

This appendix first provides definitions of terms organized in alphabetical order, and then offers a list of abbreviations commonly used in the text. The body of this document includes several additional definitions, most notably in *Section 4.2* where each performance metric is defined individually before it is used to compare IPD and non-IPD projects.

Building Information Modeling (BIM): a model-based digital design process of generating and managing project and building data so that the knowledge created throughout the model development is more usable, accessible and transparent.

Cluster Teams: interdisciplinary teams of subject matter experts, comprised of owners, designers and trade specialists, assembled to handle specific design and production areas.

Co-location: a collaborative method where project participants conduct their day-to-day work in the same physical space; otherwise known as the “Big Room”.

Core Team Leadership: a group responsible for major project decision-making, comprised of representatives from a minimum of the owner, contractor and designer teams.

Integrated Form of Agreement (IFOA): the original standard form for an IPA.

Integrated Project Agreement (IPA): a single written multiparty contract linking all key IPD project participants (owner, designer, constructor, and often key consultants and subcontractors), that specifies their respective roles, rights, obligations, and liabilities; examples include IFOA and ConsensusDOCS 300.

Integrated Project Delivery (IPD): a project delivery system distinguished by a multiparty agreement and the very early involvement of the key participants; ideally at 0% design but definitely before 10% design complete.

Last Planner System (LPS): a production control tool that smoothes construction project task workflow.

Lean: lean manufacturing, or simply lean, is a business philosophy and production practice of maximizing customer value and minimizing waste.

Lean Construction: a production management approach to construction projects, aimed at maximizing value and minimizing waste.

Lean Project Delivery System (LPDS): A project delivery system that incorporates lean thinking for managing project. LPDS has five, interconnected stages: Project Definition, Lean Design, Lean Supply, Lean Assembly and Use.

Multiparty Agreement (MPA): See “Integrated Project Agreement” above.

Percent Plan Complete (PPC): a measure of work flow reliability, calculated by dividing the number of actual task completions by the number of planned task completions.

Project Delivery System (PDS): defines the relationships, roles and responsibilities of parties involved in a project. It also establishes an execution framework sequencing design, procurement, and construction activities required to provide a facility. Practices and techniques of management that are used can also be described by the project delivery system.

Pull Planning: a scheduling tool that works backward from the desired condition to the current condition, pulling production items into the system only as needed.

Reliable Promising: a commitment management system that measures the Percent Plan Complete (PPC). A promise should not be made if there is any doubt that it can be fulfilled.

Set Based Design (SBD): an approach that defers design decisions to the “last responsible moment” to allow for the evaluation of several alternatives against all design criteria.

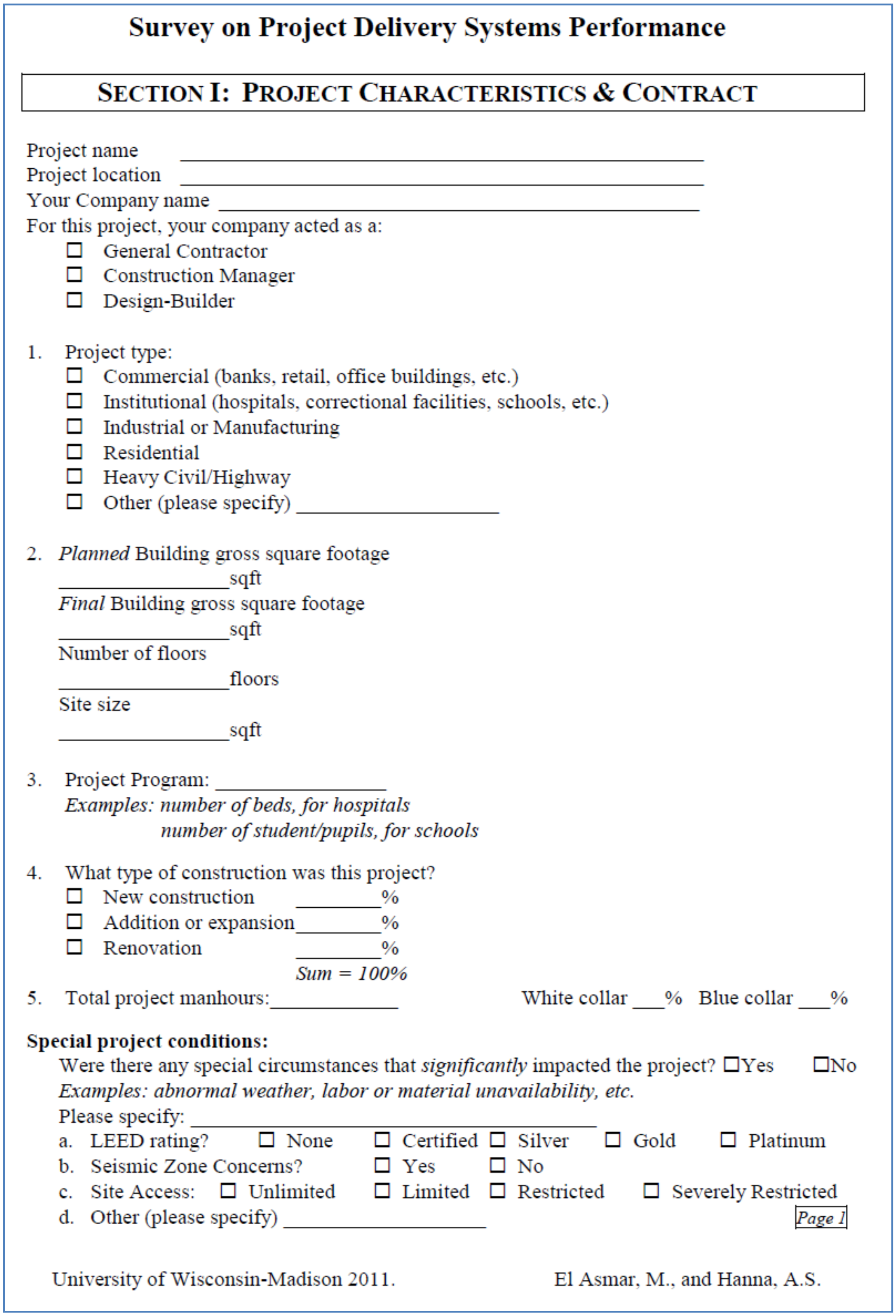
Target Value Design (TVD) or Target Costing: a collaborative process that is iterative in nature, and consists of establishing early financial targets for the project and then designing to a detailed estimate rather than estimating based on a detailed design, in order to drive innovation.

Value Stream Mapping (VSM): a tool used to visualize material and information flows of project activities in order to minimize waste. VSM helps understand work processes and identify sources of waste by distinguishing between value-added and non-value-added activities.

List of Abbreviations

2S	Second Shift Work
AEC	Architecture / Engineering / Construction
A/E	Architect / Engineer
AIA	American Institute of Architects
CM/GC	Construction Manager or General Contractor
CMR	Construction Management at Risk
CII	Construction Industry Institute
CO	Change Orders
DB	Design-Build
DBB	Design-Bid-Build
GC	General Contractor
IPD	Integrated Project Delivery
IPD-ish	IPD project with no multiparty contract
LCI	Lean Construction Institute
LCJ	Lean Construction Journal
MWW	Mann-Whitney-Wilcoxon
OM	Over-manning
OSHA	Occupational Safety and Health Administration
OT	Overtime
PCA	Principal Component Analysis
Q-Q Plots	Quantile-to-Quantile Plots
RFI	Request for Information

Appendix C – DATA COLLECTION QUESTIONNAIRE



Survey on Project Delivery Systems Performance

SECTION I: PROJECT CHARACTERISTICS & CONTRACT

Project name ____
Project location ____
Your Company name ____
For this project, your company acted as a:

- ☐ General Contractor
- ☐ Construction Manager
- ☐ Design-Builder

1. Project type:
   - ☐ Commercial (banks, retail, office buildings, etc.)
   - ☐ Institutional (hospitals, correctional facilities, schools, etc.)
   - ☐ Industrial or Manufacturing
   - ☐ Residential
   - ☐ Heavy Civil/Highway
   - ☐ Other (please specify) ____________________

2. *Planned* Building gross square footage
   ________________sqft
   Final Building gross square footage
   ________________sqft
   Number of floors
   ________________floors
   Site size
   ________________sqft

3. Project Program: ________________
   Examples: number of beds, for hospitals
   number of student/pupils, for schools

4. What type of construction was this project?
   - ☐ New construction ________%
   - ☐ Addition or expansion________%
   - ☐ Renovation ________%

   Sum = 100%

5. Total project manhours:____________ White collar ___% Blue collar ___%

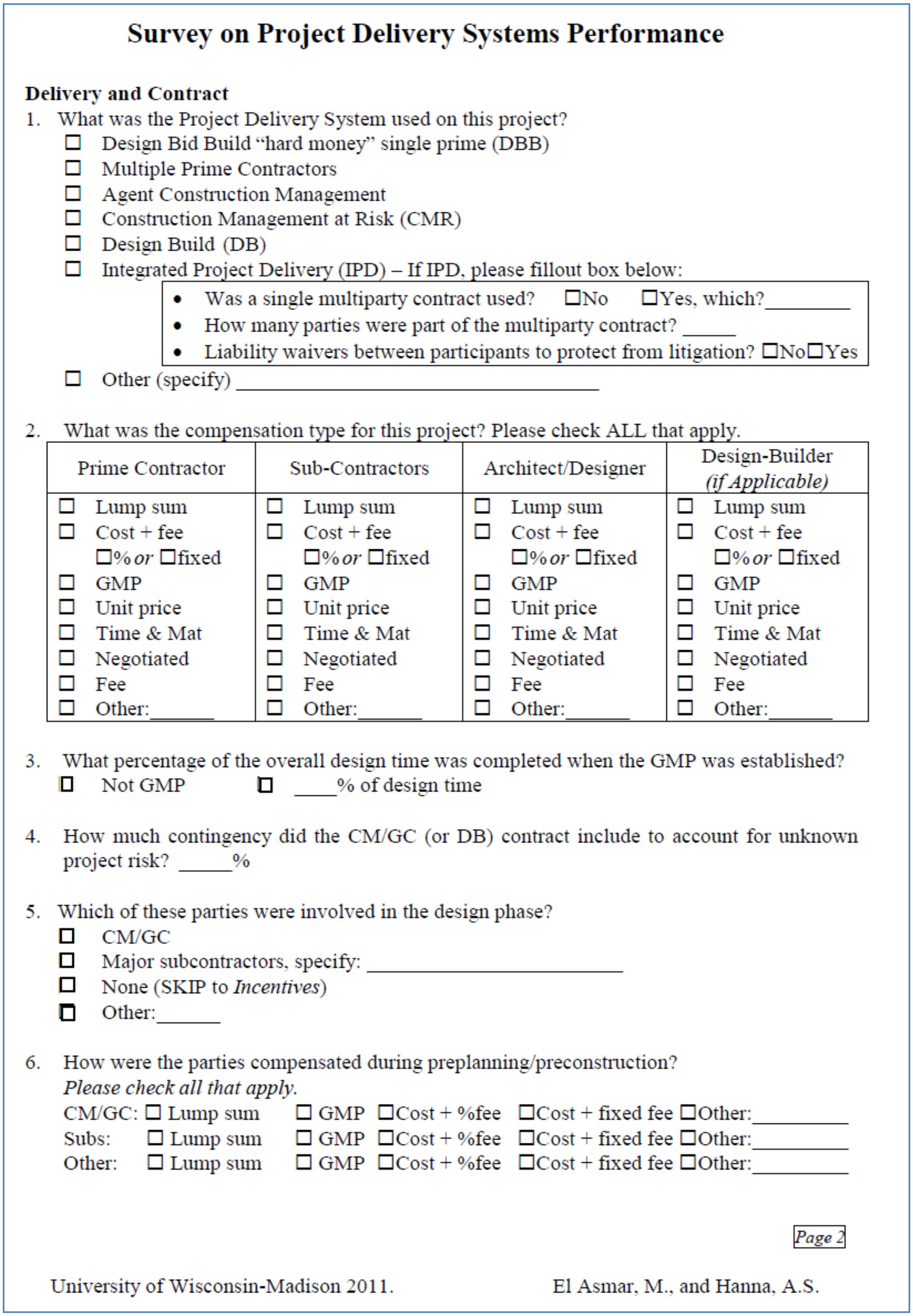
Special project conditions:

Were there any special circumstances that *significantly* impacted the project? ☐Yes ☐No
Examples: abnormal weather, labor or material unavailability, etc.
Please specify: ______________________________________

a. LEED rating? ☐ None ☐ Certified ☐ Silver ☐ Gold ☐ Platinum
b. Seismic Zone Concerns? ☐ Yes ☐ No
c. Site Access: ☐ Unlimited ☐ Limited ☐ Restricted ☐ Severely Restricted
d. Other (please specify) ____________________

Page 1

University of Wisconsin-Madison 2011. El Asmar, M., and Hanna, A.S.



Survey on Project Delivery Systems Performance

Delivery and Contract

1. What was the Project Delivery System used on this project?
   - ☐ Design Bid Build "hard money" single prime (DBB)
   - ☐ Multiple Prime Contractors
   - ☐ Agent Construction Management
   - ☐ Construction Management at Risk (CMR)
   - ☐ Design Build (DB)
   - ☐ Integrated Project Delivery (IPD) – If IPD, please fillout box below:
     - Was a single multiparty contract used? ☐No ☐Yes, which?_________
     - How many parties were part of the multiparty contract? _____
     - Liability waivers between participants to protect from litigation? ☐No☐Yes
   - ☐ Other (specify) ________________________________

2. What was the compensation type for this project? Please check ALL that apply.

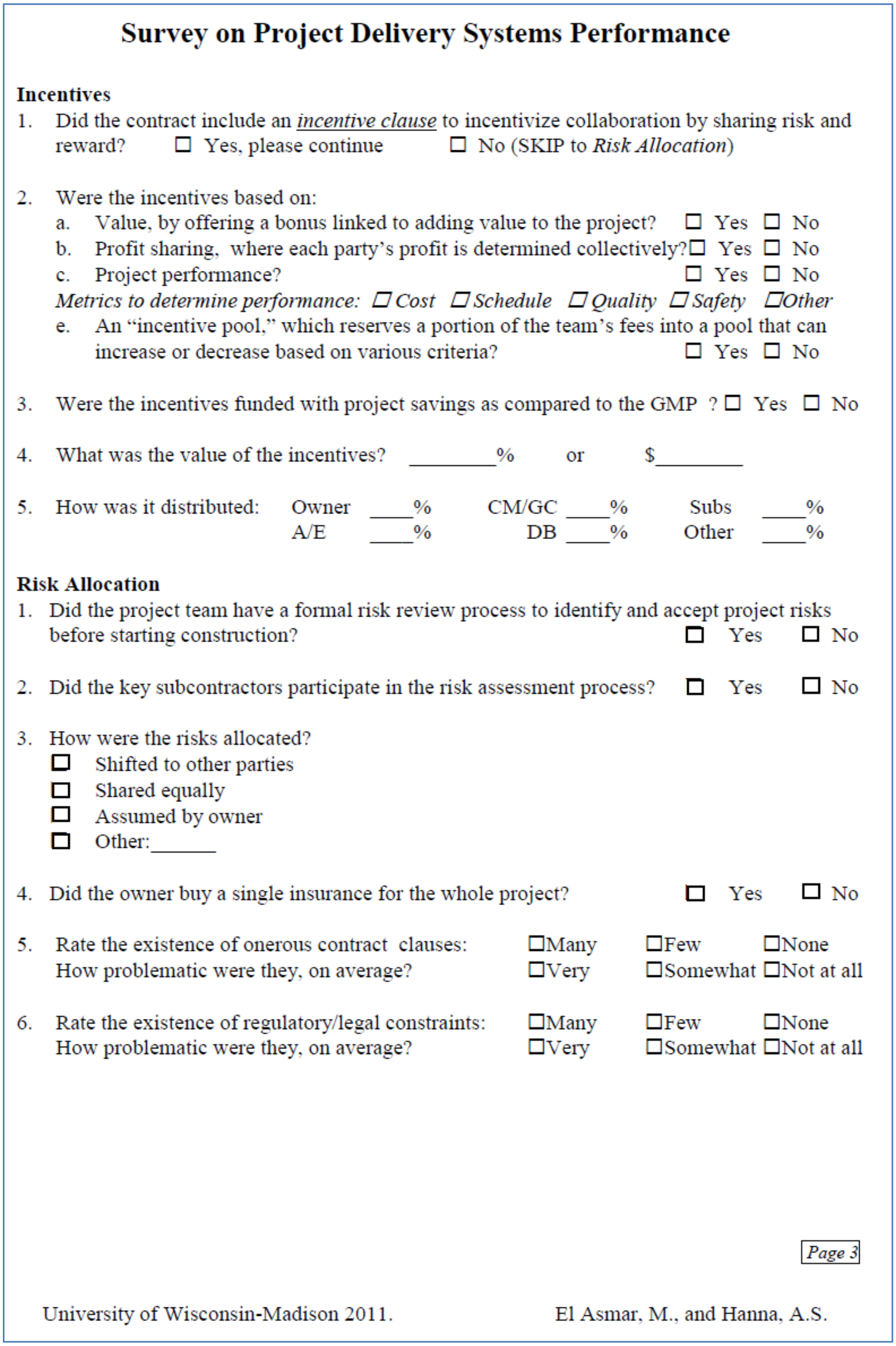
Prime Contractor	Sub-Contractors	Architect/Designer	Design-Builder *(if Applicable)*
☐ Lump sum	☐ Lump sum	☐ Lump sum	☐ Lump sum
☐ Cost + fee	☐ Cost + fee	☐ Cost + fee	☐ Cost + fee
☐% *or* ☐fixed	☐% *or* ☐fixed	☐% *or* ☐fixed	☐% *or* ☐fixed
☐ GMP	☐ GMP	☐ GMP	☐ GMP
☐ Unit price	☐ Unit price	☐ Unit price	☐ Unit price
☐ Time & Mat	☐ Time & Mat	☐ Time & Mat	☐ Time & Mat
☐ Negotiated	☐ Negotiated	☐ Negotiated	☐ Negotiated
☐ Fee	☐ Fee	☐ Fee	☐ Fee
☐ Other:______	☐ Other:______	☐ Other:______	☐ Other:______

3. What percentage of the overall design time was completed when the GMP was established?
   ☐ Not GMP ☐ ____% of design time

4. How much contingency did the CM/GC (or DB) contract include to account for unknown project risk? _____%

5. Which of these parties were involved in the design phase?
   - ☐ CM/GC
   - ☐ Major subcontractors, specify: ________________________
   - ☐ None (SKIP to *Incentives*)
   - ☐ Other:______

6. How were the parties compensated during preplanning/preconstruction?
   Please check all that apply.
   CM/GC: ☐ Lump sum ☐ GMP ☐Cost + %fee ☐Cost + fixed fee ☐Other:_________
   Subs: ☐ Lump sum ☐ GMP ☐Cost + %fee ☐Cost + fixed fee ☐Other:_________
   Other: ☐ Lump sum ☐ GMP ☐Cost + %fee ☐Cost + fixed fee ☐Other:_________



Survey on Project Delivery Systems Performance

Incentives

1. Did the contract include an *incentive clause* to incentivize collaboration by sharing risk and reward? ☐ Yes, please continue ☐ No (SKIP to *Risk Allocation*)

2. Were the incentives based on:
   a. Value, by offering a bonus linked to adding value to the project? ☐ Yes ☐ No
   b. Profit sharing, where each party's profit is determined collectively? ☐ Yes ☐ No
   c. Project performance? ☐ Yes ☐ No
   Metrics to determine performance: ☐ Cost ☐ Schedule ☐ Quality ☐ Safety ☐ Other
   e. An "incentive pool," which reserves a portion of the team's fees into a pool that can increase or decrease based on various criteria? ☐ Yes ☐ No

3. Were the incentives funded with project savings as compared to the GMP ? ☐ Yes ☐ No

4. What was the value of the incentives? ________% or $________

5. How was it distributed: Owner ____% CM/GC ____% Subs ____%
   A/E ____% DB ____% Other ____%

Risk Allocation

1. Did the project team have a formal risk review process to identify and accept project risks before starting construction? ☐ Yes ☐ No

2. Did the key subcontractors participate in the risk assessment process? ☐ Yes ☐ No

3. How were the risks allocated?
   ☐ Shifted to other parties
   ☐ Shared equally
   ☐ Assumed by owner
   ☐ Other:______

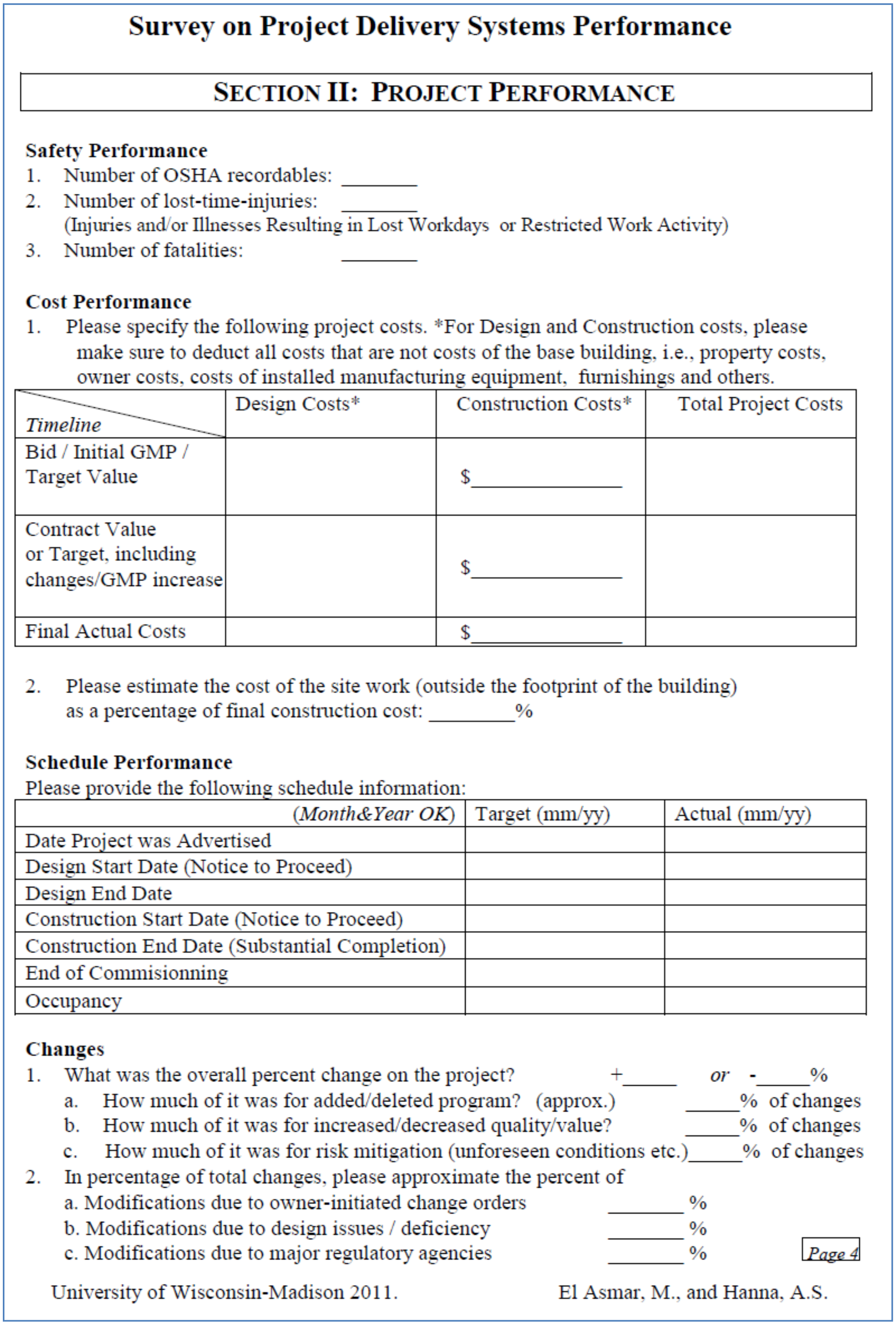
4. Did the owner buy a single insurance for the whole project? ☐ Yes ☐ No

5. Rate the existence of onerous contract clauses: ☐Many ☐Few ☐None
   How problematic were they, on average? ☐Very ☐Somewhat ☐Not at all

6. Rate the existence of regulatory/legal constraints: ☐Many ☐Few ☐None
   How problematic were they, on average? ☐Very ☐Somewhat ☐Not at all

Page 3

University of Wisconsin-Madison 2011. El Asmar, M., and Hanna, A.S.



Survey on Project Delivery Systems Performance

SECTION II: PROJECT PERFORMANCE

Safety Performance

1. Number of OSHA recordables: ________
2. Number of lost-time-injuries: ________
   (Injuries and/or Illnesses Resulting in Lost Workdays or Restricted Work Activity)
3. Number of fatalities: ________

Cost Performance

1. Please specify the following project costs. *For Design and Construction costs, please make sure to deduct all costs that are not costs of the base building, i.e., property costs, owner costs, costs of installed manufacturing equipment, furnishings and others.

Timeline	Design Costs*	Construction Costs*	Total Project Costs
Bid / Initial GMP / Target Value		$______________	
Contract Value or Target, including changes/GMP increase		$______________	
Final Actual Costs		$______________	

2. Please estimate the cost of the site work (outside the footprint of the building) as a percentage of final construction cost: ________%

Schedule Performance

Please provide the following schedule information:

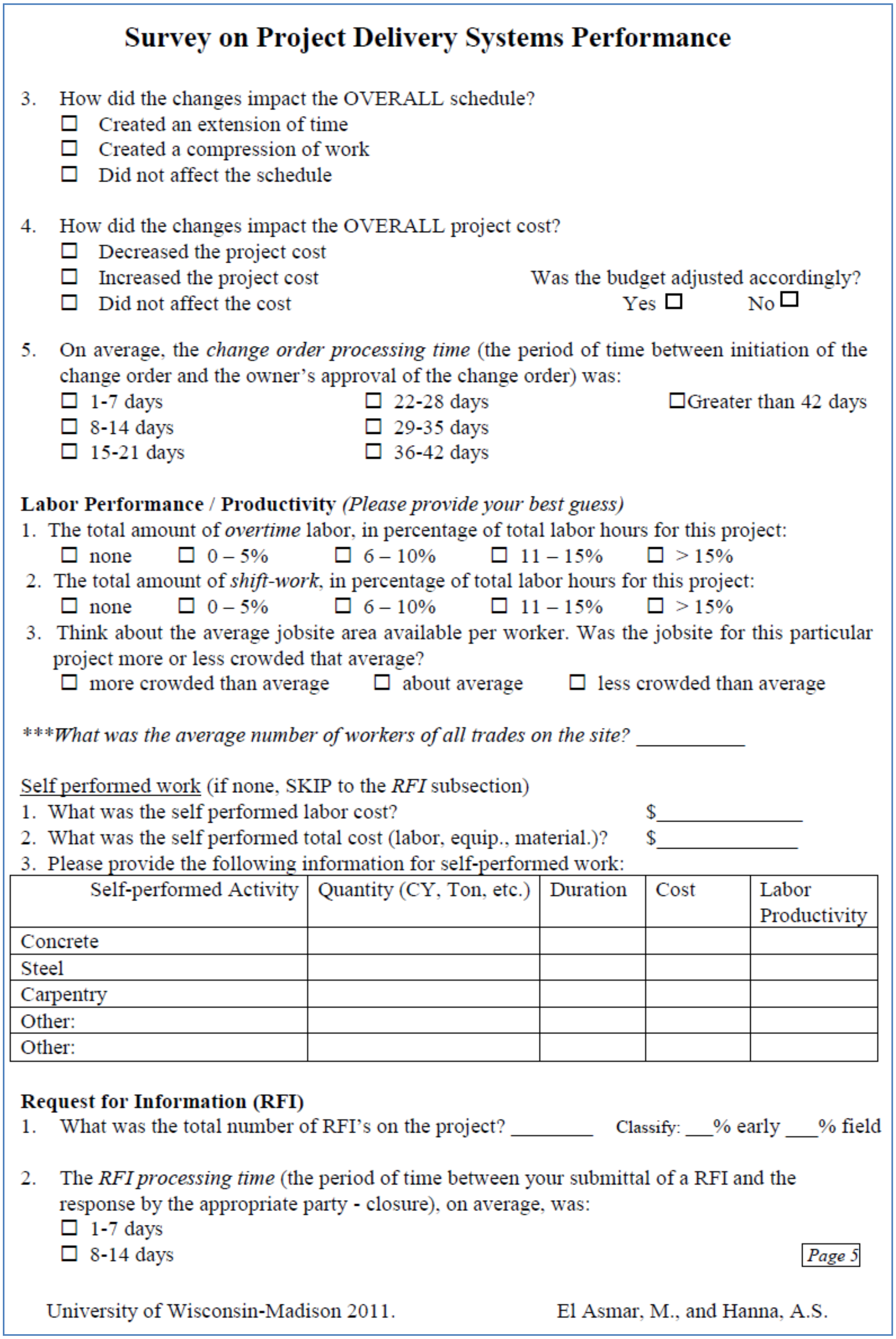
(*Month&Year OK*)	Target (mm/yy)	Actual (mm/yy)
Date Project was Advertised		
Design Start Date (Notice to Proceed)		
Design End Date		
Construction Start Date (Notice to Proceed)		
Construction End Date (Substantial Completion)		
End of Commisionning		
Occupancy		

Changes

1. What was the overall percent change on the project? +_____ *or* -_____%
   a. How much of it was for added/deleted program? (approx.) _____% of changes
   b. How much of it was for increased/decreased quality/value? _____% of changes
   c. How much of it was for risk mitigation (unforeseen conditions etc.)_____% of changes
2. In percentage of total changes, please approximate the percent of
   a. Modifications due to owner-initiated change orders ________ %
   b. Modifications due to design issues / deficiency ________ %
   c. Modifications due to major regulatory agencies ________ %

Page 4

University of Wisconsin-Madison 2011. El Asmar, M., and Hanna, A.S.



Survey on Project Delivery Systems Performance

3. How did the changes impact the OVERALL schedule?
   - ☐ Created an extension of time
   - ☐ Created a compression of work
   - ☐ Did not affect the schedule

4. How did the changes impact the OVERALL project cost?
   - ☐ Decreased the project cost
   - ☐ Increased the project cost
   - ☐ Did not affect the cost

   Was the budget adjusted accordingly? Yes ☐ No ☐

5. On average, the *change order processing time* (the period of time between initiation of the change order and the owner's approval of the change order) was:
   - ☐ 1-7 days
   - ☐ 8-14 days
   - ☐ 15-21 days
   - ☐ 22-28 days
   - ☐ 29-35 days
   - ☐ 36-42 days
   - ☐Greater than 42 days

Labor Performance / Productivity *(Please provide your best guess)*

1. The total amount of *overtime* labor, in percentage of total labor hours for this project:
   ☐ none ☐ 0 – 5% ☐ 6 – 10% ☐ 11 – 15% ☐ > 15%
2. The total amount of *shift-work*, in percentage of total labor hours for this project:
   ☐ none ☐ 0 – 5% ☐ 6 – 10% ☐ 11 – 15% ☐ > 15%
3. Think about the average jobsite area available per worker. Was the jobsite for this particular project more or less crowded that average?
   ☐ more crowded than average ☐ about average ☐ less crowded than average

****What was the average number of workers of all trades on the site?* ___________

Self performed work (if none, SKIP to the *RFI* subsection)

1. What was the self performed labor cost? $______________
2. What was the self performed total cost (labor, equip., material.)? $______________
3. Please provide the following information for self-performed work:

Self-performed Activity	Quantity (CY, Ton, etc.)	Duration	Cost	Labor Productivity
Concrete				
Steel				
Carpentry				
Other:				
Other:				

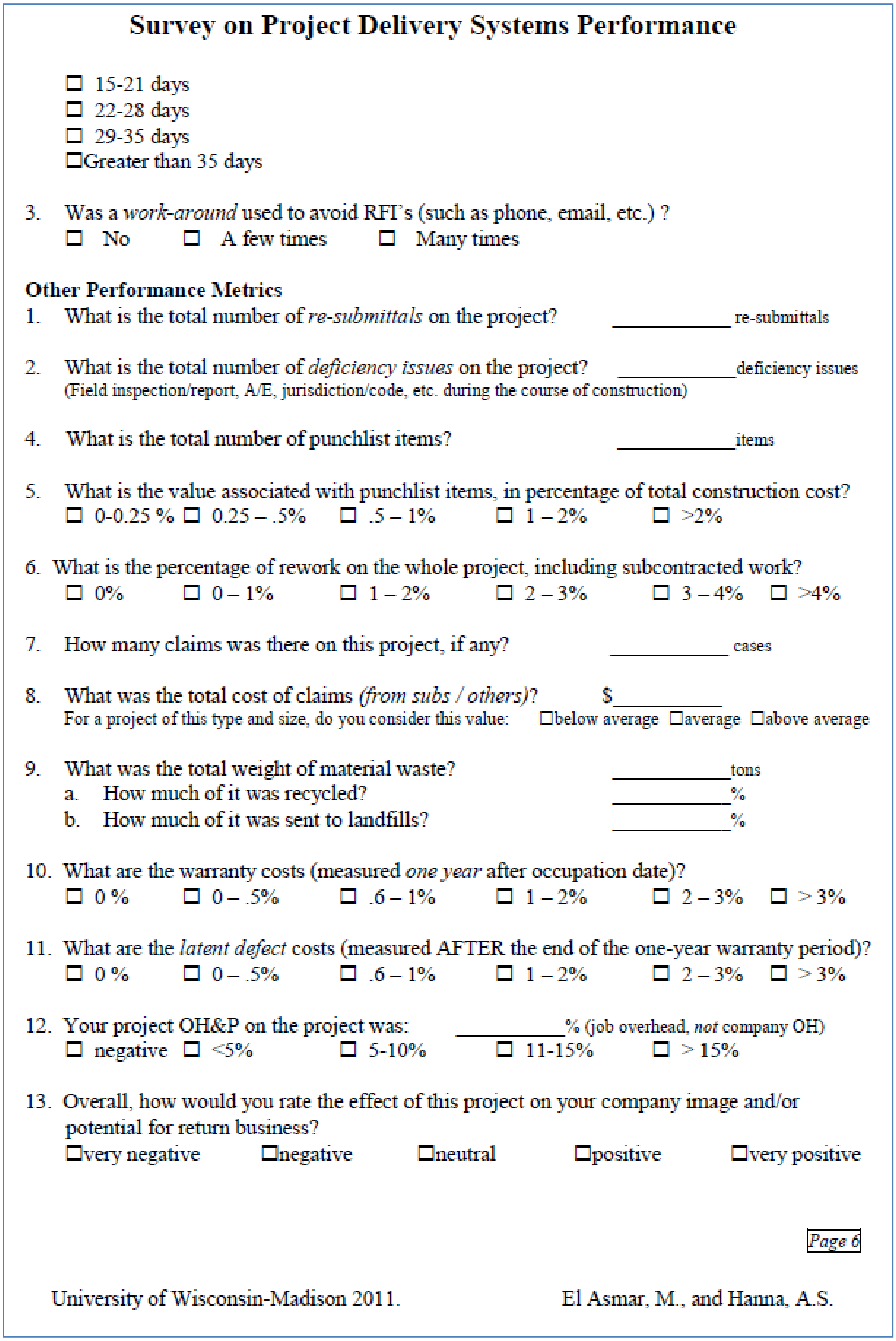
Request for Information (RFI)

1. What was the total number of RFI's on the project? ________ Classify: ___% early ___% field

2. The *RFI processing time* (the period of time between your submittal of a RFI and the response by the appropriate party - closure), on average, was:
   - ☐ 1-7 days
   - ☐ 8-14 days

Page 5

University of Wisconsin-Madison 2011. El Asmar, M., and Hanna, A.S.



Survey on Project Delivery Systems Performance

☐ 15-21 days
☐ 22-28 days
☐ 29-35 days
☐Greater than 35 days

3. Was a *work-around* used to avoid RFI's (such as phone, email, etc.) ?
☐ No ☐ A few times ☐ Many times

Other Performance Metrics

1. What is the total number of *re-submittals* on the project? ____________ re-submittals

2. What is the total number of *deficiency issues* on the project? ____________deficiency issues
(Field inspection/report, A/E, jurisdiction/code, etc. during the course of construction)

4. What is the total number of punchlist items? ____________items

5. What is the value associated with punchlist items, in percentage of total construction cost?
☐ 0-0.25 % ☐ 0.25 – .5% ☐ .5 – 1% ☐ 1 – 2% ☐ >2%

6. What is the percentage of rework on the whole project, including subcontracted work?
☐ 0% ☐ 0 – 1% ☐ 1 – 2% ☐ 2 – 3% ☐ 3 – 4% ☐ >4%

7. How many claims was there on this project, if any? ____________ cases

8. What was the total cost of claims *(from subs / others)*? $____________
For a project of this type and size, do you consider this value: ☐below average ☐average ☐above average

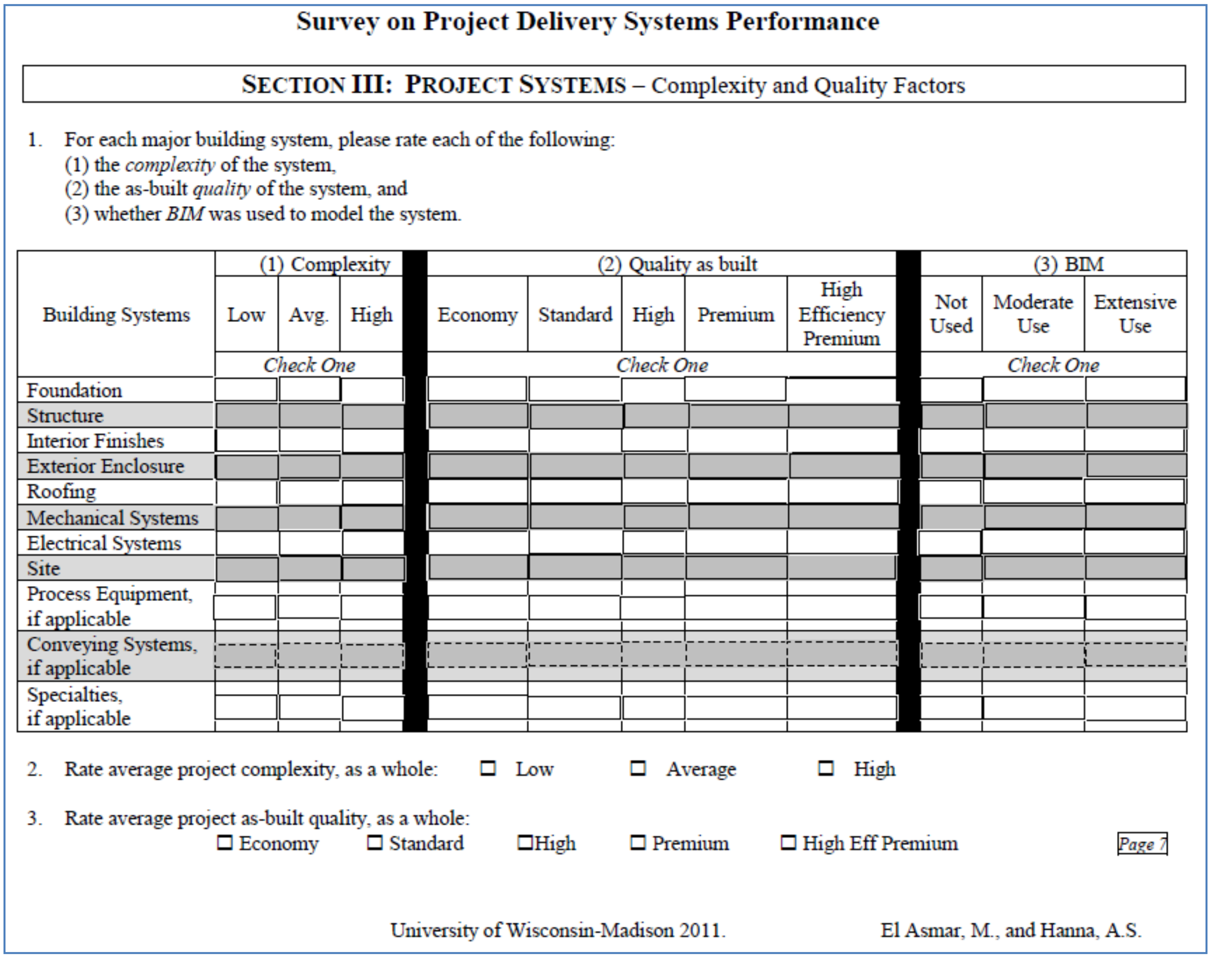
9. What was the total weight of material waste? ____________tons
a. How much of it was recycled? ____________%
b. How much of it was sent to landfills? ____________%

10. What are the warranty costs (measured *one year* after occupation date)?
☐ 0 % ☐ 0 – .5% ☐ .6 – 1% ☐ 1 – 2% ☐ 2 – 3% ☐ > 3%

11. What are the *latent defect* costs (measured AFTER the end of the one-year warranty period)?
☐ 0 % ☐ 0 – .5% ☐ .6 – 1% ☐ 1 – 2% ☐ 2 – 3% ☐ > 3%

12. Your project OH&P on the project was: ____________% (job overhead, *not* company OH)
☐ negative ☐ <5% ☐ 5-10% ☐ 11-15% ☐ > 15%

13. Overall, how would you rate the effect of this project on your company image and/or potential for return business?
☐very negative ☐negative ☐neutral ☐positive ☐very positive



Survey on Project Delivery Systems Performance

SECTION III: PROJECT SYSTEMS – Complexity and Quality Factors

1. For each major building system, please rate each of the following:
   (1) the *complexity* of the system,
   (2) the as-built *quality* of the system, and
   (3) whether *BIM* was used to model the system.

Building Systems	(1) Complexity			(2) Quality as built					(3) BIM		
	Low	Avg.	High	Economy	Standard	High	Premium	High Efficiency Premium	Not Used	Moderate Use	Extensive Use
	Check One			*Check One*					*Check One*		
Foundation											
Structure											
Interior Finishes											
Exterior Enclosure											
Roofing											
Mechanical Systems											
Electrical Systems											
Site											
Process Equipment, if applicable											
Conveying Systems, if applicable											
Specialties, if applicable											

2. Rate average project complexity, as a whole: ☐ Low ☐ Average ☐ High

3. Rate average project as-built quality, as a whole:
   ☐ Economy ☐ Standard ☐High ☐ Premium ☐ High Eff Premium

Page 7

University of Wisconsin-Madison 2011. El Asmar, M., and Hanna, A.S.

Survey on Project Delivery Systems Performance

SECTION IV: PROJECT TEAM AND COLLABORATION

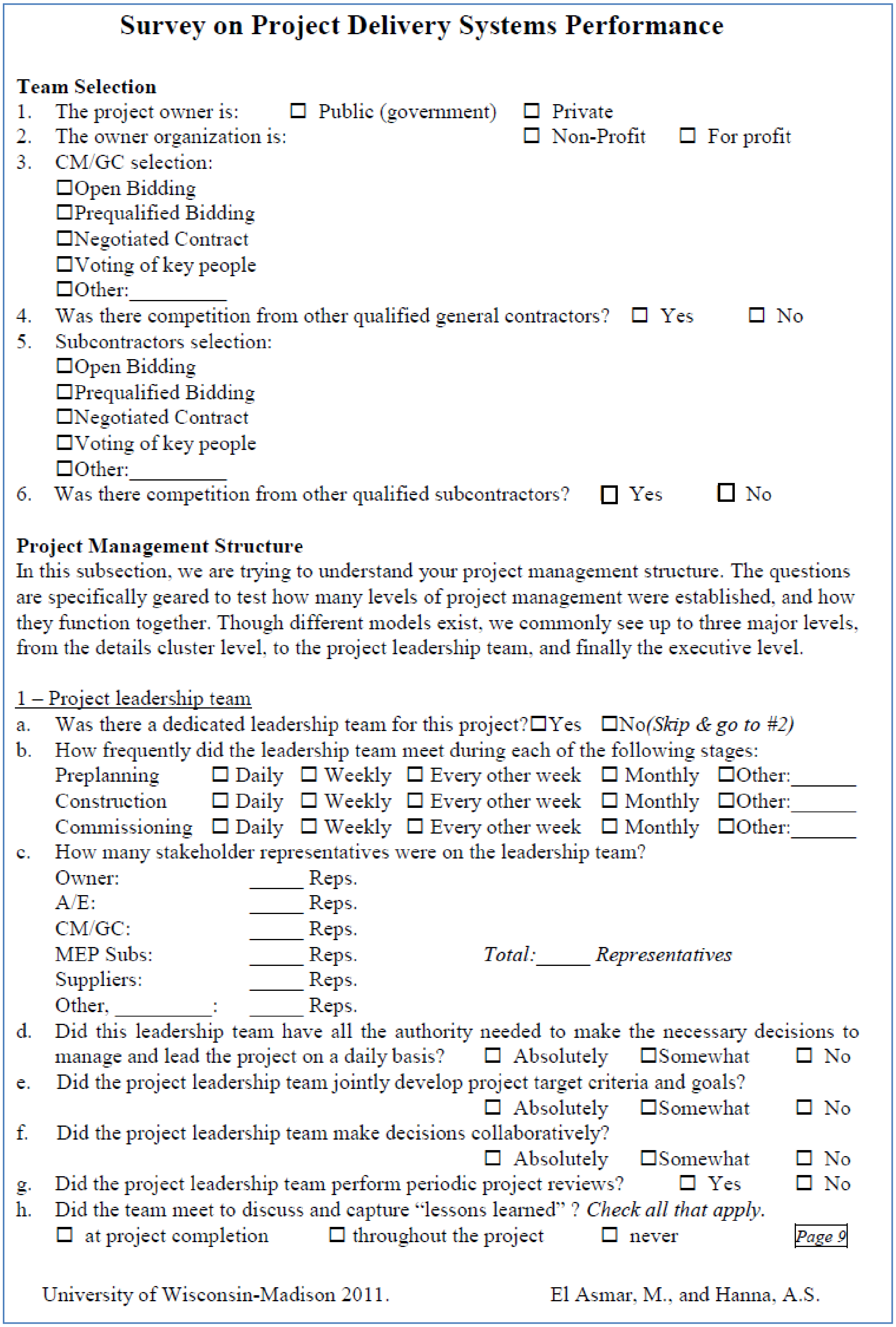
Experience

1. Please tell us about the prior experience of the stakeholders before the start of this project. Example: *is the CM/GC experienced with building hospitals?*

	A lot	Some	A little	None
Past experience with this *type* of construction:				
CM/GC				
Subcontractors				
Owner				
A/E				
DB firm (if applicable)				
Past experience with this *size* of construction:				
CM/GC				
Subcontractors				
Owner				
A/E				
DB firm (if applicable)				
Past experience with this ***project delivery system***:				
CM/GC				
Subcontractors				
Owner				
A/E				
DB firm (if applicable)				
Past experience with ***BIM***:				
CM/GC				
Subcontractors				
Owner				
A/E				
DB firm (if applicable)				
Your past experience with the ***other stakeholders***:				
Subcontractors				
Owner				
A/E				
Past experience of the ***team as a unit***:				

2. Now looking back, what is your overall satisfaction working with this project team? In other words, rate your *current* experience working as a team for this project:

Project team with Architect and Owner: ☐Excellent ☐Very Good ☐Good ☐Fair ☐Poor
Construction team with Subcontractors: ☐Excellent ☐Very Good ☐Good ☐Fair ☐Poor



Survey on Project Delivery Systems Performance

Team Selection

1. The project owner is: ☐ Public (government) ☐ Private
2. The owner organization is: ☐ Non-Profit ☐ For profit
3. CM/GC selection:
   ☐Open Bidding
   ☐Prequalified Bidding
   ☐Negotiated Contract
   ☐Voting of key people
   ☐Other:__________
4. Was there competition from other qualified general contractors? ☐ Yes ☐ No
5. Subcontractors selection:
   ☐Open Bidding
   ☐Prequalified Bidding
   ☐Negotiated Contract
   ☐Voting of key people
   ☐Other:__________
6. Was there competition from other qualified subcontractors? ☐ Yes ☐ No

Project Management Structure

In this subsection, we are trying to understand your project management structure. The questions are specifically geared to test how many levels of project management were established, and how they function together. Though different models exist, we commonly see up to three major levels, from the details cluster level, to the project leadership team, and finally the executive level.

<u>1 – Project leadership team</u>

a. Was there a dedicated leadership team for this project?☐Yes ☐No*(Skip & go to #2)*

b. How frequently did the leadership team meet during each of the following stages:

Preplanning	☐ Daily	☐ Weekly	☐ Every other week	☐ Monthly	☐Other:______
Construction	☐ Daily	☐ Weekly	☐ Every other week	☐ Monthly	☐Other:______
Commissioning	☐ Daily	☐ Weekly	☐ Every other week	☐ Monthly	☐Other:______

c. How many stakeholder representatives were on the leadership team?

Owner: _____ Reps.
A/E: _____ Reps.
CM/GC: _____ Reps.
MEP Subs: _____ Reps. *Total:_____ Representatives*
Suppliers: _____ Reps.
Other, __________: _____ Reps.

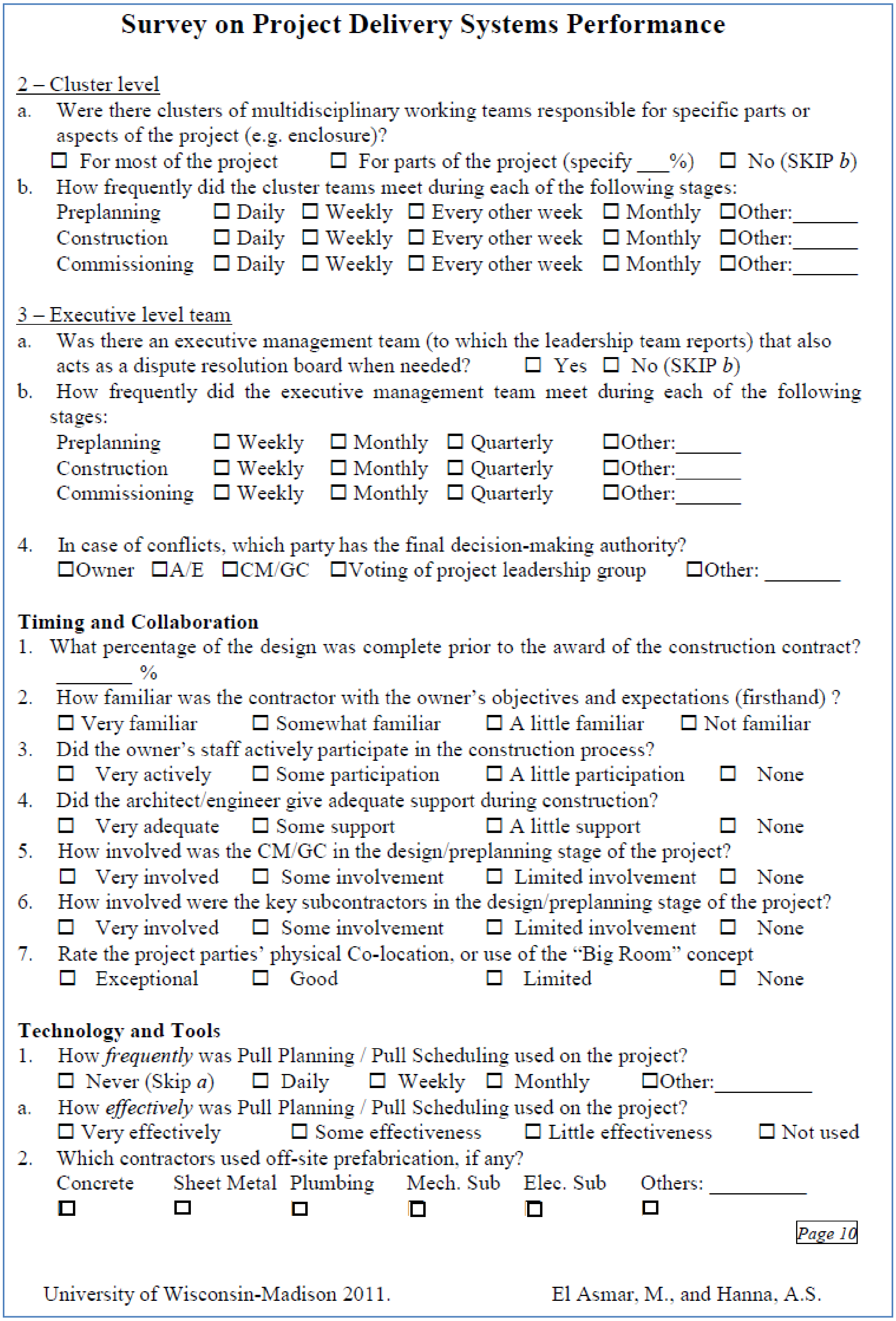
d. Did this leadership team have all the authority needed to make the necessary decisions to manage and lead the project on a daily basis? ☐ Absolutely ☐Somewhat ☐ No

e. Did the project leadership team jointly develop project target criteria and goals?
☐ Absolutely ☐Somewhat ☐ No

f. Did the project leadership team make decisions collaboratively?
☐ Absolutely ☐Somewhat ☐ No

g. Did the project leadership team perform periodic project reviews? ☐ Yes ☐ No

h. Did the team meet to discuss and capture "lessons learned" ? *Check all that apply.*
☐ at project completion ☐ throughout the project ☐ never



Survey on Project Delivery Systems Performance

2 – Cluster level

a. Were there clusters of multidisciplinary working teams responsible for specific parts or aspects of the project (e.g. enclosure)?

☐ For most of the project ☐ For parts of the project (specify ___%) ☐ No (SKIP *b*)

b. How frequently did the cluster teams meet during each of the following stages:

Preplanning	☐ Daily	☐ Weekly	☐ Every other week	☐ Monthly	☐Other:______
Construction	☐ Daily	☐ Weekly	☐ Every other week	☐ Monthly	☐Other:______
Commissioning	☐ Daily	☐ Weekly	☐ Every other week	☐ Monthly	☐Other:______

3 – Executive level team

a. Was there an executive management team (to which the leadership team reports) that also acts as a dispute resolution board when needed? ☐ Yes ☐ No (SKIP *b*)

b. How frequently did the executive management team meet during each of the following stages:

Preplanning	☐ Weekly	☐ Monthly	☐ Quarterly	☐Other:______
Construction	☐ Weekly	☐ Monthly	☐ Quarterly	☐Other:______
Commissioning	☐ Weekly	☐ Monthly	☐ Quarterly	☐Other:______

4. In case of conflicts, which party has the final decision-making authority?

☐Owner ☐A/E ☐CM/GC ☐Voting of project leadership group ☐Other: _______

Timing and Collaboration

1. What percentage of the design was complete prior to the award of the construction contract?

________ %

2. How familiar was the contractor with the owner's objectives and expectations (firsthand) ?

☐ Very familiar ☐ Somewhat familiar ☐ A little familiar ☐ Not familiar

3. Did the owner's staff actively participate in the construction process?

☐ Very actively ☐ Some participation ☐ A little participation ☐ None

4. Did the architect/engineer give adequate support during construction?

☐ Very adequate ☐ Some support ☐ A little support ☐ None

5. How involved was the CM/GC in the design/preplanning stage of the project?

☐ Very involved ☐ Some involvement ☐ Limited involvement ☐ None

6. How involved were the key subcontractors in the design/preplanning stage of the project?

☐ Very involved ☐ Some involvement ☐ Limited involvement ☐ None

7. Rate the project parties' physical Co-location, or use of the "Big Room" concept

☐ Exceptional ☐ Good ☐ Limited ☐ None

Technology and Tools

1. How *frequently* was Pull Planning / Pull Scheduling used on the project?

☐ Never (Skip *a*) ☐ Daily ☐ Weekly ☐ Monthly ☐Other:_________

a. How *effectively* was Pull Planning / Pull Scheduling used on the project?

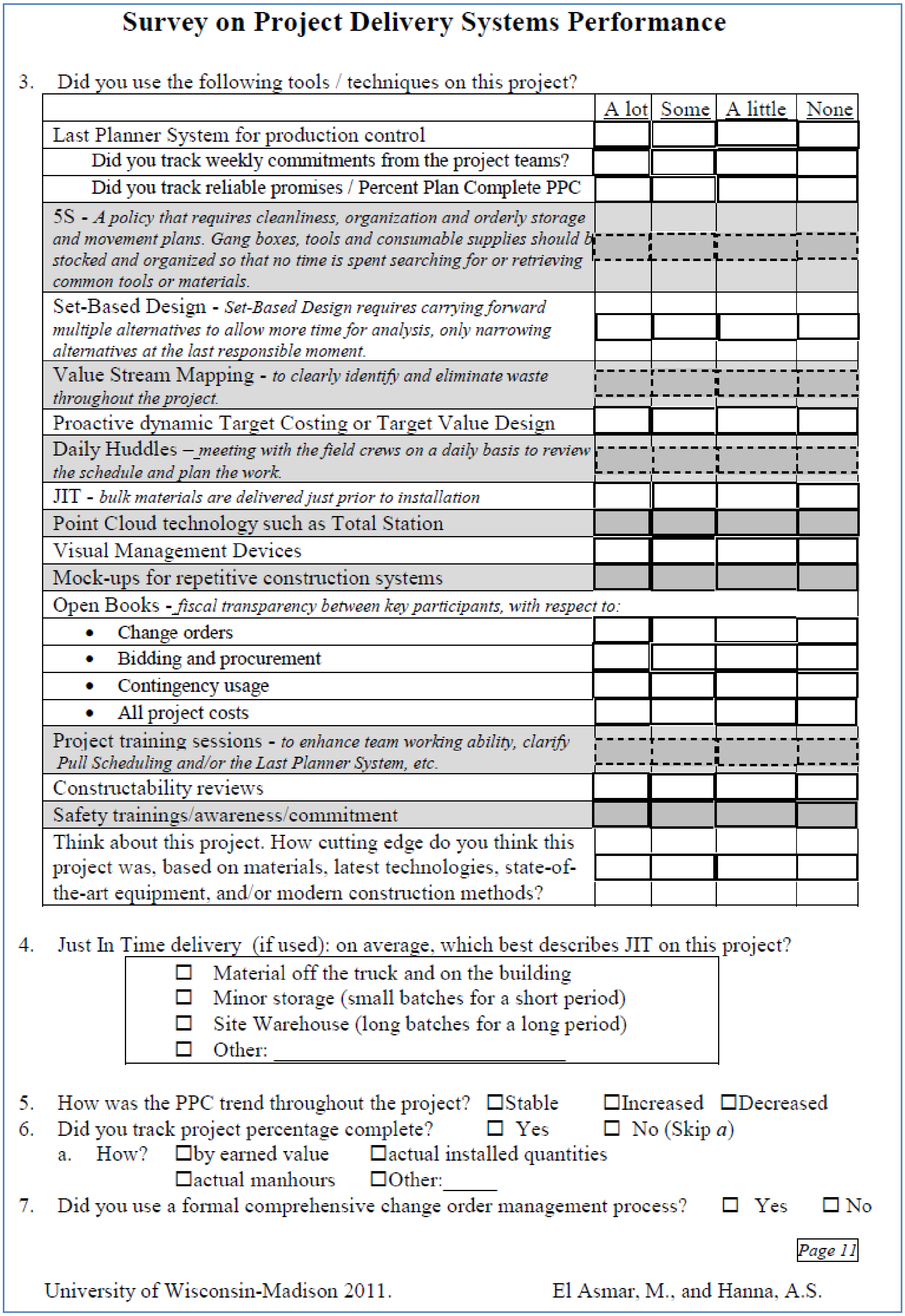
☐ Very effectively ☐ Some effectiveness ☐ Little effectiveness ☐ Not used

2. Which contractors used off-site prefabrication, if any?

Concrete	Sheet Metal	Plumbing	Mech. Sub	Elec. Sub	Others: _________
☐	☐	☐	☐	☐	☐

Page 10

University of Wisconsin-Madison 2011. El Asmar, M., and Hanna, A.S.



Survey on Project Delivery Systems Performance

3. Did you use the following tools / techniques on this project?

	A lot	Some	A little	None
Last Planner System for production control				
Did you track weekly commitments from the project teams?				
Did you track reliable promises / Percent Plan Complete PPC				
5S - *A policy that requires cleanliness, organization and orderly storage and movement plans. Gang boxes, tools and consumable supplies should be stocked and organized so that no time is spent searching for or retrieving common tools or materials.*				
Set-Based Design - *Set-Based Design requires carrying forward multiple alternatives to allow more time for analysis, only narrowing alternatives at the last responsible moment.*				
Value Stream Mapping - *to clearly identify and eliminate waste throughout the project.*				
Proactive dynamic Target Costing or Target Value Design				
Daily Huddles – *meeting with the field crews on a daily basis to review the schedule and plan the work.*				
JIT - *bulk materials are delivered just prior to installation*				
Point Cloud technology such as Total Station				
Visual Management Devices				
Mock-ups for repetitive construction systems				
Open Books - *fiscal transparency between key participants, with respect to:*				
• Change orders				
• Bidding and procurement				
• Contingency usage				
• All project costs				
Project training sessions - *to enhance team working ability, clarify Pull Scheduling and/or the Last Planner System, etc.*				
Constructability reviews				
Safety trainings/awareness/commitment				
Think about this project. How cutting edge do you think this project was, based on materials, latest technologies, state-of-the-art equipment, and/or modern construction methods?				

4. Just In Time delivery (if used): on average, which best describes JIT on this project?
   - ☐ Material off the truck and on the building
   - ☐ Minor storage (small batches for a short period)
   - ☐ Site Warehouse (long batches for a long period)
   - ☐ Other: ______________________________

5. How was the PPC trend throughout the project? ☐Stable ☐Increased ☐Decreased
6. Did you track project percentage complete? ☐ Yes ☐ No (Skip *a*)
   a. How? ☐by earned value ☐actual installed quantities
   ☐actual manhours ☐Other:_____
7. Did you use a formal comprehensive change order management process? ☐ Yes ☐ No

Page 11

University of Wisconsin-Madison 2011. El Asmar, M., and Hanna, A.S.

Survey on Project Delivery Systems Performance

Building Information Modeling (BIM)

1. Was BIM used on this project? ☐ Yes ☐ No (Skip and go to *Project Team Contacts*)
2. Was a BIM protocol manual used/developed? ☐ Yes ☐ No
3. Did the contract allow the right of reliance on the 3D models? ☐ Yes ☐ No
4. Were the project parties using joint servers for the building model? ☐ Yes ☐ No

Rate the use of the BIM model for the following tasks. *Please check the appropriate box to answer if BIM was used a lot, if there was some BIM use, only little BIM use, or not BIM use at all for the respective tasks.*

	A lot	Some	A little	None
Visualization				
Space Validation				
Site Logistics				
Environmental Analysis				
Early Design Coordination				
MEP Coordination				
Design Collaboration				
Clash Detection				
Submittals				
Estimating / Quantity Take-off				
4D Scheduling				
Digital Fabrication / Prefab				
Construction Simulation / Virtual Mock-ups				
Proj. Turnover and Closeout				
Facilities Management, O&M				
Rule / Code Checking				
Other:________________				
Other:________________				

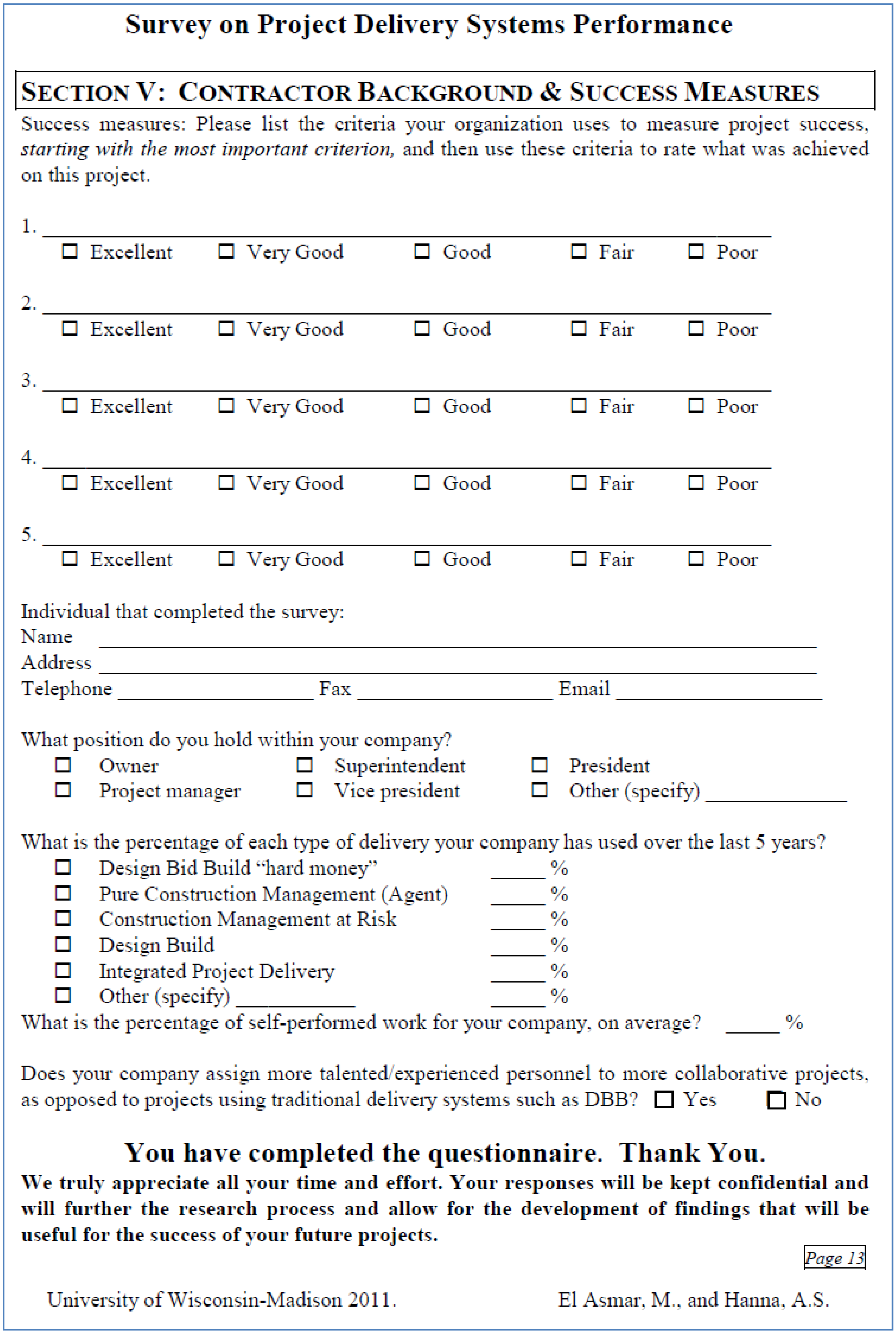
Project Team Contact Information

Project Owner ________________________ Contact Person ________________
E-mail ________________________ Phone ________________

Architect ________________________ Contact Person ________________
E-mail ________________________ Phone ________________

Mechanical Sub. ________________________ Contact Person ________________
E-mail ________________________ Phone ________________

Electrical Sub. ________________________ Contact Person ________________
E-mail ________________________ Phone ________________



Survey on Project Delivery Systems Performance

SECTION V: CONTRACTOR BACKGROUND & SUCCESS MEASURES

Success measures: Please list the criteria your organization uses to measure project success, *starting with the most important criterion,* and then use these criteria to rate what was achieved on this project.

1. ______________________________
☐ Excellent ☐ Very Good ☐ Good ☐ Fair ☐ Poor

2. ______________________________
☐ Excellent ☐ Very Good ☐ Good ☐ Fair ☐ Poor

3. ______________________________
☐ Excellent ☐ Very Good ☐ Good ☐ Fair ☐ Poor

4. ______________________________
☐ Excellent ☐ Very Good ☐ Good ☐ Fair ☐ Poor

5. ______________________________
☐ Excellent ☐ Very Good ☐ Good ☐ Fair ☐ Poor

Individual that completed the survey:
Name ______________________________
Address ______________________________
Telephone ________________ Fax ________________ Email ________________

What position do you hold within your company?

☐ Owner	☐ Superintendent	☐ President
☐ Project manager	☐ Vice president	☐ Other (specify) ____________

What is the percentage of each type of delivery your company has used over the last 5 years?

☐ Design Bid Build "hard money"	_____ %
☐ Pure Construction Management (Agent)	_____ %
☐ Construction Management at Risk	_____ %
☐ Design Build	_____ %
☐ Integrated Project Delivery	_____ %
☐ Other (specify) ___________	_____ %

What is the percentage of self-performed work for your company, on average? _____ %

Does your company assign more talented/experienced personnel to more collaborative projects, as opposed to projects using traditional delivery systems such as DBB? ☐ Yes ☐ No

You have completed the questionnaire. Thank You.

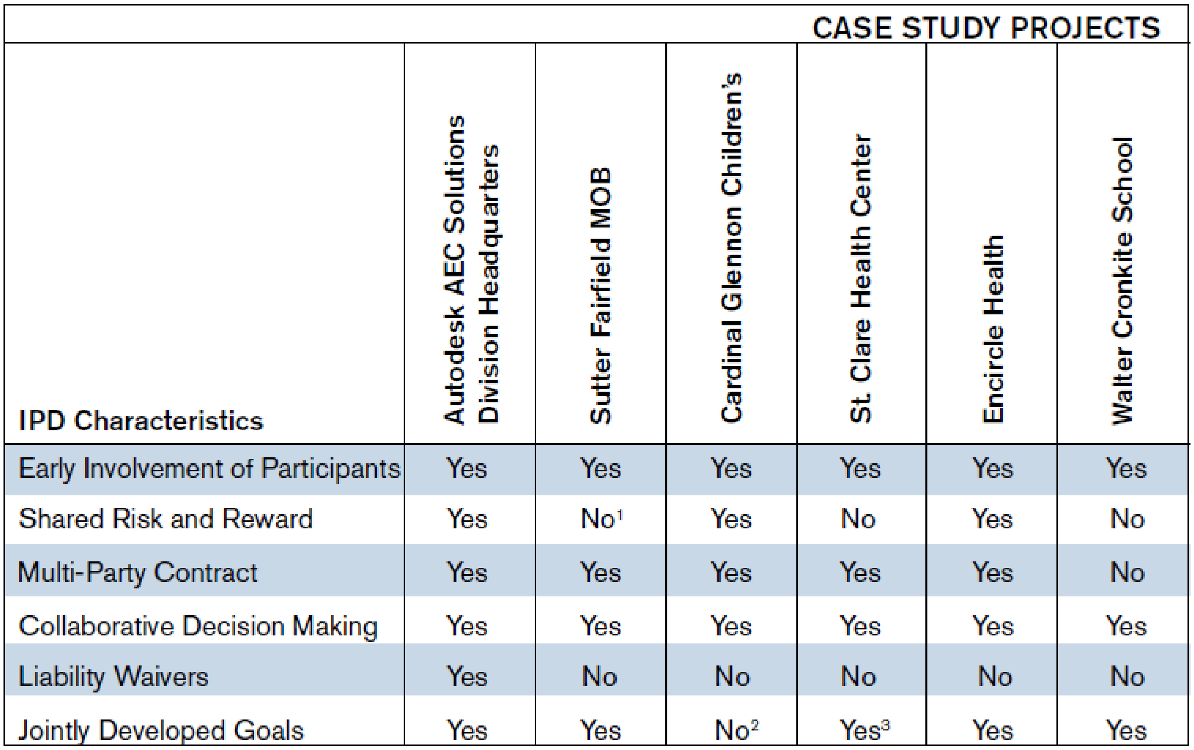
We truly appreciate all your time and effort. Your responses will be kept confidential and will further the research process and allow for the development of findings that will be useful for the success of your future projects.

Appendix D – PRELIMINARY ANALYSIS ON AIA DATASET

Before initiating the data collection phase of this research, and in order to complete a proof of concept, a preliminary analysis was performed based on a modest dataset that was publicly available. Although not compiled specifically for this research, some published project data can be found in the American Institute of Architects (AIA) 2010 report on IPD case studies, as discussed in *Chapter 2*. The AIA dataset did not satisfy all the needed elements of this study; however, it is consistent enough to demonstrate the applicability of the proposed research.

As shown in *Figure 71*, the AIA dataset includes six projects, where six independent variables labeled *IPD Characteristics* are tracked. For the purpose of this analysis, a simplified version of the 3T delivery characteristics is used as a single independent variable, scoring one point for each of the six characteristics used on any given project. The minimum and maximum scores obtained for this dataset were 3 and 6, respectively.

The published data allows for the calculation of seven performance metrics consistent with the ones established for this study: unit cost and construction cost growth, delivery speed and schedule growth, added program, total number of request for information, and change orders. These were tested against the total number of IPD characteristics, to show the effect of "*IPD-ness*" on project performance metrics. Due to the limited dataset, the analysis only consists of simple linear regressions. However, some of the results experienced what can be referred to as the law of diminishing returns: the marginal benefit or output from one input unit progressively decreases as the amount of input units is increased. Therefore, an additional logarithmic regression line is shown for metrics where this diminishing return effect is evident.



	CASE STUDY PROJECTS					
IPD Characteristics	**Autodesk AEC Solutions Division Headquarters**	**Sutter Fairfield MOB**	**Cardinal Glennon Children's**	**St. Clare Health Center**	**Encircle Health**	**Walter Cronkite School**
Early Involvement of Participants	Yes	Yes	Yes	Yes	Yes	Yes
Shared Risk and Reward	Yes	No[1]	Yes	No	Yes	No
Multi-Party Contract	Yes	Yes	Yes	Yes	Yes	No
Collaborative Decision Making	Yes	Yes	Yes	Yes	Yes	Yes
Liability Waivers	Yes	No	No	No	No	No
Jointly Developed Goals	Yes	Yes	No[2]	Yes[3]	Yes	Yes

FIGURE 70: Preliminary Analysis – AIA 2010 Case Studies (AIA 2010)

First, the results for *Added Program* are discussed. Added Program is calculated by comparing the final building area in gross square feet, to the initial programmed area. This metric was only calculated in terms of area for this preliminary analysis. However, in the final study, the definition of this metric was broadened to include metrics like added systems quality.

The regression in *Figure 71* shows, with a relatively high R^2 value, an increase in program when more IPD characteristics are used on the project. One interpretation is that

IPD delivers more value for building owners. However, this result is best explained when combined with the cost results, discussed next.

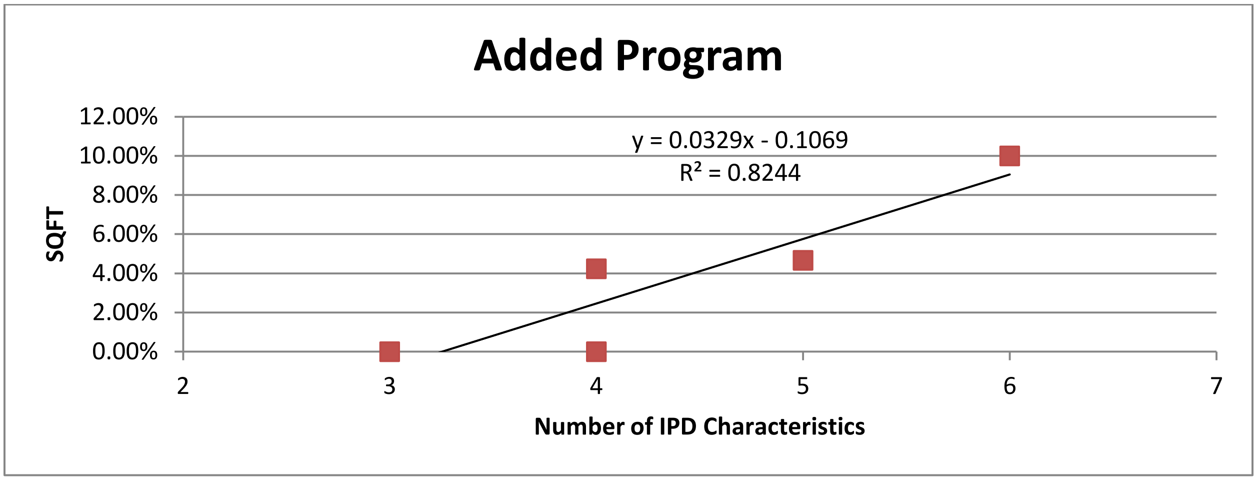


FIGURE 71: Preliminary Analysis – Added Program

The regression for *Unit Cost* is shown in the top part of *Figure 72*. Unit costs are calculated by dividing the building construction cost in dollars, by the final building area in square feet. The regression line for *Construction Cost Growth* is shown on the bottom of *Figure 72*. Cost growth is calculated by dividing the final actual construction costs, by the estimated construction costs at the time of the award. Both cost relationships show the projects with more IPD characteristics outperforming the projects with less IPD characteristics.

While the R^2 values are not especially high for the cost regressions, they are enough to confirm the validity of this research and preview the expected type of results. As mentioned earlier, the program and cost results can be combined to confirm that the measured IPD characteristics are correlated with providing building owners with more value at a lower cost.

It is interesting that the single relatively high value for cost growth (about 10%) corresponds to the only project lacking a multiparty contract. The other positive cost growth seen in this small database is a value of 1.23%, while the remaining projects either met or beat their respective cost estimates. These findings are not completely validated in the final research results. In fact, owners were receiving significant value increases for their facilities, but there were no statistically significant cost differences between the IPD and non-IPD projects.

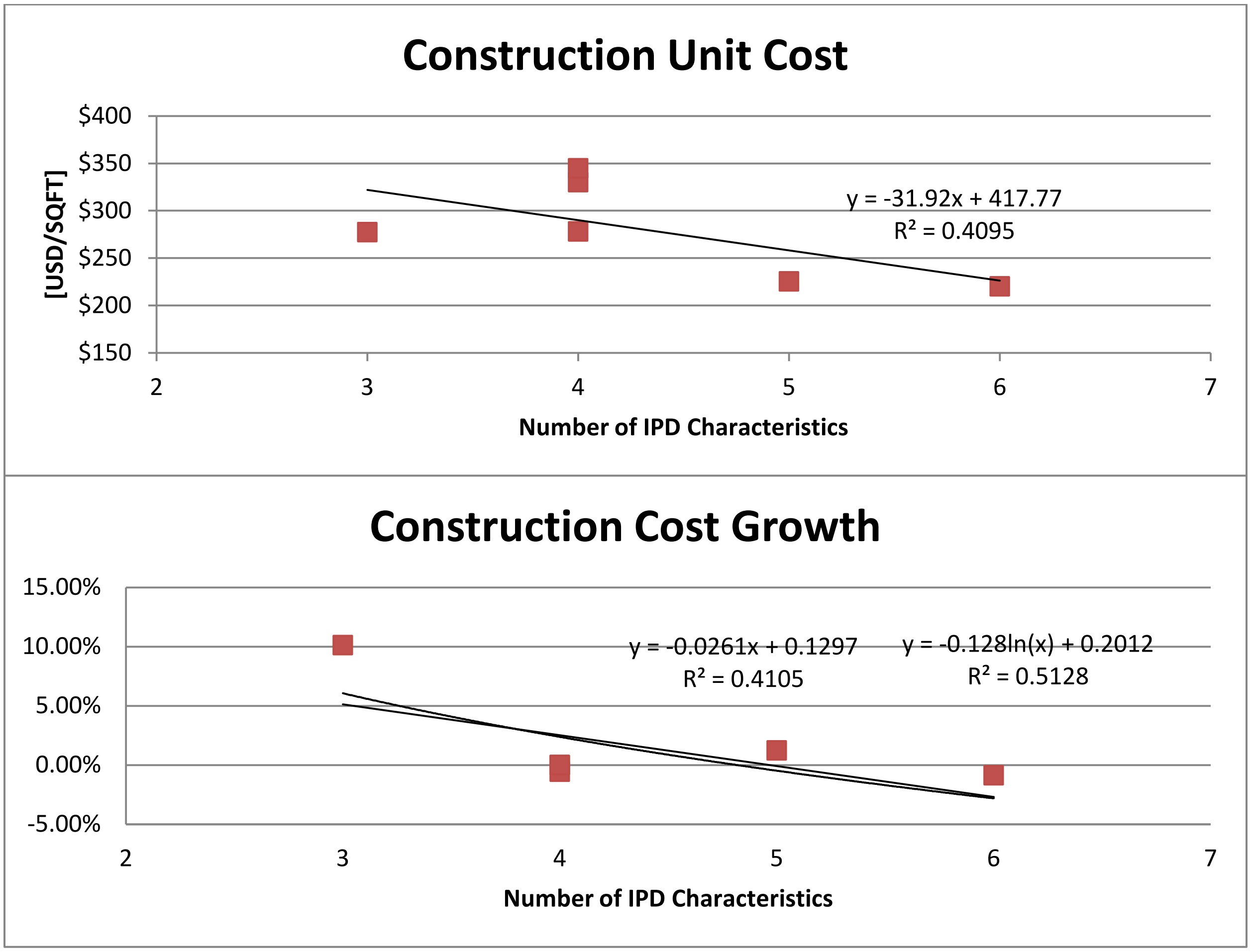


FIGURE 72: Preliminary Analysis – Unit Cost and Cost Growth

The next set of performance metrics tested includes change orders (CO), in million U.S. dollars, and the total number of requests for information (RFI). These two metrics were considered as process waste and inefficiencies in this study, given that a perfect coordination effort at the outset of a project should eliminate the need for both RFI and CO. The regressions in *Figure 73* show a decrease in RFI and CO as more IPD characteristics are used in a given project. As with the cost regressions, and as expected due to the small sample size, the model does not account for all the data variability. However, it is satisfactory enough for the purpose of this preliminary analysis.

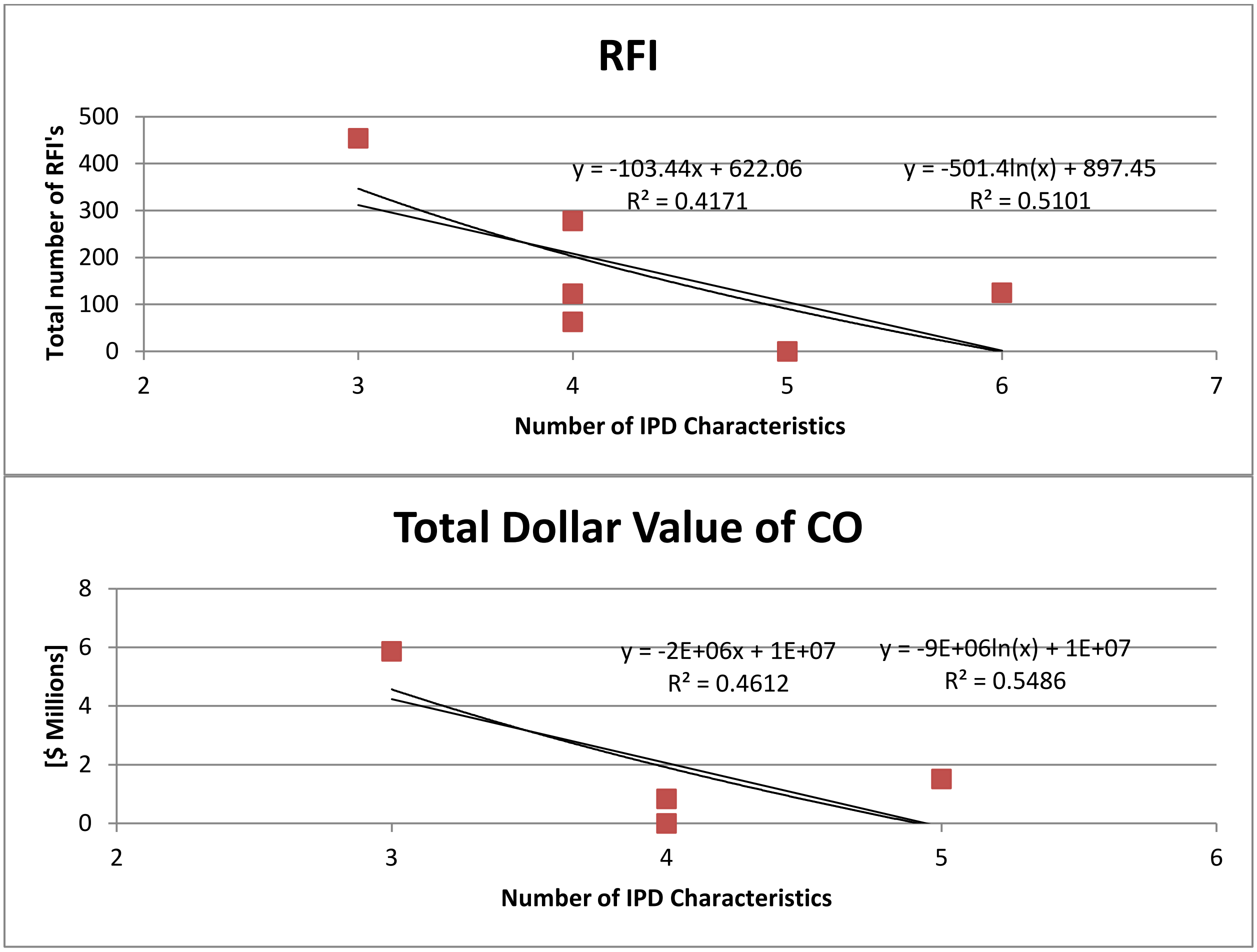


FIGURE 73: Preliminary Analysis – RFI & CO

Finally, the schedule metrics tested in this preliminary analysis are delivery speed and schedule growth. The delivery speed is calculated by dividing the final building area, in square feet, by the actual total project duration, in days. The schedule growth is calculated by comparing the achieved project duration to the initial estimate of project duration. Both of these metrics did not show an improvement for IPD projects, as illustrated in *Figure 74*.

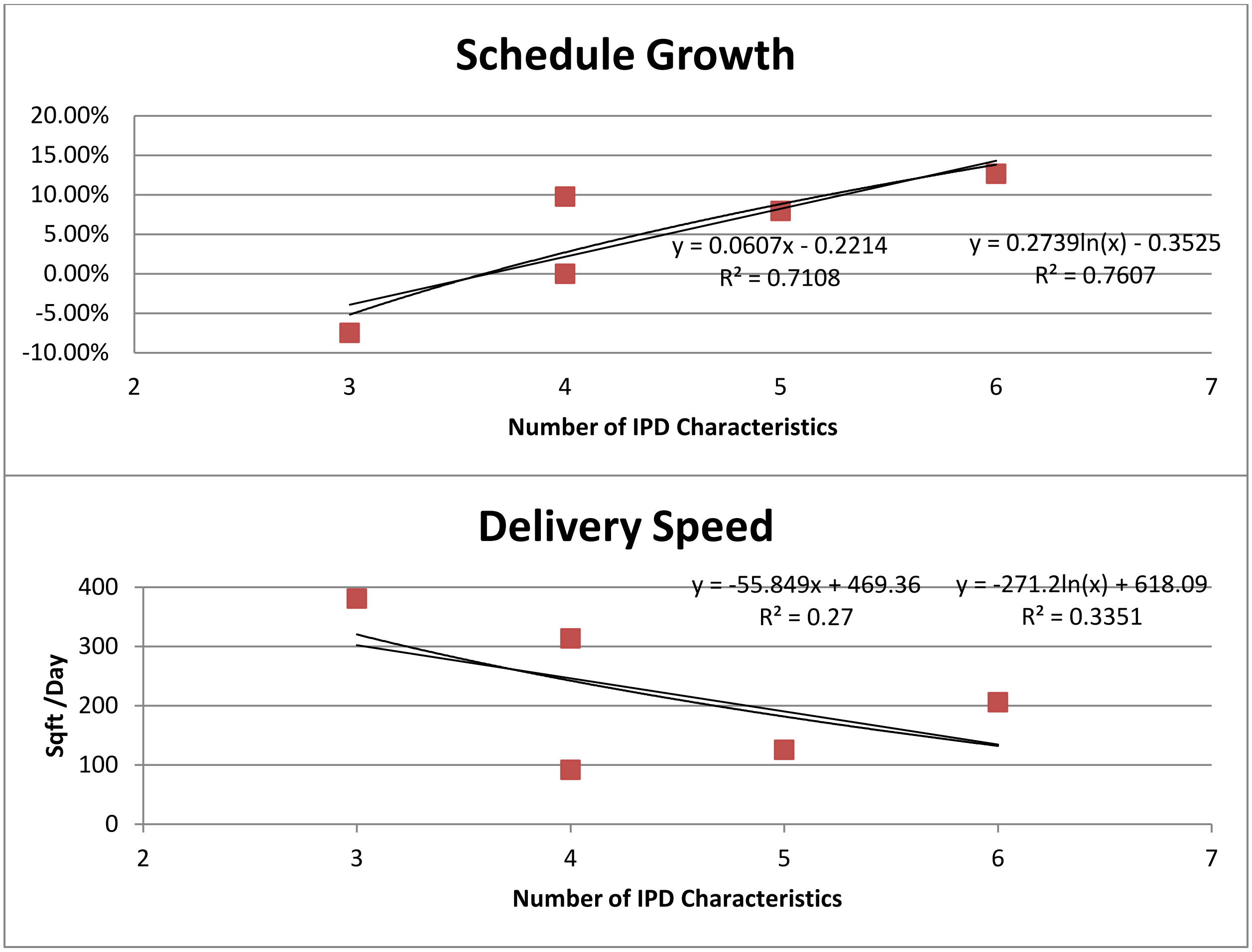


FIGURE 74: Preliminary Analysis – Schedule Growth and Delivery Speed

On the contrary, the *Schedule Growth* regressions show a larger schedule growth with more IPD characteristics, with a relatively high R^2 value. The delivery speed regression also shows inferior performance for projects with more IPD characteristics. One notable observation is that the only project not using a multiparty contract outperforms all other projects in both schedule metrics. This unexpected schedule performance is intriguing and was given special attention while conducting the research study. The final research results with a dataset of 35 projects eneded up refuting these preliminary results, and provided strong statistical proof that IPD projects result in a significantly faster delivery process.

Appendix E – MINITAB UNIVARIATE RESULTS (SELECTED)

The univariate results shown in this appendix are outputs of the Minitab software that was used to perform MWW tests, t-tests, ANOVA, and Kruskal-Wallis tests. For space purposes, only a selected set of results will be shown here. The total number of projects was 35; however, the number of observation varies between the tests based on the missing values and because the IPD-ish projects were excluded from the MWW and t-tests. The selected outputs are presented in the following order: first MWW results, then t-test results, Kruskal-Wallis results, and finally ANOVA results.

Selected MWW Results

Mann-Whitney Test and CI: UnitCost, UCIPD

```
           N  Median
UnitCost  23   306.1
UCIPD      7   371.8

Point estimate for ETA1-ETA2 is -40.3
95.0 Percent CI for ETA1-ETA2 is (-236.1,75.4)
W = 347.0
Test of ETA1 = ETA2 vs ETA1 < ETA2 is significant at 0.3295
```

Mann-Whitney Test and CI: CostGrowth, CGIPD

```
             N  Median
CostGrowth  23  0.0313
CGIPD        7  0.0300

Point estimate for ETA1-ETA2 is -0.0039
95.0 Percent CI for ETA1-ETA2 is (-0.1211,0.0709)
W = 354.5
Test of ETA1 = ETA2 vs ETA1 < ETA2 is significant at 0.4707
The test is significant at 0.4707 (adjusted for ties)
```

Mann-Whitney Test and CI: ConstSpeed, CSIPD

```
             N  Median
ConstSpeed  17   271.0
CSIPD        8   315.7

Point estimate for ETA1-ETA2 is -53.6
95.6 Percent CI for ETA1-ETA2 is (-167.5,65.4)
W = 204.0
Test of ETA1 = ETA2 vs ETA1 < ETA2 is significant at 0.1682
```

Mann-Whitney Test and CI: DelivSpeed, DSIPD

```
             N  Median
DelivSpeed  18   155.0
DSIPD        8   216.8

Point estimate for ETA1-ETA2 is -57.9
95.1 Percent CI for ETA1-ETA2 is (-125.5,14.1)
W = 214.0
Test of ETA1 = ETA2 vs ETA1 < ETA2 is significant at 0.0567
```

Mann-Whitney Test and CI: Intensity, IIPD

```
            N  Median
Intensity  17   76360
IIPD        8   87702

Point estimate for ETA1-ETA2 is -22830
95.6 Percent CI for ETA1-ETA2 is (-66455,32190)
W = 202.0
Test of ETA1 = ETA2 vs ETA1 < ETA2 is significant at 0.1406
```

Mann-Whitney Test and CI: ConSchedGrowth, CSGIPD

```
                 N  Median
ConSchedGrowth  15   0.000
CSGIPD           7   0.000

Point estimate for ETA1-ETA2 is -0.126
95.2 Percent CI for ETA1-ETA2 is (-0.451,0.001)
W = 152.0
Test of ETA1 = ETA2 vs ETA1 < ETA2 is significant at 0.0793
The test is significant at 0.0760 (adjusted for ties)
```

Mann-Whitney Test and CI: DelivSchedGrowth, DSGIPD

```
                   N  Median
DelivSchedGrowth  10  0.0000
DSGIPD             7  0.0000

Point estimate for ETA1-ETA2 is 0.0000
95.5 Percent CI for ETA1-ETA2 is (-0.3517,0.0628)
W = 84.0
Test of ETA1 = ETA2 vs ETA1 < ETA2 is significant at 0.2957
The test is significant at 0.2806 (adjusted for ties)
```

Mann-Whitney Test and CI: Systems Quality, QIPD

```
          N  Median
Quality  23   3.167
QIPD      8   3.871

Point estimate for ETA1-ETA2 is -0.698
95.0 Percent CI for ETA1-ETA2 is (-1.194,0.022)
W = 326.5
Test of ETA1 = ETA2 vs ETA1 < ETA2 is significant at 0.0321
The test is significant at 0.0320 (adjusted for ties)
```

Mann-Whitney Test and CI: Complexity, CIPD

```
             N  Median
Complexity  23  1.6818
CIPD         8  1.6938

Point estimate for ETA1-ETA2 is -0.0250
95.0 Percent CI for ETA1-ETA2 is (-0.5511,0.1331)
W = 359.5
Test of ETA1 = ETA2 vs ETA1 < ETA2 is significant at 0.3590
The test is significant at 0.3589 (adjusted for ties)
```

Mann-Whitney Test and CI: RIR, RIRIPD

```
         N  Median
RIR     16   1.629
RIRIPD   8   1.556

Point estimate for ETA1-ETA2 is -0.000
95.3 Percent CI for ETA1-ETA2 is (-1.456,2.546)
W = 203.5
Test of ETA1 = ETA2 vs ETA1 > ETA2 is significant at 0.4271
The test is significant at 0.4251 (adjusted for ties)
```

Mann-Whitney Test and CI: LTR, LTRIPD

```
         N  Median
LTR     17  0.0000
LTRIPD   8  0.0000

Point estimate for ETA1-ETA2 is 0.0000
95.6 Percent CI for ETA1-ETA2 is (0.0003,0.1391)
W = 227.0
Test of ETA1 = ETA2 vs ETA1 > ETA2 is significant at 0.3743
The test is significant at 0.3429 (adjusted for ties)
```

Mann-Whitney Test and CI: Recordables, RecIPD

```
         N  Median
Rec     23   4.776
RecIPD   8   4.155

Point estimate for ETA1-ETA2 is -0.000
```

```
95.0 Percent CI for ETA1-ETA2 is (-4.394,4.775)
W = 369.0
Test of ETA1 = ETA2 vs ETA1 > ETA2 is significant at 0.4910
The test is significant at 0.4909 (adjusted for ties)
```

Mann-Whitney Test and CI: LTI, LTIIPD

```
         N  Median
LTI     23   0.000
LTIIPD   8   0.000

Point estimate for ETA1-ETA2 is 0.000
95.0 Percent CI for ETA1-ETA2 is (0.001,2.739)
W = 396.0
Test of ETA1 = ETA2 vs ETA1 > ETA2 is significant at 0.1072
The test is significant at 0.0832 (adjusted for ties)
```

Mann-Whitney Test and CI: LaborFactor, LFIPD

```
              N  Median
LaborFactor  12   1.760
LFIPD         5   2.532

Point estimate for ETA1-ETA2 is -0.562
96.0 Percent CI for ETA1-ETA2 is (-1.389,0.257)
W = 95.0
Test of ETA1 = ETA2 vs ETA1 < ETA2 is significant at 0.0938
```

Mann-Whitney Test and CI: PPCtrend, PPCIPD

```
           N  Median
PPCtrend  16  0.0000
PPCIPD     8  0.0000

Point estimate for ETA1-ETA2 is 0.0000
95.3 Percent CI for ETA1-ETA2 is (-1.0003,0.0001)
W = 181.5
Test of ETA1 = ETA2 vs ETA1 < ETA2 is significant at 0.1352
The test is significant at 0.0721 (adjusted for ties)
```

Mann-Whitney Test and CI: ExtraLabor, ExtLabIPD

```
             N  Median
ExtraLabor  23  1.0000
ExtLabIPD    8  0.8333

Point estimate for ETA1-ETA2 is 0.3334
95.0 Percent CI for ETA1-ETA2 is (-0.3334,0.6665)
W = 384.5
Test of ETA1 = ETA2 vs ETA1 > ETA2 is significant at 0.2351
The test is significant at 0.2297 (adjusted for ties)
```

Mann-Whitney Test and CI: RFI, RFIIPD

```
         N  Median
RFI     21   9.610
RFIIPD   7   1.813
```

```
Point estimate for ETA1-ETA2 is 8.226
95.0 Percent CI for ETA1-ETA2 is (4.099,14.877)
W = 366.0
Test of ETA1 = ETA2 vs ETA1 > ETA2 is significant at 0.0006
The test is significant at 0.0006 (adjusted for ties)
```

Mann-Whitney Test and CI: RFI Processing Time, RFITIPD

```
          N  Median
RFIT     22   2.000
RFITIPD   8   1.000

Point estimate for ETA1-ETA2 is 1.000
95.4 Percent CI for ETA1-ETA2 is (-0.000,2.000)
W = 381.0
Test of ETA1 = ETA2 vs ETA1 > ETA2 is significant at 0.0320
The test is significant at 0.0253 (adjusted for ties)
```

Mann-Whitney Test and CI: Change, ChangeIPD

```
            N  Median
Change     21  0.0300
ChangeIPD   8  0.0450

Point estimate for ETA1-ETA2 is -0.0100
95.2 Percent CI for ETA1-ETA2 is (-0.1001,0.0600)
W = 305.0
Test of ETA1 = ETA2 vs ETA1 < ETA2 is significant at 0.3215
The test is significant at 0.3213 (adjusted for ties)
```

Mann-Whitney Test and CI: Change Absolute, ChABSipd

```
           N  Median
ChABS     21  0.0800
ChABSipd   8  0.0450

Point estimate for ETA1-ETA2 is 0.0245
95.2 Percent CI for ETA1-ETA2 is (-0.0260,0.0900)
W = 331.0
Test of ETA1 = ETA2 vs ETA1 > ETA2 is significant at 0.2247
The test is significant at 0.2244 (adjusted for ties)
```

Mann-Whitney Test and CI: ChangeProgram, ChPrIPD

```
             N  Median
ChangeProg  22  0.7750
ChPrIPD      8  0.7000

Point estimate for ETA1-ETA2 is 0.0000
95.4 Percent CI for ETA1-ETA2 is (-0.2001,0.5003)
W = 350.5
Test of ETA1 = ETA2 vs ETA1 > ETA2 is significant at 0.3365
The test is significant at 0.3336 (adjusted for ties)
```

Mann-Whitney Test and CI: Change Design, ChDeIPD

```
            N  Median
```

```
ChangeDes  22  0.1000
ChDeIPD     8  0.0185

Point estimate for ETA1-ETA2 is 0.0500
95.4 Percent CI for ETA1-ETA2 is (-0.0000,0.2500)
W = 381.0
Test of ETA1 = ETA2 vs ETA1 > ETA2 is significant at 0.0320
The test is significant at 0.0290 (adjusted for ties)
```

Mann-Whitney Test and CI: ChangeReg, ChReIPD

```
            N  Median
ChangeReg  21  0.0500
ChReIPD     8  0.0000

Point estimate for ETA1-ETA2 is 0.0120
95.2 Percent CI for ETA1-ETA2 is (0.0000,0.0500)
W = 347.5
Test of ETA1 = ETA2 vs ETA1 > ETA2 is significant at 0.0592
The test is significant at 0.0499 (adjusted for ties)
```

Mann-Whitney Test and CI: COTime, COTIPD

```
         N  Median
COTime  23   4.000
COTIPD   8   1.000

Point estimate for ETA1-ETA2 is 3.000
95.0 Percent CI for ETA1-ETA2 is (2.000,5.000)
W = 443.0
Test of ETA1 = ETA2 vs ETA1 > ETA2 is significant at 0.0004
The test is significant at 0.0003 (adjusted for ties)
```

Mann-Whitney Test and CI: Rework, ReworkIPD

```
            N  Median
Rework     22   1.000
ReworkIPD   8   1.000

Point estimate for ETA1-ETA2 is 0.000
95.4 Percent CI for ETA1-ETA2 is (-1.000,-0.000)
W = 326.5
Test of ETA1 = ETA2 vs ETA1 < ETA2 is significant at 0.2557
The test is significant at 0.1732 (adjusted for ties)
```

Mann-Whitney Test and CI: Resubmittals, ReSubIPD

```
               N  Median
Resubmittals  14   1.438
ReSubIPD       6   0.196

Point estimate for ETA1-ETA2 is 0.935
95.7 Percent CI for ETA1-ETA2 is (0.012,1.799)
W = 173.0
Test of ETA1 = ETA2 vs ETA1 > ETA2 is significant at 0.0177
```

Mann-Whitney Test and CI: Deficiencies, DefIPD

```
               N  Median
Deficiencies  13   1.425
DefIPD         8   0.207

Point estimate for ETA1-ETA2 is 1.382
95.4 Percent CI for ETA1-ETA2 is (0.506,2.897)
W = 185.0
Test of ETA1 = ETA2 vs ETA1 > ETA2 is significant at 0.0013
The test is significant at 0.0013 (adjusted for ties)
```

Mann-Whitney Test and CI: TonWaste, WasteIPD

```
           N  Median
TonWaste  12    98.6
WasteIPD   5    50.2

Point estimate for ETA1-ETA2 is 54.0
96.0 Percent CI for ETA1-ETA2 is (-26.2,226.7)
W = 122.0
Test of ETA1 = ETA2 vs ETA1 > ETA2 is significant at 0.0774
```

Mann-Whitney Test and CI: Recycled, RecyIPD

```
           N  Median
Recycled  20  0.8400
RecyIPD    6  0.8725

Point estimate for ETA1-ETA2 is -0.0325
95.2 Percent CI for ETA1-ETA2 is (-0.2100,0.0901)
W = 258.0
Test of ETA1 = ETA2 vs ETA1 < ETA2 is significant at 0.2420
The test is significant at 0.2418 (adjusted for ties)
```

Mann-Whitney Test and CI: Landfill, LandfIPD

```
           N  Median
Landfill  20  0.1600
LandfIPD   6  0.1275

Point estimate for ETA1-ETA2 is 0.0300
95.2 Percent CI for ETA1-ETA2 is (-0.0901,0.2102)
W = 282.0
Test of ETA1 = ETA2 vs ETA1 > ETA2 is significant at 0.2420
The test is significant at 0.2418 (adjusted for ties)
```

Mann-Whitney Test and CI: OHProfit, OPHIPD

```
           N  Median
OHProfit  21   1.000
OPHIPD     8   2.000

Point estimate for ETA1-ETA2 is -1.000
95.2 Percent CI for ETA1-ETA2 is (-1.000,-0.000)
W = 279.0
Test of ETA1 = ETA2 vs ETA1 < ETA2 is significant at 0.0416
```

```
The test is significant at 0.0241 (adjusted for ties)
```

Mann-Whitney Test and CI: ReturnBusiness, RBIPD

```
                 N  Median
ReturnBusiness  23  2.0000
RBIPD            8  2.0000

Point estimate for ETA1-ETA2 is 0.0000
95.0 Percent CI for ETA1-ETA2 is (-0.0000,0.0001)
W = 354.5
Test of ETA1 = ETA2 vs ETA1 < ETA2 is significant at 0.2786
The test is significant at 0.2106 (adjusted for ties)
```

Mann-Whitney Test and CI: PunchlistItems, PunItIPD

```
                 N  Median
PunchlistItems  22   32.39
PunItIPD         6    8.98

Point estimate for ETA1-ETA2 is 23.05
95.3 Percent CI for ETA1-ETA2 is (2.82,48.18)
W = 359.0
Test of ETA1 = ETA2 vs ETA1 > ETA2 is significant at 0.0135
```

Mann-Whitney Test and CI: PunchlistCost, PunCoIPD

```
                N  Median
PunchlistCost  22  2.0000
PunCoIPD        7  1.0000

Point estimate for ETA1-ETA2 is 1.0000
95.0 Percent CI for ETA1-ETA2 is (0.0001,0.9998)
W = 380.0
Test of ETA1 = ETA2 vs ETA1 > ETA2 is significant at 0.0058
The test is significant at 0.0025 (adjusted for ties)
```

Mann-Whitney Test and CI: WarrantyCost, WarrIPD

```
               N  Median
WarrantyCost  18  1.0000
WarrIPD        7  1.0000

Point estimate for ETA1-ETA2 is 0.0000
95.1 Percent CI for ETA1-ETA2 is (0.0000,1.0000)
W = 256.0
Test of ETA1 = ETA2 vs ETA1 > ETA2 is significant at 0.0966
The test is significant at 0.0402 (adjusted for ties)
```

Mann-Whitney Test and CI: LatentDefect, LatIPD

```
               N  Median
LatentDefect  17  1.0000
LatIPD         7  1.0000

Point estimate for ETA1-ETA2 is -0.0000
95.1 Percent CI for ETA1-ETA2 is (-0.9999,0.0002)
W = 210.0
```

```
Test of ETA1 = ETA2 vs ETA1 < ETA2 is significant at 0.4495
The test is significant at 0.4415 (adjusted for ties)
```

Selected t-test Results

Two-Sample T-Test and CI: UnitCost, UCIPD

```
Two-sample T for UnitCost vs UCIPD

                             SE
           N   Mean  StDev  Mean
UnitCost  23  315.7   84.6    18
UCIPD      7    402    214    81

Difference = mu (UnitCost) - mu (UCIPD)
Estimate for difference:  -86.3
95% upper bound for difference:  74.3
T-Test of difference = 0 (vs <): T-Value = -1.04  P-Value = 0.168  DF = 6
```

Two-Sample T-Test and CI: CostGrowth, CGIPD

```
Two-sample T for CostGrowth vs CGIPD

             N    Mean   StDev  SE Mean
CostGrowth  23   0.048   0.117    0.024
CGIPD        7  0.0628  0.0982    0.037

Difference = mu (CostGrowth) - mu (CGIPD)
Estimate for difference:  -0.0151
95% upper bound for difference:  0.0648
T-Test of difference = 0 (vs <): T-Value = -0.34  P-Value = 0.370  DF = 11
```

Two-Sample T-Test and CI: ConstSpeed, CSIPD

```
Two-sample T for ConstSpeed vs CSIPD

                               SE
             N   Mean  StDev  Mean
ConstSpeed  17    275    130    32
CSIPD        8  318.6   96.7    34

Difference = mu (ConstSpeed) - mu (CSIPD)
Estimate for difference:  -43.5
95% upper bound for difference:  37.1
T-Test of difference = 0 (vs <): T-Value = -0.94  P-Value = 0.181  DF = 18
```

Two-Sample T-Test and CI: DelivSpeed, DSIPD

```
Two-sample T for DelivSpeed vs DSIPD

                               SE
             N   Mean  StDev  Mean
DelivSpeed  18  164.9   81.2    19
DSIPD        8  218.7   65.3    23
```

```
Difference = mu (DelivSpeed) - mu (DSIPD)
Estimate for difference:  -53.9
95% upper bound for difference:  -1.5
T-Test of difference = 0 (vs <): T-Value = -1.80  P-Value = 0.046  DF = 16
```

Two-Sample T-Test and CI: Intensity, IIPD

```
Two-sample T for Intensity vs IIPD

            N    Mean  StDev  SE Mean
Intensity  17   84119  52654    12770
IIPD        8  106702  51763    18301

Difference = mu (Intensity) - mu (IIPD)
Estimate for difference:  -22583
95% upper bound for difference:  16723
T-Test of difference = 0 (vs <): T-Value = -1.01  P-Value = 0.164  DF = 14
```

Two-Sample T-Test and CI: DelivSchedGrowth, DSGIPD

```
Two-sample T for DelivSchedGrowth vs DSGIPD

                   N    Mean   StDev  SE Mean
DelivSchedGrowth  10  0.0005  0.0629    0.020
DSGIPD             7   0.175   0.395     0.15

Difference = mu (DelivSchedGrowth) - mu (DSGIPD)
Estimate for difference:  -0.175
95% CI for difference:  (-0.543, 0.193)
T-Test of difference = 0 (vs not =): T-Value = -1.16  P-Value = 0.289  DF = 6
```

Two-Sample T-Test and CI: Quality, QIPD

```
Two-sample T for Quality vs QIPD

          N   Mean  StDev  SE Mean
Quality  23  3.199  0.608     0.13
QIPD      8  3.835  0.814     0.29

Difference = mu (Quality) - mu (QIPD)
Estimate for difference:  -0.636
95% upper bound for difference:  -0.060
T-Test of difference = 0 (vs <): T-Value = -2.02  P-Value = 0.037  DF = 9
```

Two-Sample T-Test and CI: Complexity, CIPD

```
Two-sample T for Complexity vs CIPD

             N   Mean  StDev  SE Mean
Complexity  23  1.498  0.394    0.082
CIPD         8  1.640  0.229    0.081

Difference = mu (Complexity) - mu (CIPD)
Estimate for difference:  -0.141
95% upper bound for difference:  0.057
T-Test of difference = 0 (vs <): T-Value = -1.23  P-Value = 0.117  DF = 21
```

Two-Sample T-Test and CI: RIR, RIRIPD

```
Two-sample T for RIR vs RIRIPD

         N  Mean  StDev  SE Mean
RIR     16  2.22   2.57     0.64
RIRIPD   8  1.53   1.19     0.42


Difference = mu (RIR) - mu (RIRIPD)
Estimate for difference:  0.689
95% lower bound for difference:  -0.632
T-Test of difference = 0 (vs >): T-Value = 0.90  P-Value = 0.190  DF = 21
```

Two-Sample T-Test and CI: LTR, LTRIPD

```
Two-sample T for LTR vs LTRIPD

         N   Mean  StDev  SE Mean
LTR     17  0.370  0.798     0.19
LTRIPD   8  0.064  0.125    0.044

Difference = mu (LTR) - mu (LTRIPD)
Estimate for difference:  0.307
95% lower bound for difference:  -0.039
T-Test of difference = 0 (vs >): T-Value = 1.55  P-Value = 0.070  DF = 17
```

Two-Sample T-Test and CI: Rec, RecIPD

```
Two-sample T for Rec vs RecIPD

         N  Mean  StDev  SE Mean
Rec     23  5.99   6.04      1.3
RecIPD   8  5.70   5.05      1.8

Difference = mu (Rec) - mu (RecIPD)
Estimate for difference:  0.28
95% lower bound for difference:  -3.57
T-Test of difference = 0 (vs >): T-Value = 0.13  P-Value = 0.450  DF = 14
```

Two-Sample T-Test and CI: LTI, LTIIPD

```
Two-sample T for LTI vs LTIIPD

         N   Mean  StDev  SE Mean
LTI     23   1.94   2.78     0.58
LTIIPD   8  0.361  0.816     0.29

Difference = mu (LTI) - mu (LTIIPD)
Estimate for difference:  1.581
95% lower bound for difference:  0.481
T-Test of difference = 0 (vs >): T-Value = 2.44  P-Value = 0.011  DF = 28
```

Two-Sample T-Test and CI: LaborFactor, LFIPD

```
Two-sample T for LaborFactor vs LFIPD

              N   Mean  StDev  SE Mean
```

```
LaborFactor  12  1.857  0.589     0.17
LFIPD         5  2.408  0.734     0.33

Difference = mu (LaborFactor) - mu (LFIPD)
Estimate for difference:  -0.551
95% upper bound for difference:  0.168
T-Test of difference = 0 (vs <): T-Value = -1.49  P-Value = 0.093  DF = 6
```

Two-Sample T-Test and CI: PPCtrend, PPCIPD

```
Two-sample T for PPCtrend vs PPCIPD

           N   Mean  StDev  SE Mean
PPCtrend  16  0.063  0.443     0.11
PPCIPD     8  0.375  0.518     0.18

Difference = mu (PPCtrend) - mu (PPCIPD)
Estimate for difference:  -0.313
95% upper bound for difference:  0.069
T-Test of difference = 0 (vs <): T-Value = -1.46  P-Value = 0.085  DF = 12
```

Two-Sample T-Test and CI: ExtraLabor, ExtLabIPD

```
Two-sample T for ExtraLabor vs ExtLabIPD

             N   Mean  StDev  SE Mean
ExtraLabor  23  1.000  0.471    0.098
ExtLabIPD    8  0.792  0.589     0.21

Difference = mu (ExtraLabor) - mu (ExtLabIPD)
Estimate for difference:  0.208
95% lower bound for difference:  -0.209
T-Test of difference = 0 (vs >): T-Value = 0.90  P-Value = 0.194  DF = 10
```

Two-Sample T-Test and CI: RFI, RFIIPD

```
Two-sample T for RFI vs RFIIPD

         N   Mean  StDev  SE Mean
RFI     21  12.65   7.95      1.7
RFIIPD   7   2.82   3.21      1.2

Difference = mu (RFI) - mu (RFIIPD)
Estimate for difference:  9.83
95% lower bound for difference:  6.21
T-Test of difference = 0 (vs >): T-Value = 4.64  P-Value = 0.000  DF = 24
```

Two-Sample T-Test and CI: RFIT, RFITIPD

```
Two-sample T for RFIT vs RFITIPD

          N   Mean  StDev  SE Mean
RFIT     22  1.955  0.899     0.19
RFITIPD   8  1.125  0.835     0.30

Difference = mu (RFIT) - mu (RFITIPD)
Estimate for difference:  0.830
95% lower bound for difference:  0.207
```

```
T-Test of difference = 0 (vs >): T-Value = 2.36  P-Value = 0.017  DF = 13
```

Two-Sample T-Test and CI: Change, ChangeIPD

```
Two-sample T for Change vs ChangeIPD

            N    Mean   StDev  SE Mean
Change     21   0.056   0.126    0.027
ChangeIPD   8  0.0765  0.0978    0.035

Difference = mu (Change) - mu (ChangeIPD)
Estimate for difference:  -0.0205
95% CI for difference:  (-0.1141, 0.0731)
T-Test of difference = 0 (vs not =): T-Value = -0.46  P-Value = 0.649  DF = 16
```

Two-Sample T-Test and CI: Change, ChangeIPD

```
Two-sample T for ChABS vs ChABSipd

           N    Mean   StDev  SE Mean
ChABS     21  0.1022  0.0901    0.020
ChABSipd   8  0.0765  0.0978    0.035

Difference = mu (ChABS) - mu (ChABSipd)
Estimate for difference:  0.0257
95% lower bound for difference:  -0.0457
T-Test of difference = 0 (vs >): T-Value = 0.65  P-Value = 0.266  DF = 11
```

Two-Sample T-Test and CI: ChangeProg, ChPrIPD

```
Two-sample T for ChangeProg vs ChPrIPD

             N   Mean  StDev  SE Mean
ChangeProg  22  0.648  0.359    0.077
ChPrIPD      8  0.539  0.473     0.17

Difference = mu (ChangeProg) - mu (ChPrIPD)
Estimate for difference:  0.109
95% lower bound for difference:  -0.224
T-Test of difference = 0 (vs >): T-Value = 0.59  P-Value = 0.283  DF = 10
```

Two-Sample T-Test and CI: ChangeDes, ChDeIPD

```
Two-sample T for ChangeDes vs ChDeIPD

            N    Mean   StDev  SE Mean
ChangeDes  22   0.187   0.214    0.046
ChDeIPD     8  0.0296  0.0365    0.013

Difference = mu (ChangeDes) - mu (ChDeIPD)
Estimate for difference:  0.1570
95% lower bound for difference:  0.0760
T-Test of difference = 0 (vs >): T-Value = 3.32  P-Value = 0.001  DF = 24
```

Two-Sample T-Test and CI: ChangeReg, ChReIPD

```
Two-sample T for ChangeReg vs ChReIPD
```

```
            N    Mean   StDev  SE Mean
ChangeReg  21   0.107   0.225    0.049
ChReIPD     8  0.0173  0.0241   0.0085

Difference = mu (ChangeReg) - mu (ChReIPD)
Estimate for difference:  0.0901
95% lower bound for difference:  0.0044
T-Test of difference = 0 (vs >): T-Value = 1.81  P-Value = 0.042  DF = 21
```

Two-Sample T-Test and CI: COTime, COTIPD

```
Two-sample T for COTime vs COTIPD

         N  Mean  StDev  SE Mean
COTime  23  4.52   2.04     0.43
COTIPD   8  1.31   1.10     0.39

Difference = mu (COTime) - mu (COTIPD)
Estimate for difference:  3.209
95% lower bound for difference:  2.221
T-Test of difference = 0 (vs >): T-Value = 5.57  P-Value = 0.000  DF = 23
```

Two-Sample T-Test and CI: Rework, ReworkIPD

```
Two-sample T for Rework vs ReworkIPD

            N   Mean  StDev  SE Mean
Rework     22  1.091  0.426    0.091
ReworkIPD   8   1.50   1.07     0.38

Difference = mu (Rework) - mu (ReworkIPD)
Estimate for difference:  -0.409
95% upper bound for difference:  0.327
T-Test of difference = 0 (vs <): T-Value = -1.05  P-Value = 0.164  DF = 7
```

Two-Sample T-Test and CI: Resubmittals, ReSubIPD

```
Two-sample T for Resubmittals vs ReSubIPD

               N   Mean  StDev  SE Mean
Resubmittals  14   2.68   4.32      1.2
ReSubIPD       6  0.370  0.376     0.15

Difference = mu (Resubmittals) - mu (ReSubIPD)
Estimate for difference:  2.31
95% lower bound for difference:  0.25
T-Test of difference = 0 (vs >): T-Value = 1.99  P-Value = 0.034  DF = 13
```

Two-Sample T-Test and CI: Deficiencies, DefIPD

```
Two-sample T for Deficiencies vs DefIPD

               N   Mean  StDev  SE Mean
Deficiencies  13   2.15   1.72     0.48
DefIPD         8  0.328  0.380     0.13

Difference = mu (Deficiencies) - mu (DefIPD)
Estimate for difference:  1.818
```

```
95% lower bound for difference:  0.942
T-Test of difference = 0 (vs >): T-Value = 3.67  P-Value = 0.001  DF = 13
```

Two-Sample T-Test and CI: TonWaste, WasteIPD

```
Two-sample T for TonWaste vs WasteIPD

                              SE
           N  Mean  StDev  Mean
TonWaste  12   135    123    36
WasteIPD   5  48.0   38.8    17

Difference = mu (TonWaste) - mu (WasteIPD)
Estimate for difference:  87.2
95% lower bound for difference:  17.5
T-Test of difference = 0 (vs >): T-Value = 2.20  P-Value = 0.022  DF = 14
```

Two-Sample T-Test and CI: Recycled, RecyIPD

```
Two-sample T for Recycled vs RecyIPD

           N   Mean  StDev  SE Mean
Recycled  20  0.688  0.345    0.077
RecyIPD    6  0.823  0.153    0.062

Difference = mu (Recycled) - mu (RecyIPD)
Estimate for difference:  -0.1341
95% upper bound for difference:  0.0377
T-Test of difference = 0 (vs <): T-Value = -1.35  P-Value = 0.096  DF = 19
```

Two-Sample T-Test and CI: OHProfit, OPHIPD

```
Two-sample T for OHProfit vs OPHIPD

           N   Mean  StDev  SE Mean
OHProfit  21  1.238  0.539     0.12
OPHIPD     8  1.875  0.835     0.30

Difference = mu (OHProfit) - mu (OPHIPD)
Estimate for difference:  -0.637
95% upper bound for difference:  -0.055
T-Test of difference = 0 (vs <): T-Value = -2.01  P-Value = 0.038  DF = 9
```

Two-Sample T-Test and CI: ReturnBusiness, RBIPD

```
Two-sample T for ReturnBusiness vs RBIPD

                 N   Mean  StDev  SE Mean
ReturnBusiness  23   1.52   1.04     0.22
RBIPD            8  1.875  0.354     0.13

Difference = mu (ReturnBusiness) - mu (RBIPD)
Estimate for difference:  -0.353
95% upper bound for difference:  0.072
T-Test of difference = 0 (vs <): T-Value = -1.41  P-Value = 0.084  DF = 28
```

Two-Sample T-Test and CI: PunchlistItems, PunItIPD

```
Two-sample T for PunchlistItems vs PunItIPD

                 N  Mean  StDev  SE Mean
PunchlistItems  22  40.1   36.2      7.7
PunItIPD         6  8.75   5.40      2.2

Difference = mu (PunchlistItems) - mu (PunItIPD)
Estimate for difference:  31.32
95% lower bound for difference:  17.56
T-Test of difference = 0 (vs >): T-Value = 3.90  P-Value = 0.000  DF = 23
```

Two-Sample T-Test and CI: PunchlistCost, PunCoIPD

```
Two-sample T for PunchlistCost vs PunCoIPD

                N   Mean  StDev  SE Mean
PunchlistCost  22  1.682  0.646     0.14
PunCoIPD        7  0.857  0.378     0.14

Difference = mu (PunchlistCost) - mu (PunCoIPD)
Estimate for difference:  0.825
95% lower bound for difference:  0.479
T-Test of difference = 0 (vs >): T-Value = 4.15  P-Value = 0.000  DF = 17
```

Two-Sample T-Test and CI: WarrantyCost, WarrIPD

```
Two-sample T for WarrantyCost vs WarrIPD

               N   Mean  StDev  SE Mean
WarrantyCost  18  0.944  0.416    0.098
WarrIPD        7  0.571  0.535     0.20

Difference = mu (WarrantyCost) - mu (WarrIPD)
Estimate for difference:  0.373
95% lower bound for difference:  -0.045
T-Test of difference = 0 (vs >): T-Value = 1.66  P-Value = 0.068  DF = 8
```

Two-Sample T-Test and CI: LatentDefect, LatIPD

```
Two-sample T for LatentDefect vs LatIPD

               N   Mean  StDev  SE Mean
LatentDefect  17  0.529  0.514     0.12
LatIPD         7  0.571  0.535     0.20

Difference = mu (LatentDefect) - mu (LatIPD)
Estimate for difference:  -0.042
95% upper bound for difference:  0.388
T-Test of difference = 0 (vs <): T-Value = -0.18  P-Value = 0.432  DF = 10
```

Selected Kruskal-Wallis Results

Kruskal-Wallis Test: Quality versus PDS

```
Kruskal-Wallis Test on Quality

PDS       N  Median  Ave Rank      Z
0         7   2.778       8.7  -2.68
1        13   3.556      20.2   0.96
2         5   3.111      15.3  -0.64
3        10   3.871      23.1   1.84
Overall  35              18.0

H = 9.10  DF = 3  P = 0.028
H = 9.10  DF = 3  P = 0.028  (adjusted for ties)
```

Kruskal-Wallis Test: Unit Cost versus PDS

```
34 cases were used
1 cases contained missing values

Kruskal-Wallis Test on UnitCost

PDS       N  Median  Ave Rank      Z
0         7   354.2      19.1   0.49
1        13   304.1      15.5  -0.94
2         5   300.1      15.2  -0.56
3         9   371.8      20.4   1.03
Overall  34              17.5

H = 1.79  DF = 3  P = 0.617
```

Kruskal-Wallis Test: Cost Growth versus PDS

```
34 cases were used
1 cases contained missing values

Kruskal-Wallis Test on CostGrowth

PDS       N    Median  Ave Rank      Z
0         7   0.10100      23.9   1.92
1        13   0.02857      16.2  -0.62
2         5  -0.04082      11.0  -1.58
3         9   0.03000      18.1   0.20
Overall  34                17.5

H = 5.31  DF = 3  P = 0.150
H = 5.31  DF = 3  P = 0.150  (adjusted for ties)
```

Kruskal-Wallis Test: Recordables versus PDS

```
Kruskal-Wallis Test on RecPer100Mil

PDS       N  Median  Ave Rank      Z
0         7   4.852      19.3   0.37
1        13   5.797      20.6   1.16
2         5   3.266      11.4  -1.56
3        10   3.413      17.0  -0.37
Overall  35              18.0

H = 3.13  DF = 3  P = 0.373
H = 3.23  DF = 3  P = 0.358  (adjusted for ties)
```

Kruskal-Wallis Test: LTI versus PDS

```
Kruskal-Wallis Test on LTIper100Mil

PDS       N       Median  Ave Rank      Z
0         7  0.000000000      17.6  -0.12
1        13  1.591976439      22.1   1.81
2         5  0.000000000      15.4  -0.61
3        10  0.000000000      14.3  -1.35
Overall  35                   18.0

H = 3.70  DF = 3  P = 0.296
H = 4.91  DF = 3  P = 0.178  (adjusted for ties)
```

Kruskal-Wallis Test: PPC trend versus PDS

```
27 cases were used
8 cases contained missing values

Kruskal-Wallis Test on PPCtrend

PDS       N       Median  Ave Rank      Z
0         5  0.000000000      12.0  -0.62
1         9  0.000000000      11.5  -1.16
2         3  0.000000000      15.8   0.42
3        10  0.000000000      16.7   1.36
Overall  27                   14.0

H = 2.53  DF = 3  P = 0.470
H = 4.33  DF = 3  P = 0.228  (adjusted for ties)
```

Kruskal-Wallis Test: Extra Labor versus PDS

```
Kruskal-Wallis Test on ExtraLabor

PDS       N  Median  Ave Rank      Z
0         7  1.0000      20.6   0.76
1        13  1.0000      21.1   1.37
2         5  0.3333       9.0  -2.12
3        10  0.8333      16.6  -0.49
Overall  35              18.0

H = 5.67  DF = 3  P = 0.129
H = 5.95  DF = 3  P = 0.114  (adjusted for ties)
```

Kruskal-Wallis Test: RFI versus PDS

```
32 cases were used
3 cases contained missing values

Kruskal-Wallis Test on RFIperMil

PDS        N  Median  Ave Rank      Z
0          6  20.376      26.3   2.85
1         12   9.264      20.2   1.71
2          5   3.609      10.6  -1.53
3          9   1.942       8.3  -3.08
Overall   32              16.5

H = 17.22  DF = 3  P = 0.001
H = 17.23  DF = 3  P = 0.001  (adjusted for ties)
```

Kruskal-Wallis Test: RFI Processing Time versus PDS

```
34 cases were used
1 cases contained missing values

Kruskal-Wallis Test on RFITime

PDS        N  Median  Ave Rank      Z
0          6   2.500      24.8   1.99
1         13   2.000      17.4  -0.05
2          5   2.000      18.6   0.27
3         10   1.000      12.7  -1.81
Overall   34              17.5

H = 5.64  DF = 3  P = 0.131
H = 6.37  DF = 3  P = 0.095  (adjusted for ties)
```

Kruskal-Wallis Test: Total Change versus PDS

```
33 cases were used
2 cases contained missing values

Kruskal-Wallis Test on ChangeABS

PDS        N   Median  Ave Rank      Z
0          6  0.12125      24.3   2.05
1         12  0.02750      14.8  -0.99
2          5  0.08000      19.3   0.58
3         10  0.04500      14.1  -1.14
Overall   33               17.0

H = 5.26  DF = 3  P = 0.154
H = 5.27  DF = 3  P = 0.153  (adjusted for ties)
```

Kruskal-Wallis Test: Additions/Program Changes versus PDS

```
34 cases were used
1 cases contained missing values

Kruskal-Wallis Test on ChangeProg
```

```
PDS       N  Median  Ave Rank      Z
0         6  0.7750      19.4   0.52
1        13  0.7500      18.4   0.43
2         5  0.7000      14.8  -0.66
3        10  0.7000      16.5  -0.38
Overall  34              17.5

H = 0.80  DF = 3  P = 0.849
H = 0.83  DF = 3  P = 0.843  (adjusted for ties)
```

Kruskal-Wallis Test: Design Changes versus PDS

```
34 cases were used
1 cases contained missing values

Kruskal-Wallis Test on ChangeDes

PDS       N       Median  Ave Rank      Z
0         7  0.200000000      19.9   0.72
1        12  0.150000000      20.3   1.19
2         5  0.000000000      11.7  -1.41
3        10  0.043500000      15.4  -0.79
Overall  34                   17.5

H = 3.47  DF = 3  P = 0.324
H = 3.62  DF = 3  P = 0.306  (adjusted for ties)
```

Kruskal-Wallis Test: Regulatory Changes versus PDS

```
32 cases were used
3 cases contained missing values

Kruskal-Wallis Test on ChangeReg

PDS       N       Median  Ave Rank      Z
0         7  0.010000000      17.2   0.23
1        12  0.050000000      21.0   2.08
2         3  0.000000000      14.2  -0.45
3        10  0.000000000      11.4  -2.09
Overall  32                   16.5

H = 5.95  DF = 3  P = 0.114
H = 6.65  DF = 3  P = 0.084  (adjusted for ties)
```

Kruskal-Wallis Test: CO Processing Time versus PDS

```
Kruskal-Wallis Test on COTime

PDS       N  Median  Ave Rank      Z
0         7   7.000      26.9   2.58
1        13   4.000      18.2   0.09
2         5   5.000      25.6   1.79
3        10   1.000       7.7  -3.76
Overall  35              18.0

H = 18.17  DF = 3  P = 0.000
H = 18.61  DF = 3  P = 0.000  (adjusted for ties)
```

Kruskal-Wallis Test: Rework versus PDS

```
34 cases were used
1 cases contained missing values

Kruskal-Wallis Test on Rework

PDS       N  Median  Ave Rank      Z
0         6   1.000      17.7   0.05
1        13   1.000      16.4  -0.51
2         5   1.000      15.0  -0.61
3        10   1.000      20.1   0.98
Overall  34              17.5

H = 1.16  DF = 3  P = 0.762
H = 2.34  DF = 3  P = 0.505  (adjusted for ties)
```

Kruskal-Wallis Test: Resubmittals versus PDS

```
23 cases were used
12 cases contained missing values

Kruskal-Wallis Test on Resubmittals

PDS       N  Median  Ave Rank      Z
0         4  5.2386      17.3   1.70
1         9  1.0383      13.6   0.88
2         3  0.4083      10.0  -0.55
3         7  0.2427       7.9  -1.94
Overall  23              12.0

H = 5.74  DF = 3  P = 0.125
```

Kruskal-Wallis Test: Deficiency Issues versus PDS

```
25 cases were used
10 cases contained missing values

Kruskal-Wallis Test on Deficiencies

PDS       N  Median  Ave Rank      Z
0         5  1.9604      18.0   1.70
1         7  1.3889      17.1   1.76
2         3  0.3266       9.8  -0.79
3        10  0.2071       8.6  -2.47
Overall  25              13.0

H = 8.74  DF = 3  P = 0.033
H = 8.77  DF = 3  P = 0.033  (adjusted for ties)
```

Kruskal-Wallis Test: Claims versus PDS

```
34 cases were used
1 cases contained missing values

Kruskal-Wallis Test on Claims

PDS       N       Median  Ave Rank      Z
```

```
0         6  0.000000000      18.8   0.36
1        13  0.000000000      18.6   0.51
2         5  0.000000000      16.0  -0.36
3        10  0.000000000      16.0  -0.57
Overall  34                   17.5

H = 0.61  DF = 3  P = 0.894
H = 2.53  DF = 3  P = 0.470  (adjusted for ties)
```

Kruskal-Wallis Test: Recycled versus PDS

```
28 cases were used
7 cases contained missing values

Kruskal-Wallis Test on Recycled

PDS       N  Median  Ave Rank      Z
0         7  0.7000      11.1  -1.27
1        11  0.8700      16.6   1.11
2         3  0.8000      12.5  -0.45
3         7  0.8700      15.4   0.34
Overall  28              14.5

H = 2.22  DF = 3  P = 0.527
H = 2.23  DF = 3  P = 0.526  (adjusted for ties)
```

Kruskal-Wallis Test: OH&Profit versus PDS

```
33 cases were used
2 cases contained missing values

Kruskal-Wallis Test on OHProfit

PDS       N  Median  Ave Rank      Z
0         6   1.000       8.9  -2.26
1        12   1.000      16.8  -0.11
2         5   1.000      16.5  -0.13
3        10   2.000      22.4   2.12
Overall  33              17.0

H = 7.33  DF = 3  P = 0.062
H = 9.28  DF = 3  P = 0.026  (adjusted for ties)
```

Kruskal-Wallis Test: Return Business versus PDS

```
Kruskal-Wallis Test on ReturnBusiness

PDS       N  Median  Ave Rank      Z
0         7   2.000      13.4  -1.32
1        13   2.000      19.0   0.43
2         5   2.000      18.2   0.05
3        10   2.000      19.9   0.68
Overall  35              18.0

H = 1.84  DF = 3  P = 0.607
H = 3.78  DF = 3  P = 0.286  (adjusted for ties)
```

Kruskal-Wallis Test: Punchlist Items versus PDS

```
32 cases were used
3 cases contained missing values

Kruskal-Wallis Test on PunchlistItems

PDS       N  Median  Ave Rank      Z
0         7  41.265      23.6   2.26
1        12  22.977      17.9   0.66
2         5   9.786      15.2  -0.34
3         8   7.228       9.0  -2.61
Overall  32              16.5

H = 9.46  DF = 3  P = 0.024
H = 9.46  DF = 3  P = 0.024  (adjusted for ties)
```

Kruskal-Wallis Test: Punchlist Cost versus PDS

```
33 cases were used
2 cases contained missing values

Kruskal-Wallis Test on PunchlistCost

PDS       N  Median  Ave Rank      Z
0         7   2.000      18.8   0.55
1        12   2.000      20.0   1.37
2         5   1.000      16.3  -0.18
3         9   1.000      11.9  -1.84
Overall  33              17.0

H = 3.91  DF = 3  P = 0.271
H = 4.88  DF = 3  P = 0.181  (adjusted for ties)
```

Kruskal-Wallis Test: Warranty Cost versus PDS

```
29 cases were used
6 cases contained missing values

Kruskal-Wallis Test on WarrantyCost

PDS       N       Median  Ave Rank      Z
0         4  1.000000000      18.0   0.76
1        11  1.000000000      16.5   0.72
2         5  1.000000000      18.0   0.87
3         9  0.000000000      10.2  -2.03
Overall  29                   15.0

H = 4.27  DF = 3  P = 0.234
H = 7.04  DF = 3  P = 0.071  (adjusted for ties)

* NOTE * One or more small samples
```

Kruskal-Wallis Test: Latent Defect Cost versus PDS

```
28 cases were used
7 cases contained missing values
```

```
Kruskal-Wallis Test on LatentDefect

PDS       N       Median  Ave Rank       Z
0         4  0.500000000      14.0   -0.13
1        10  1.000000000      15.4    0.43
2         5  1.000000000      15.4    0.27
3         9  0.000000000      13.2   -0.57
Overall  28                   14.5

H = 0.41  DF = 3  P = 0.938
H = 0.55  DF = 3  P = 0.908  (adjusted for ties)
```

Selected Kruskal-Wallis and ANOVA Results for PQR

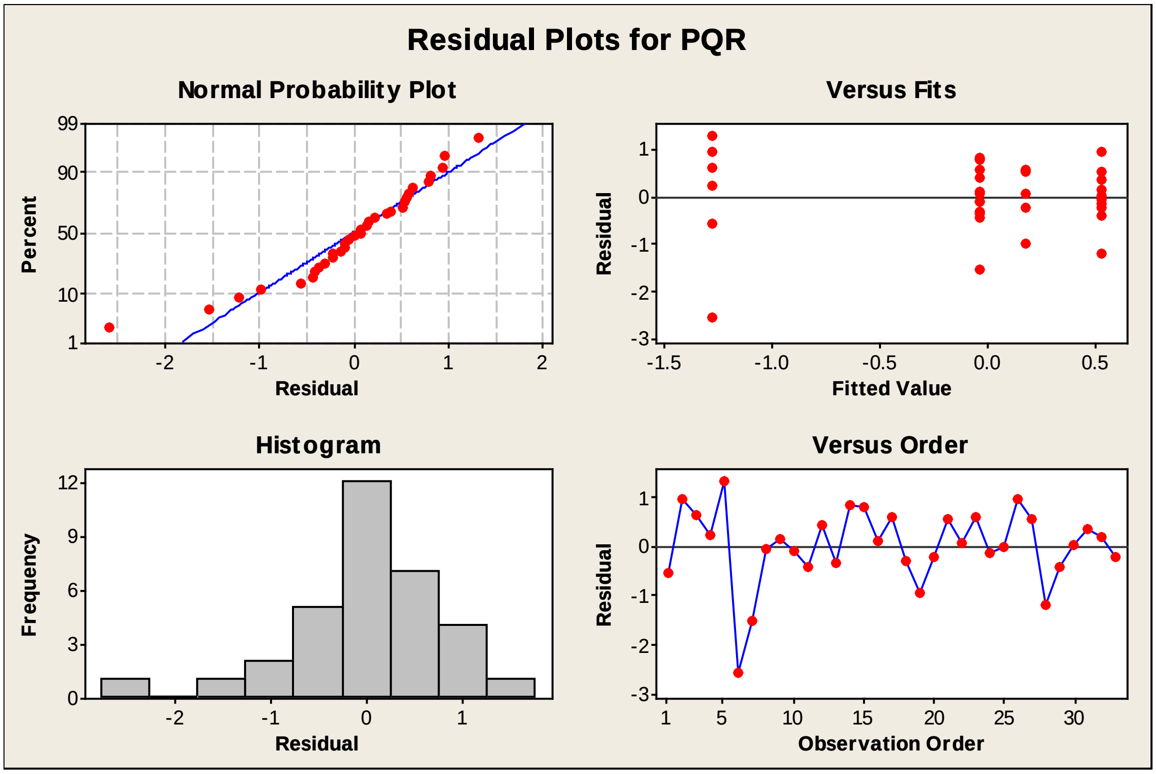
```
Welcome to Minitab, press F1 for help.
Retrieving project from file: 'I:\Win\Desktop\KruskalWallis.MPJ'
```

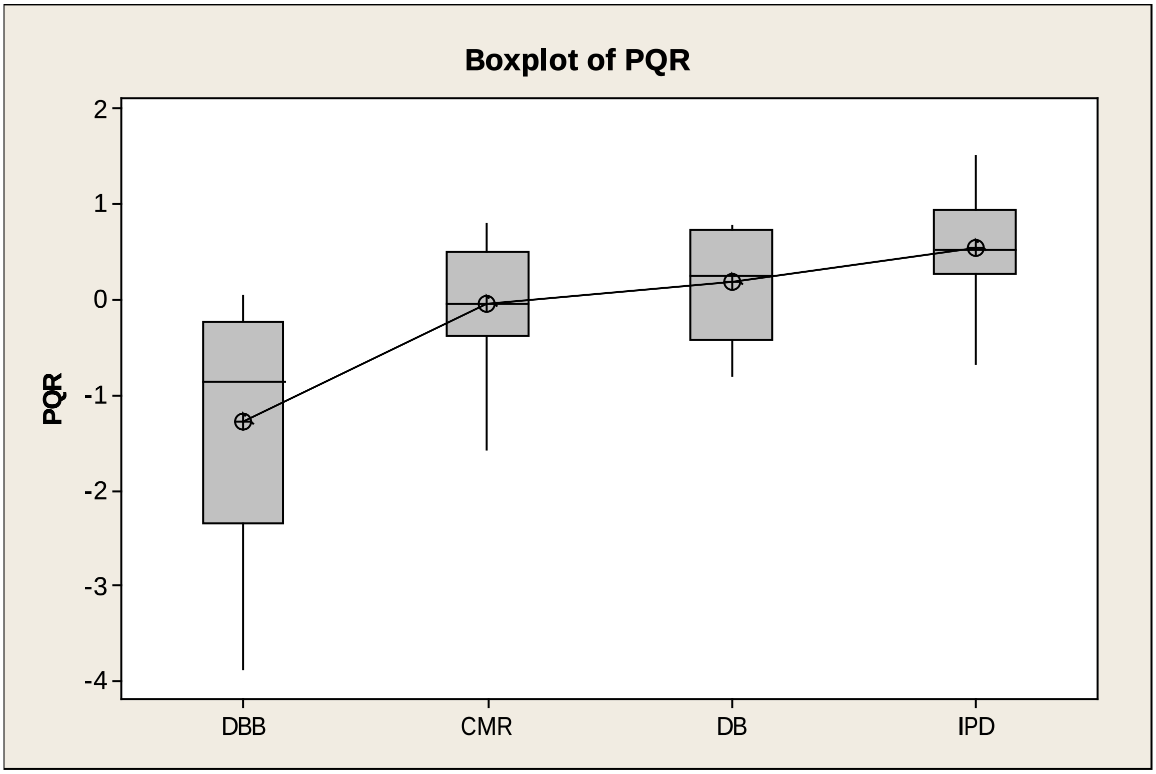
Kruskal-Wallis Test: PQR versus PDS

```
Kruskal-Wallis Test on PQR

PDS--2    N     Median  Ave Rank      Z
0         6  -0.85589       6.7  -2.89
1        12  -0.03358      15.9  -0.49
2         5   0.25064      18.8   0.45
3        10   0.53005      23.6   2.59
Overall  33                17.0

H = 11.83  DF = 3  P = 0.008
```

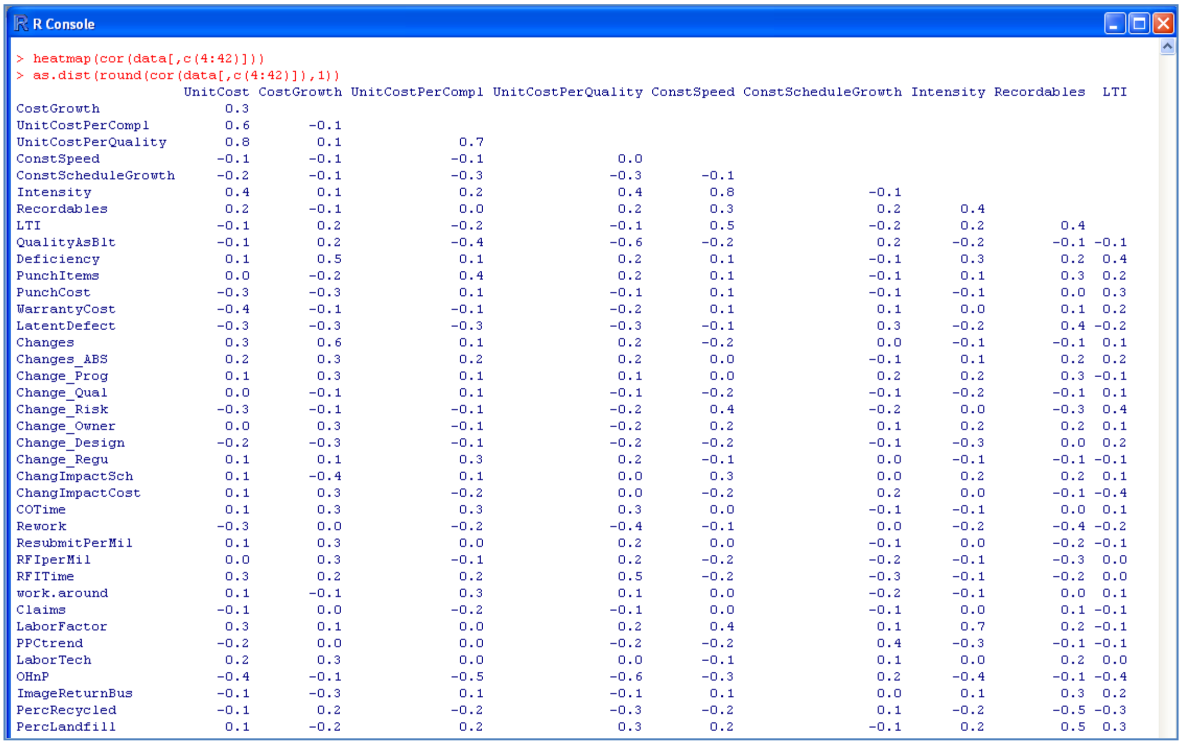
One-way ANOVA: PQR versus PDS

```
Source  DF      SS     MS     F      P
PDS--2   3  12.817  4.272  6.41  0.002
Error   29  19.340  0.667
Total   32  32.157

S = 0.8166   R-Sq = 39.86%   R-Sq(adj) = 33.64%


                            Individual 95% CIs For Mean Based on
                            Pooled StDev
Level    N     Mean   StDev  -----+---------+---------+---------+----
0        6  -1.2819  1.4159  (--------*-------)
1       12  -0.0366  0.6467                   (-----*-----)
2        5   0.1786  0.6391                  (--------*---------)
3       10   0.5373  0.5852                         (------*-----)
                             -----+---------+---------+---------+----
                               -1.60     -0.80     -0.00      0.80

Pooled StDev = 0.8166
```

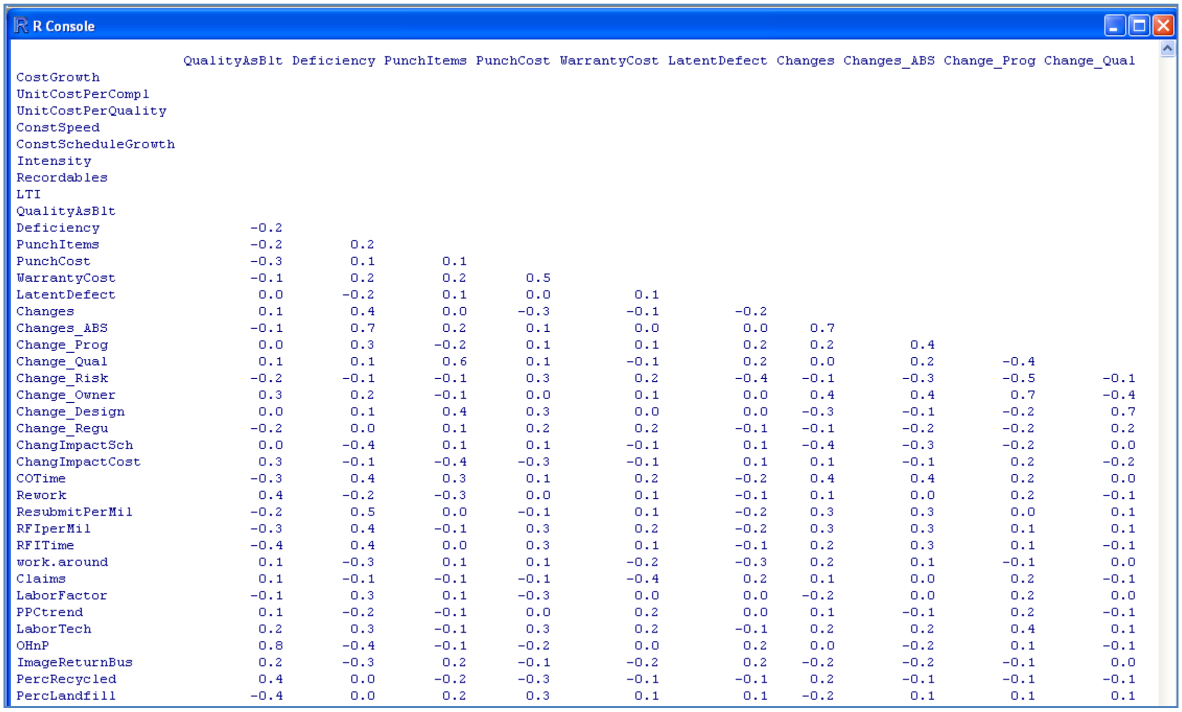

Residual Plots for PQR
Normal Probability Plot
Percent
99
90
50
10
1
-2
-1
0
1
2
Residual
Versus Fits
Residual
1
0
-1
-2
-3
-1.5
-1.0
-0.5
0.0
0.5
Fitted Value
Histogram
Frequency
12
9
6
3
0
-2
-1
0
1
Residual
Versus Order
Residual
1
0
-1
-2
-3
1
5
10
15
20
25
30
Observation Order


Boxplot of PQR
PQR
2
1
0
-1
-2
-3
-4
DBB
CMR
DB
IPD

Appendix F – PERFORMANCE METRICS CORRELATIONS

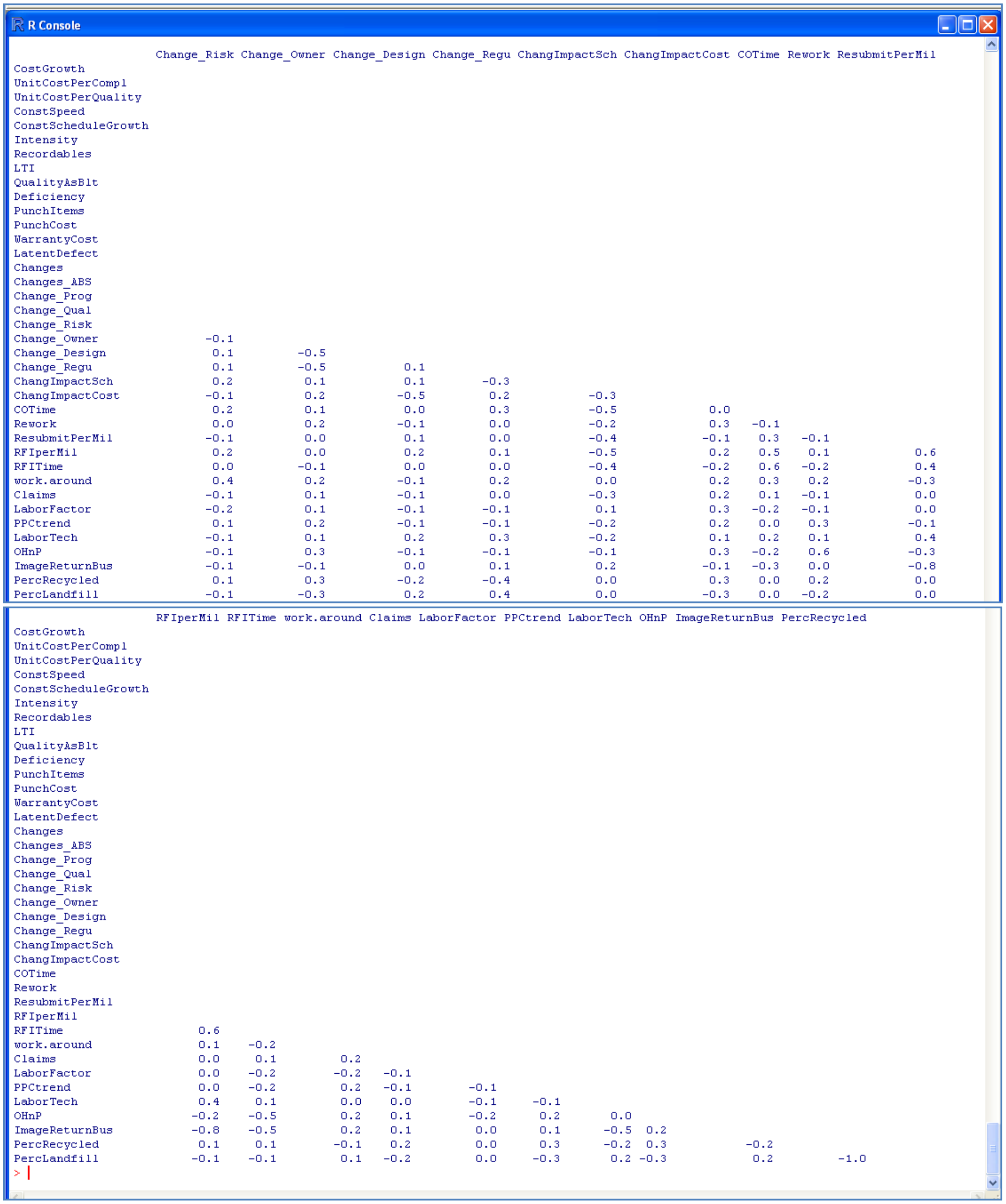


```
> heatmap(cor(data[,c(4:42)]))
> as.dist(round(cor(data[,c(4:42)]),1))
                    UnitCost CostGrowth UnitCostPerCompl UnitCostPerQuality ConstSpeed ConstScheduleGrowth Intensity Recordables  LTI
CostGrowth               0.3
UnitCostPerCompl         0.6       -0.1
UnitCostPerQuality       0.8        0.1              0.7
ConstSpeed              -0.1       -0.1             -0.1                0.0
ConstScheduleGrowth     -0.2       -0.1             -0.3               -0.3       -0.1
Intensity                0.4        0.1              0.2                0.4        0.8                -0.1
Recordables              0.2       -0.1              0.0                0.2        0.3                 0.2       0.4
LTI                     -0.1        0.2             -0.2               -0.1        0.5                -0.2       0.2         0.4
QualityAsBlt            -0.1        0.2             -0.4               -0.6       -0.2                 0.2      -0.2        -0.1 -0.1
Deficiency               0.1        0.5              0.1                0.2        0.1                -0.1       0.3         0.2  0.4
PunchItems               0.0       -0.2              0.4                0.2        0.1                -0.1       0.1         0.3  0.2
PunchCost               -0.3       -0.3              0.1               -0.1        0.1                -0.1      -0.1         0.0  0.3
WarrantyCost            -0.4       -0.1             -0.1               -0.2        0.1                 0.1       0.0         0.1  0.2
LatentDefect            -0.3       -0.3             -0.3               -0.3       -0.1                 0.3      -0.2         0.4 -0.2
Changes                  0.3        0.6              0.1                0.2       -0.2                 0.0      -0.1        -0.1  0.1
Changes_ABS              0.2        0.3              0.2                0.2        0.0                -0.1       0.1         0.2  0.2
Change_Prog              0.1        0.3              0.1                0.1        0.0                 0.2       0.2         0.3 -0.1
Change_Qual              0.0       -0.1              0.1               -0.1       -0.2                -0.1      -0.2        -0.1  0.1
Change_Risk             -0.3       -0.1             -0.1               -0.2        0.4                -0.2       0.0        -0.3  0.4
Change_Owner             0.0        0.3             -0.1               -0.2        0.2                 0.1       0.2         0.2  0.1
Change_Design           -0.2       -0.3             -0.1               -0.2       -0.2                -0.1      -0.3         0.0  0.2
Change_Regu              0.1        0.1              0.3                0.2       -0.1                 0.0      -0.1        -0.1 -0.1
ChangImpactSch           0.1       -0.4              0.1                0.0        0.3                 0.0       0.2         0.2  0.1
ChangImpactCost          0.1        0.3             -0.2                0.0       -0.2                 0.2       0.0        -0.1 -0.4
COTime                   0.1        0.3              0.3                0.3        0.0                -0.1      -0.1         0.0  0.1
Rework                  -0.3        0.0             -0.2               -0.4       -0.1                 0.0      -0.2        -0.4 -0.2
ResubmitPerMil           0.1        0.3              0.0                0.2        0.0                -0.1       0.0        -0.2 -0.1
RFIperMil                0.0        0.3             -0.1                0.2       -0.2                -0.2      -0.1        -0.3  0.0
RFITime                  0.3        0.2              0.2                0.5       -0.2                -0.3      -0.1        -0.2  0.0
work.around              0.1       -0.1              0.3                0.1        0.0                -0.2      -0.1         0.0  0.1
Claims                  -0.1        0.0             -0.2               -0.1        0.0                -0.1       0.0         0.1 -0.1
LaborFactor              0.3        0.1              0.0                0.2        0.4                 0.1       0.7         0.2 -0.1
PPCtrend                -0.2        0.0              0.0               -0.2       -0.2                 0.4      -0.3        -0.1 -0.1
LaborTech                0.2        0.3              0.0                0.0       -0.1                 0.1       0.0         0.2  0.0
OHnP                    -0.4       -0.1             -0.5               -0.6       -0.3                 0.2      -0.4        -0.1 -0.4
ImageReturnBus          -0.1       -0.3              0.1               -0.1        0.1                 0.0       0.1         0.3  0.2
PercRecycled            -0.1        0.2             -0.2               -0.3       -0.2                 0.1      -0.2        -0.5 -0.3
PercLandfill             0.1       -0.2              0.2                0.3        0.2                -0.1       0.2         0.5  0.3
```



```
                    QualityAsBlt Deficiency PunchItems PunchCost WarrantyCost LatentDefect Changes Changes_ABS Change_Prog Change_Qual
CostGrowth
UnitCostPerCompl
UnitCostPerQuality
ConstSpeed
ConstScheduleGrowth
Intensity
Recordables
LTI
QualityAsBlt
Deficiency                  -0.2
PunchItems                  -0.2        0.2
PunchCost                   -0.3        0.1        0.1
WarrantyCost                -0.1        0.2        0.2       0.5
LatentDefect                 0.0       -0.2        0.1       0.0          0.1
Changes                      0.1        0.4        0.0      -0.3         -0.1         -0.2
Changes_ABS                 -0.1        0.7        0.2       0.1          0.0          0.0     0.7
Change_Prog                  0.0        0.3       -0.2       0.1          0.1          0.2     0.2         0.4
Change_Qual                  0.1        0.1        0.6       0.1         -0.1          0.2     0.0         0.2        -0.4
Change_Risk                 -0.2       -0.1       -0.1       0.3          0.2         -0.4    -0.1        -0.3        -0.5        -0.1
Change_Owner                 0.3        0.2       -0.1       0.0          0.1          0.0     0.4         0.4         0.7        -0.4
Change_Design                0.0        0.1        0.4       0.3          0.0          0.0    -0.3        -0.1        -0.2         0.7
Change_Regu                 -0.2        0.0        0.1       0.2          0.2         -0.1    -0.1        -0.2        -0.2         0.2
ChangImpactSch               0.0       -0.4        0.1       0.1         -0.1          0.1    -0.4        -0.3        -0.2         0.0
ChangImpactCost              0.3       -0.1       -0.4      -0.3         -0.1          0.1     0.1        -0.1         0.2        -0.2
COTime                      -0.3        0.4        0.3       0.1          0.2         -0.2     0.4         0.4         0.2         0.0
Rework                       0.4       -0.2       -0.3       0.0          0.1         -0.1     0.1         0.0         0.2        -0.1
ResubmitPerMil              -0.2        0.5        0.0      -0.1          0.1         -0.2     0.3         0.3         0.0         0.1
RFIperMil                   -0.3        0.4       -0.1       0.3          0.2         -0.2     0.3         0.3         0.1         0.1
RFITime                     -0.4        0.4        0.0       0.3          0.1         -0.1     0.2         0.3         0.1        -0.1
work.around                  0.1       -0.3        0.1       0.1         -0.2         -0.3     0.2         0.1        -0.1         0.0
Claims                       0.1       -0.1       -0.1      -0.1         -0.4          0.2     0.1         0.0         0.2        -0.1
LaborFactor                 -0.1        0.3        0.1      -0.3          0.0          0.0    -0.2         0.0         0.2         0.0
PPCtrend                     0.1       -0.2       -0.1       0.0          0.2          0.0     0.1        -0.1         0.2        -0.1
LaborTech                    0.2        0.3       -0.1       0.3          0.2         -0.1     0.2         0.2         0.4         0.1
OHnP                         0.8       -0.4       -0.1      -0.2          0.0          0.2     0.0        -0.2         0.1        -0.1
ImageReturnBus               0.2       -0.3        0.2      -0.1         -0.2          0.2    -0.2        -0.2        -0.1         0.0
PercRecycled                 0.4        0.0       -0.2      -0.3         -0.1         -0.1     0.2        -0.1        -0.1        -0.1
PercLandfill                -0.4        0.0        0.2       0.3          0.1          0.1    -0.2         0.1         0.1         0.1
```

FIGURE 75: Performance Metrics Correlations Part 1

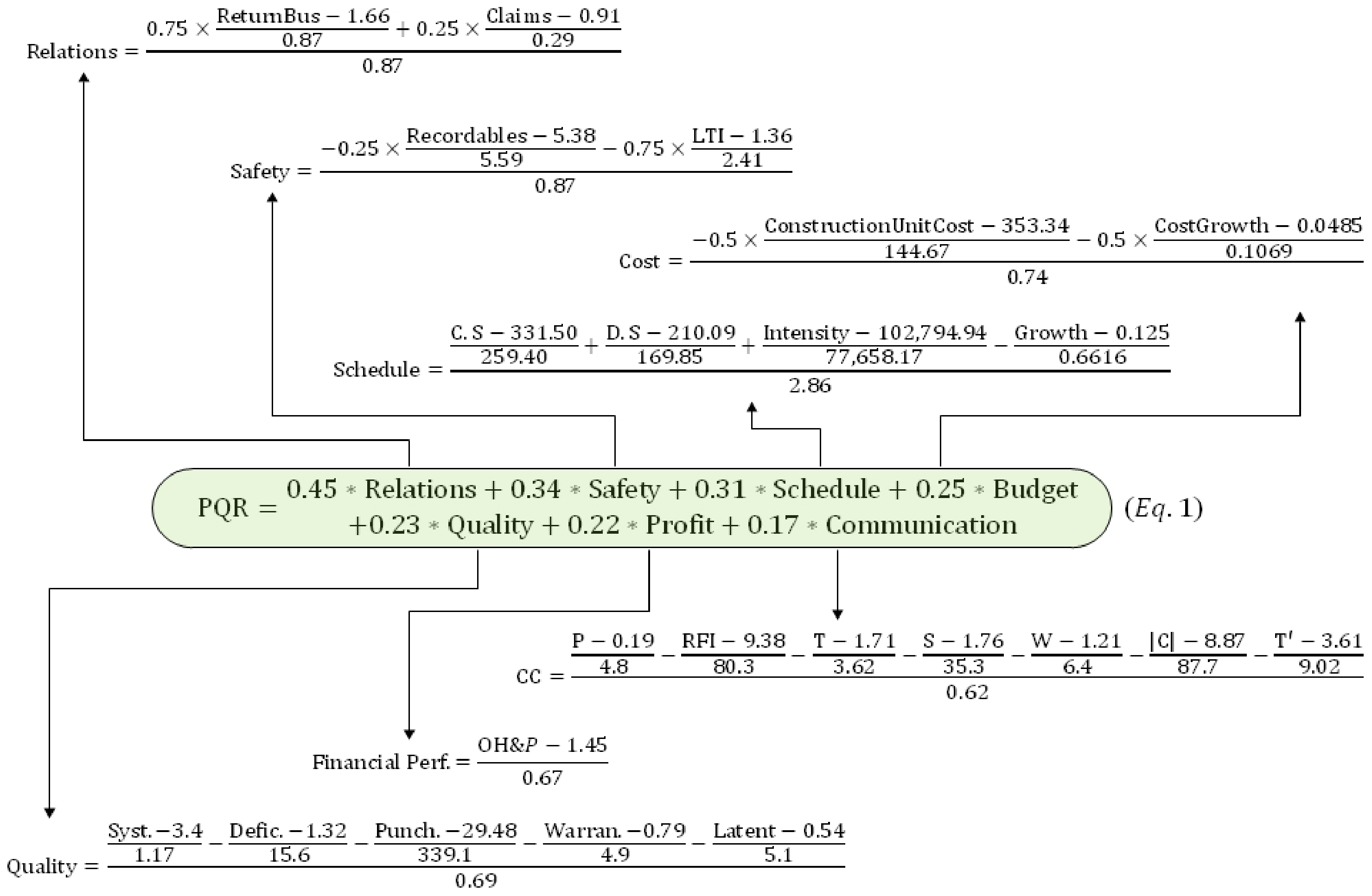


```
                    Change_Risk Change_Owner Change_Design Change_Regu ChangImpactSch ChangImpactCost COTime Rework ResubmitPerMil
CostGrowth
UnitCostPerCompl
UnitCostPerQuality
ConstSpeed
ConstScheduleGrowth
Intensity
Recordables
LTI
QualityAsBlt
Deficiency
PunchItems
PunchCost
WarrantyCost
LatentDefect
Changes
Changes_ABS
Change_Prog
Change_Qual
Change_Risk
Change_Owner               -0.1
Change_Design               0.1         -0.5
Change_Regu                 0.1         -0.5           0.1
ChangImpactSch              0.2          0.1           0.1        -0.3
ChangImpactCost            -0.1          0.2          -0.5         0.2           -0.3
COTime                      0.2          0.1           0.0         0.3           -0.5             0.0
Rework                      0.0          0.2          -0.1         0.0           -0.2             0.3   -0.1
ResubmitPerMil             -0.1          0.0           0.1         0.0           -0.4            -0.1    0.3   -0.1
RFIperMil                   0.2          0.0           0.2         0.1           -0.5             0.2    0.5    0.1            0.6
RFITime                     0.0         -0.1           0.0         0.0           -0.4            -0.2    0.6   -0.2            0.4
work.around                 0.4          0.2          -0.1         0.2            0.0             0.2    0.3    0.2           -0.3
Claims                     -0.1          0.1          -0.1         0.0           -0.3             0.2    0.1   -0.1            0.0
LaborFactor                -0.2          0.1          -0.1        -0.1            0.1             0.3   -0.2   -0.1            0.0
PPCtrend                    0.1          0.2          -0.1        -0.1           -0.2             0.2    0.0    0.3           -0.1
LaborTech                  -0.1          0.1           0.2         0.3           -0.2             0.1    0.2    0.1            0.4
OHnP                       -0.1          0.3          -0.1        -0.1           -0.1             0.3   -0.2    0.6           -0.3
ImageReturnBus             -0.1         -0.1           0.0         0.1            0.2            -0.1   -0.3    0.0           -0.8
PercRecycled                0.1          0.3          -0.2        -0.4            0.0             0.3    0.0    0.2            0.0
PercLandfill               -0.1         -0.3           0.2         0.4            0.0            -0.3    0.0   -0.2            0.0
```

```
                    RFIperMil RFITime work.around Claims LaborFactor PPCtrend LaborTech OHnP ImageReturnBus PercRecycled
CostGrowth
UnitCostPerCompl
UnitCostPerQuality
ConstSpeed
ConstScheduleGrowth
Intensity
Recordables
LTI
QualityAsBlt
Deficiency
PunchItems
PunchCost
WarrantyCost
LatentDefect
Changes
Changes_ABS
Change_Prog
Change_Qual
Change_Risk
Change_Owner
Change_Design
Change_Regu
ChangImpactSch
ChangImpactCost
COTime
Rework
ResubmitPerMil
RFIperMil
RFITime                   0.6
work.around               0.1    -0.2
Claims                    0.0     0.1         0.2
LaborFactor               0.0    -0.2        -0.2   -0.1
PPCtrend                  0.0    -0.2         0.2   -0.1        -0.1
LaborTech                 0.4     0.1         0.0    0.0        -0.1     -0.1
OHnP                     -0.2    -0.5         0.2    0.1        -0.2      0.2       0.0
ImageReturnBus           -0.8    -0.5         0.2    0.1         0.0      0.1      -0.5  0.2
PercRecycled              0.1     0.1        -0.1    0.2         0.0      0.3      -0.2  0.3           -0.2
PercLandfill             -0.1    -0.1         0.1   -0.2         0.0     -0.3       0.2 -0.3            0.2         -1.0
> 
```

FIGURE 76: Performance Metrics Correlations Part 2

Appendix G – PQR ***APPLICATION***



$$\text{Relations} = \frac{0.75 \times \frac{\text{ReturnBus} - 1.66}{0.87} + 0.25 \times \frac{\text{Claims} - 0.91}{0.29}}{0.87}$$

$$\text{Safety} = \frac{-0.25 \times \frac{\text{Recordables} - 5.38}{5.59} - 0.75 \times \frac{\text{LTI} - 1.36}{2.41}}{0.87}$$

$$\text{Cost} = \frac{-0.5 \times \frac{\text{ConstructionUnitCost} - 353.34}{144.67} - 0.5 \times \frac{\text{CostGrowth} - 0.0485}{0.1069}}{0.74}$$

$$\text{Schedule} = \frac{\frac{\text{C}.S - 331.50}{259.40} + \frac{\text{D}.S - 210.09}{169.85} + \frac{\text{Intensity} - 102{,}794.94}{77{,}658.17} - \frac{\text{Growth} - 0.125}{0.6616}}{2.86}$$

$$\text{PQR} = \begin{matrix} 0.45 * \text{Relations} + 0.34 * \text{Safety} + 0.31 * \text{Schedule} + 0.25 * \text{Budget} \\ +0.23 * \text{Quality} + 0.22 * \text{Profit} + 0.17 * \text{Communication} \end{matrix} \quad (Eq.\ 1)$$

$$\text{CC} = \frac{\frac{\text{P} - 0.19}{4.8} - \frac{\text{RFI} - 9.38}{80.3} - \frac{\text{T} - 1.71}{3.62} - \frac{S - 1.76}{35.3} - \frac{\text{W} - 1.21}{6.4} - \frac{|\text{C}| - 8.87}{87.7} - \frac{\text{T}' - 3.61}{9.02}}{0.62}$$

$$\text{Financial Perf.} = \frac{\text{OH\&}P - 1.45}{0.67}$$

$$\text{Quality} = \frac{\frac{\text{Syst.} - 3.4}{1.17} - \frac{\text{Defic.} - 1.32}{15.6} - \frac{\text{Punch.} - 29.48}{339.1} - \frac{\text{Warran.} - 0.79}{4.9} - \frac{\text{Latent} - 0.54}{5.1}}{0.69}$$

FIGURE 77: Summary of PQR Formulas

TABLE 36: PQR Numerical Example

Performance Metrics	Initial Inputs	Performance Area Scores	PQR
Return Business	2	**0.4281**	0.6851
Claims	1		
Recordables per $100million	11.961	**0.1492**	
LTI per $100 million	0		
Construction Speed	399.871	**-1.5313**	
Delivery Speed	259.785		
Intensity	90837.670		
Construction Schedule Growth	3.293		
Unit Cost	245.508	**0.9671**	
Cost Growth	-0.024		
Systems Quality	3.938	**0.6533**	
Deficiencies per $100 million	0.299		
Punchlist Items per $100 million	10.466		
Warranty Cost	1		
Latent Defect Cost	1		
OH & Profit	2	**0.8193**	
RFI per $100 million	0	**1.8769**	
RFI Processing Time	0		
Rework	1		
Resubmittals per $100 million	0.090		
Percent Change (absolute value)	0.050		
Change Order Processing Time	1		
PPC Trend	1		