THE DEVELOPMENT OF MATHEMATICAL MODELS FOR PRELIMINARY PREDICTION OF HIGHWAY CONSTRUCTION DURATION

Robert C. Williams

Dissertation submitted to the faculty of

Virginia Polytechnic Institute and State University

in partial fulfillment of the requirements for degree of

Doctor of Philosophy

in

Civil Engineering

Michael C. Vorster, Co-chair
John C. Hildreth, Co-chair
Jesús M. de la Garza
Ray D. Pethtel
Anthony D. Songer

November 4th, 2008

Blacksburg, VA

Keywords: Highway Construction Duration, Multiple Linear Regression Analysis, Duration-Influential Project Factors

The Development of Mathematical Models for Preliminary Prediction of Highway Construction Duration

Robert C. Williams

Abstract

Knowledge of construction duration is pertinent to a number of project planning functions prior to detailed design development. Funding, financing, and resource allocation decisions take place early in project design development and are significantly influenced by the construction duration. Currently, there is not an understanding of the project factors having a statistically significant relationship with highway construction duration. Other industry sectors have successfully used statistical regression analysis to identify and model the project parameters related to construction duration. While the need is seen for such work in highway construction, there are very few studies which attempt to identify duration-influential parameters and their relationship with the highway construction duration.

This research identifies the project factors, known early in design development, which influence highway construction duration. The factors identified are specific to their respective project types and are those factors which demonstrate a statistically-significant relationship with construction duration. This work also quantifies the relationship between the duration-influential factors and highway construction duration. The quantity, magnitude, and sign of the factor coefficient yields evidence regarding the importance of the project factor to highway construction duration. Finally, the research incorporates the duration-influential project factors and their relationship with highway construction duration into mathematical models which assist in the prediction of construction duration. Full and condensed models are presented for Full-Depth Section and Highway Improvement project types. This research uses statistical regression analysis to identify, quantify, and model these early-known, duration-influential project factors.

The results of this research contribute to the body of knowledge of the sponsoring organization (Virginia Department of Transportation), the highway construction industry, and the general construction industry at large.

Dedication

For Jill,

Who had faith, when I did not.

Author’s Acknowledgements

I would like to offer my sincerest thanks to my dissertation co-chairs, Dr. Mike Vorster and Dr. John Hildreth, for their guidance, support, and assistance throughout the PhD process. They have been exceptional role models and I know my work and my future have been positively influenced by them both. Throughout the process, I’ve learned many lessons that go beyond the field of construction and I am ever-grateful for them allowing me to work with them over the last several years. In my research, Dr. Vorster has always kept me focused on the contributions of my work. In my teaching, Dr. Vorster showed me how to not ‘teach’ but to ‘cause to learn’. He impressed upon me the value of treating students as adults and professionals and expecting the same level of professionalism in return. Dr. Hildreth has grown to be as much a friend as an advisor and has taught me countless lessons about responsibility, hard work, and life-long learning. We’ve shared many laughs over the four years I’ve known him and I know there will be many more in the future.

I would also like to recognize my committee members Dr. Jesús de la Garza, Mr. Ray Pethtel, and Dr. Anthony Songer personally for their feedback, guidance, and suggestions throughout the PhD process. Dr. de la Garza is a renowned researcher and has always pushed me to maintain a scholarly approach to the work and focus on the importance of sound methodologies. Mr. Pethtel has lent his years of experience as a Commissioner of VDOT and in the highway construction industry to improve the applicability aspects of my research, keeping me attuned to the needs of the highway construction industry. As a committee member, Dr. Songer showed me the value of establishing, following, and presenting a formal research methodology. As an instructor, Dr. Songer’s Construction Research Presentation course provided a sound basis for conducting my research and presenting my results.

Virginia Department of Transportation (VDOT) and its many employees deserve acknowledgement as well. VDOT is filled with motivated people who have continually supported my work and the work of the former *VDOT – VT Partnership for Project Scheduling*. VDOT is truly a leader among highway agencies in their dedication to learning, advancement, and construction excellence.

Many friends and acquaintances have been made here at Virginia Tech and in Blacksburg. The students, faculty, and staff of the Civil Engineering Department and Vecellio Construction Engineering and Management Program are some of the greatest I've known and have made my time at Virginia Tech a pleasure. Also, the pastors and parishioners of Blacksburg United Methodist Church have welcomed me and my family and made us feel right at home over the past few years. Their friendship and love will be greatly missed.

My family is credited with keeping me grounded and inspired. Coming from the coalfields of southern West Virginia, I learned early-on the value of an education and the doors that are unlocked with the key of knowledge. Their love and support has never altered and I am eternally grateful for them.

Finally, my wife, Jill deserves the greatest acknowledgement. She constantly encouraged me to pursue my dreams, even when it meant personal sacrifice throughout the first years of our marriage. Jill believed in me, even when I did not. She showed me how to have faith in all other aspects of my life as well. And now, she has blessed me with a baby boy - Jacob. This is why I have dedicated this work to her. I hope that I can one day repay her for such love and support.

Table of Contents

Table of Figures

List of Tables

List of Equations

Chapter 1 – The Development of Mathematical Models for Preliminary Prediction of Highway Construction Duration

1.1 Introduction and Background

Construction duration is the number of days needed to perform the work required in the contract. This amount of time considers only the number of days actually worked and does not consider non-working days such as weekends, holidays, or weather-impacted days. Construction duration commences upon the notice to proceed and is concluded once all contract work items are complete. This duration estimate is generally presented in working days.

Contract duration, on the other hand, is the time allotted to the contractor, by the owner, for the performance of all duties within a contract. This duration permits the contractor to perform his work as he sees fit, even finishing early, while guarding the owner against damages suffered due to late completion [Fourie 2003]. The contract duration estimate is created by applying a calendar (considering expected impacts such as weekends and adverse weather) to the construction duration estimate. This yields duration in calendar days.

Because construction duration is directly related to contract time, the importance of contract time is shared by construction duration. Accurate and realistic duration estimates are important to every aspect of the construction project. Unreasonably short contract durations raise the bid price, restrict qualified bidders from submitting bids, have potential to reduce the quality of the work, and increase the potential for legal disputes [FHWA 2002]. Conversely, unreasonably long contract durations are a general inconvenience to the traveling public and encourage less qualified contractors to submit a bid[FHWA 2002].

Numerous authors have repeated the importance for accurate and reliable duration estimates over the past four decades. Bromilow [1969] discovered the need for more reliable methods in the Australian construction industry. More recently, Hancher and Rowings [1981] emphasized this importance within the United States transportation agency and the need for systematic methods to estimate contract times. Walker [1995] put forth a regression model to assist the United Kingdom construction industry in benchmarking its own performance against that of the rest of the world. Finally, Chan and Kumaraswamy [1996] evaluated construction duration performance in Hong Kong and determined that a more reliable, systematic method (such as a mathematical model) for estimating duration was needed.

As a pre-advertisement process, construction and contract duration estimation is performed by the construction owner or engineer. The construction owner is the "instigating party that gets the project financed, designed, and built" [Clough et al. 2000]. These entities may be public or private. Public owners include federal, state, county, and other local entities. Private entities are typically individuals, companies, or corporations but the options are almost infinite. Regardless of their status, their primary role in the construction process is to establish the scope of work [Schexnayder and Mayo 2004]. Public owners, however, have a greater responsibility to represent, provide for, and protect the best interests of their represented citizens.

One of the most visible public construction owners today is the state highway agency (SHA). These are the entities charged with constructing and maintaining highway systems throughout an entire state. Their role in the construction process goes far beyond setting the project scope and financing the project. SHAs usually design the project, plan the general sequence of work to take place, prepare detailed cost estimates, and estimate the time necessary to complete the work. The first such duration estimate is the construction duration estimate.

For state highway agencies, it is imperative that accurate and reasonable construction and contract durations be established. The Federal Highway Administration (FHWA) emphasizes the importance of establishing an accurate contract duration. Through 23CFR635.121, the FHWA requires individual states to develop and implement contract duration determination procedures for construction projects. The FHWA Guide for Construction Contract Time Determination Procedures (T5080.15) offers suggestions to assist states in the development and implementation of these procedures [FHWA 2002]. These suggestions include the development of a historical production rate database, a system for adjusting these rates for individual projects, the use of systematic scheduling methods such as CPM or bar charts, and the development of written contract time determination procedures that include suggestions for monitoring and documenting productivity [FHWA 2002].

1.2 Statement of the Problem

Despite the importance of accurately and reliably estimating construction duration, early in the project development lifecycle, few tools, practices, and procedures exist for doing so. The building construction industry has explored this area and found results in statistical regression analysis. Such analysis has lead to an understanding of the contract level components that influence construction duration and the relationships therein [Bromilow 1980, Nkado 1992, Chan and Kumaraswamy 1995, Chan 1999, Skitmore and Ng 2001, Burrows et al. 2005].

The highway construction industry has not demonstrated similar results. While most SHAs recognize a need for improved conceptual design level estimating practices, few have established processes, methods, or tools for preparing this estimate. Highway construction duration estimates are often prepared very early in project design and based on individual past experience on similar sized projects [Williams et al. 2008]. Such practice has shown to be inaccurate and biased [Farqahl and Everett 1997]. The alternative method employed by SHAs is using a more detailed scheduling tool and assuming pertinent project information such as material quantities and activity sequences [Williams et al. 2008]. Such an estimate is time consuming and based on numerous assumptions. Regardless of the method used, these estimates are often left unrefined until completed project design where a more accurate duration estimate can be developed.

Further, there is no consensus regarding the highway construction duration-influential parameters or factors. Authors have proposed such parameters and anecdotal evidence exists from experienced personnel. However, there has not been a study to determine the statistical significance of these factors. Additional research is also needed to quantify and model these relationships, the interactions between factors, and model the development of construction duration given this conceptual level project data.

1.3 Research Overview

This section gives a brief overview of the research undertaken. The intent here is to outline the overarching questions that the research answers. The expected benefits of the research and the applicability of these benefits to the construction community are addressed as well.

1.3.1 Research Questions

In performance of this research, three main questions are answered. These questions pertain to the research package as a whole. More detailed questions regarding individual elements of the research follow in the manuscripts. The broad research questions answered in fulfillment of the research are:

1. Can construction duration-influential factors be identified through analysis of historical conceptual level project data and statistical regression analysis?

2. What are the relationships between the duration-influential factors identified and highway construction duration?

3. How do these factors interact to describe and model the construction duration, given conceptual level project data?

1.3.2 Benefits of the Research

A number of benefits arise from answering these research questions. These benefits are most immediately recognizable to the highway construction owner. Generalized benefits will reach other highway construction entities and the construction industry at large.

Identification of construction duration-influential project factors has not been performed within the highway construction sector. Industry practitioners have a general understanding of these factors based on personal experience. These expectations have not, to date, been validated through a historically-based regression analysis methodology. Identification of highway construction duration-influential project factors will enhance current planning and budgeting practices. With knowledge of these factors, the practitioner may focus efforts on further developing and understanding project details for the factors known to be the most influential on construction duration.

Describing the relationships between these construction duration-influential factors and highway construction duration will provide an understanding of the influence or importance of various project factors on the project duration. Since these factors will, by definition, be known early in project development, understanding their relative impact will assist the practitioner in evaluating the consequences of defining, measuring, and altering the project scope, complexity, or size.

Finally, modeling the interactions between these factors and highway construction duration will allow the practitioner to incorporate this information into a useable and reliable process for predicting construction duration during the conceptual design phase. The primary benefit to the user is a historically-based, statistically-sound process and tool for predicting construction duration, early in project development, using few project factors. Such a process or tool will enhance the project constructability, financing, safety, and programming evaluation processes that take place early in project design.

1.3.3 Scope and Limitations

The analysis performed in this research is a retrospective study of Virginia Department of Transportation (VDOT) construction data. Of interest to this research are projects classified as Full-Depth Section and Highway Improvement project types. The projects studied are those started and completed between January 1, 2000 and January 1, 2007, respectively.

As a retrospective study, this research 'photographs' the highway construction industry. In other words, the projects analyzed are not scrutinized for potential unwarranted delays, special contracting strategies used, or any other circumstances that could have impacted the response (construction duration). Data outliers are only identified through statistical regression analysis techniques. As a photograph of the industry, the findings of this work are expected to be a useful benchmark of current practices and future improvements or alterations.

The contributions of this work to the body of knowledge are detailed in the final chapter. However, it bears mentioning here, that the results of this work are *specific* to

VDOT and *general* to the highway sector and construction industries. Particularly, the factors identified and the coefficient estimates quantified are *specific* to VDOT as they are developed using VDOT data. The duration-influential factors identified are useful to the entire highway construction industry, as it is expected that the duration-influential factors identified in Virginia, are similar to those in states across the United States. Finally, the methodology employed, especially the data collection, treatment, and regression analysis are pertinent to the construction industry at large.

1.4 Research Methodology

The methodology undertaken in this research is summarized in the following steps:

1. Establish the need for the research by understanding the current duration estimating practices and procedures employed by the highway construction industry,
2. Identify and select a construction duration information data source,
3. Collect and compile duration related data,
4. Seasonally adjust the contract duration collected to calculate a historically-typical construction duration,
5. Statistical regression analysis to identify duration-influential project factors and develop models for predicting construction duration
6. Validate the models developed,
7. Describe the proper use and bounds of the models developed, and
8. Discuss the applicability of the research methodology to the highway construction and general construction industries.

The figure below demonstrates how the methodology is documented through the four manuscripts included in this dissertation. These manuscripts have been submitted to engineering and construction research journals and are in various stages of review. The next section gives a brief overview of the dissertation document, including additional information on the manuscripts.

1. Establish the need for the research by understanding the current duration estimating practices and procedures employed by the highway construction industry,

Manuscript 1
Chapter 2

2. Identify and select a construction duration information data source,
3. Collect and compile duration related data,
4. Seasonally adjust the contract duration collected to calculate a historically-typical construction duration,

Manuscript 2
Chapter 3

5. Statistical regression analysis to identify duration-influential project factors and develop models for predicting construction duration
6. Validate the models developed,

Manuscript 3
Chapter 4

7. Describe the proper use and bounds of the models developed, and
8. Discuss the applicability of the research methodology to the highway construction and general construction industries.

Manuscript 4
Chapter 5

Figure 1.1 – Methodology to Manuscript Mapping

1.5 Document Overview

The dissertation is comprised of seven chapters. Chapters 1 and 6 form the 'bookends' of the dissertation document and highlight the overall research questions, methodology, benefits, and contributions to the body of knowledge. Chapters 2 through 5 are standalone research manuscripts which present individual elements of the research methodology and results. As such, the chapters overlap somewhat, particularly in the introduction, background, and literature review sections. However, the manuscripts generally flow chronologically in the order in which the research has been performed. Finally, Chapter 7 presents a bibliography of the research reviewed throughout the dissertation development process.

The following is an overview of the dissertation chapters.

- Chapter One – provides an introduction to the research performed and discusses the overall research questions and benefits. The research methodology is also presented and the dissertation document explained.
- Chapter Two – *The Need for Conceptual Highway Construction Duration Estimating Using Early-known Design Details*. This manuscript presents a construction industry review highlighting the current practices and tools used for estimating construction duration. A literature review of similar studies which addressed the problem in other industries is also presented.

- Chapter Three – *Highway Construction Duration Data Collection and Treatment for Statistical Regression Analysis*. This manuscript discusses the potential construction duration data sources available, the selection of the optimal data source, collection of the data, and the treatment of the data in preparation for analysis. Data collection involves obtaining both contract and project level data for analysis. The data treatment process includes topical screening for erroneous entries and the seasonal adjustment process used to ascertain the historically typical construction duration for each project.
- Chapter Four – *Statistical Regression Analysis to Identify and Model Relationships between Conceptual Design Parameters and Highway Construction Duration*. This manuscript presents the results of the statistical regression analysis. The statistically-significant project factors that influence highway construction duration are identified and quantified. Finally, the models developed are analyzed and the optimal prediction model selected.
- Chapter Five – *Prediction of Highway Construction Duration Using Conceptual Design Parameters*. This manuscript describes the proper use and limitations of the models developed, discussing the model factors and how the models may be incorporated into duration estimating practices. The application of the research methodology is discussed in regards to the entity to which they pertain. The entities to which the research is applicable includes VDOT, highway construction, and the general construction industry.
- Chapter Six – provides research conclusions and overall findings of the work. Each manuscript is summarized and the research findings documented. The contributions to the body of knowledge are described, citing the various research findings and the entities which benefit most readily from the findings.
- Chapter Seven – Bibliography of literature reviewed throughout the research and dissertation development process.

1.6 References

23 Code of Federal Regulations, (2005). CFR635.21. United States of America.

Bromilow, F. J. (1969). "Contract time performance expectations and the reality." *Building Forum*, 1(3).

Bromilow, F. J., Hinds, M. F., and Moody, N. F. (1980). "AIQS Survey of Building Contract Time Performance." *The Building Economist*, 19(2), 79-82.

Burrows, T. K., Pegg, I., and Martin, J. (2005). "Predicting Building Construction Duration." *AACE International Transactions*, PS.14.11-PS.14-19.

Chan, A. P. C. (1999). "Modelling building durations in Hong Kong." *Construction Management and Economics*, 17, 189-196.

Chan, D. W. M., and Kumaraswamy, M., M. (1995). "A study of the factors affecting construction durations in Hong Kong." *Construction Management and Economics*, 13, 319-333.

Chan, D. W. M., and Kumaraswamy, M., M. (1996). "An Evaluation of Construction Time Performance in the Building Industry." *Building and Environment*, 31(6), 569-578.

Clough, R. H., Sears, G. A., and Sears, S. K. (2000). *Construction Project Management*, 4th Ed., John Wiley & Sons, Inc., New York.

Farqahl, S. H., and Everett, J. G. (1997). "Learning Curves: Accuracy in Predicting Future Performance." *ASCE Journal of Construction Engineering and Management*, 123(1), 41-45.

Federal Highway Administration. (2002). "FHWA Guide for Construction Contract Time Determination Procedures." Publication FHWA - TA 5080.15, U.S. Department of Transportation.

Fourie, F. (2003). "Time-for Versus Time-of Performance." *AACE International Transactions*, CDR211-CDR215.

Hancher, D. E., and Rowings, J. E. (1981). "Setting Highway Construction Contract Duration." *ASCE Journal of the Construction Division*, 107(CO2), 169-179.

Nkado, R. N. (1992). "Construction time information system for the building industry." *Construction Management and Economics*, 10, 489-509.

Schexnayder, C. J., and Mayo, R. E. (2004). *Construction Management Fundamentals*, McGraw Hill Higher Education, Boston, MA.

Skitmore, R. M., and Ng, S. T. (2001). "Australian project time-cost analysis: statistical analysis of intertemporal trends." *Construction Management and Economics*, 19, 455-458.

Walker, D. H. T. (1995). "An investigation into construction time performance." *Construction Management and Economics*, 13, 263-274.

Williams, R. C., Vorster, M. C., and Hildreth, J. C. (2008). "The Need for Highway Construction Duration Estimating Using Early-known Design Details." *Transportation Research Record: Journal of the Transportation Research Board*, Submitted for review.

Chapter 2 – The Need for Highway Construction Duration Estimating Using Early-known Design Details

2.1 Introduction

Knowledge of construction duration is pertinent to a number of project planning functions prior to detailed design development. Important project funding and financing decisions are made during the preliminary design stage using estimated construction duration. These decisions, such as the duration of financing, date of occupancy, and the beginning of positive cash flows arising from the construction investment, are imperative to project success. Resource allocation decisions are also critical to project, program, and organizational success. The construction owner must be able to adequately plan and allocate management and oversight personnel and funding. Further, resources must be allocated to multiple projects concurrently. This presents challenges to entire organization, and requires an understanding of construction duration early in project design.

The need for improved preliminary construction duration estimates has been noted by numerous authors and research has taken place in a number of construction industry sectors to support the development of these estimates [Bromilow 1980, Nkado 1992, Chan and Kumaraswamy 1995, Chan 1999, Skitmore and Ng 2001, Burrows et al. 2005]. However, industry research suggests that there is a noticeable lack of research in the ability to reliably and historically predict highway construction duration using early-known project design details. Work has been performed in the residential and building construction arenas internationally [Bromilow 1980, Nkado 1992, Chan and Kumaraswamy 1995, Chan 1999, Skitmore and Ng 2001, Burrows et al. 2005]. However, the domestic highway construction industry has not seen the same level of attention.

2.2 Early-Design Duration Estimates

Early estimates (for cost and duration) allow the construction owner or designer to check the project scope early in the project lifecycle, while the cumulative project cost is relatively low, and while the ability to influence the outcomes of the project are relatively high [Hendrickson 2000]. This concept is depicted in Figure 2.1 below.

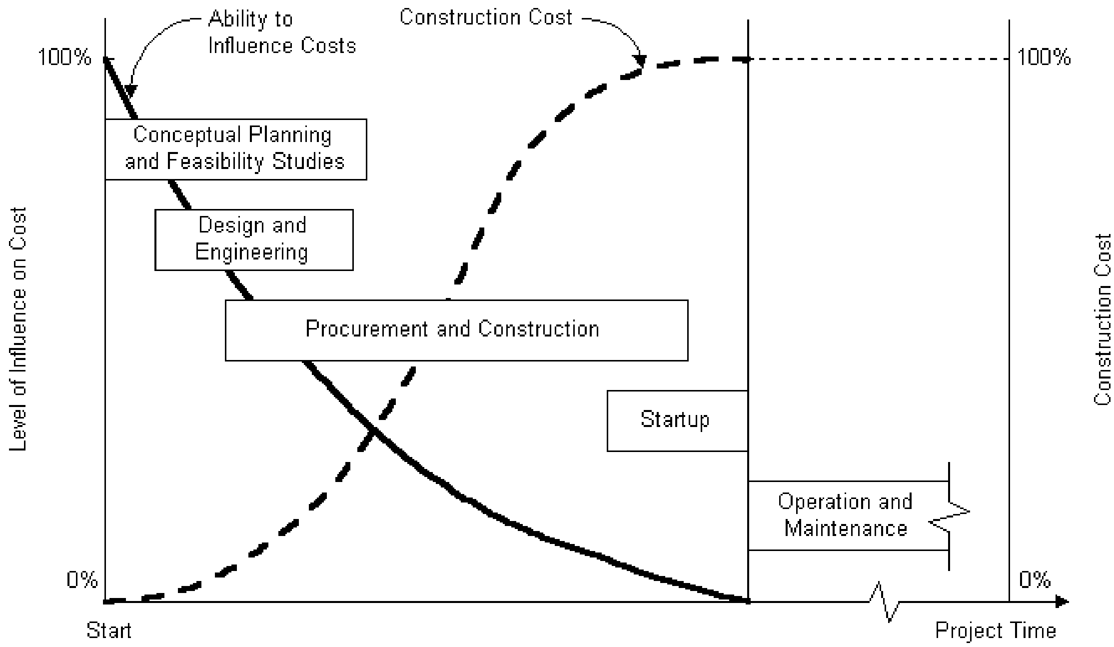

Figure 2.1 – Level of Influence on Cost over Time [Hendrickson 2000]*

*From: Hendrickson, C. (2000). Project Management for Construction, 2nd Edition. Available at: http://pmbook.ce.cmu.edu

The construction owner has a need for both cost and duration estimates that take place during initial design. These estimates are the basis for evaluating future, more refined estimates. Krokowski [1992] defines the primary objective of these early, conceptual, estimates being:

> *"to provide a basis for the capital budget medium term plan, and to provide a basis for the definition technical package (±20%) for project feasibility studies and upper management approval."*

In highway construction, preliminary duration estimates are generally the responsibility of the owner. The most prominent highway construction owner is the state highway agency (SHA). While SHAs maintain large construction budgets and are responsible for construction projects of many types, complexities, and sizes, planning procedures and operations (such as cost and duration estimating) are specific to the state and vary greatly amongst SHAs.

Figure 2.2 shows the general relationship between the accumulation of project cost and the progression of project duration from design through construction. Along the vertical axis of Figure 2.2, project costs begin to accrue during project design and are fully realized at the completion of construction. Similarly, project duration progresses throughout design, advertisement, award, and construction. For the SHA, this process may contain breaks during which time projects are reconsidered or shelved until the necessary funding becomes available. This is depicted in Figure 2.2 at the "Approval" phase of the project. Despite the break in work, project duration and cost continually accrue. Therefore, the construction owner must be able to accurately and reliably predict both the cost (vertical) and time (horizontal) differential between the beginning and end of the construction project phase. The construction cost is regularly and iteratively estimated throughout design, while construction duration is estimated upon design commencement and again just before advertisement. As the basis for budgeting and planning, the importance of accurate and reliable early-design duration estimates cannot be overstated.

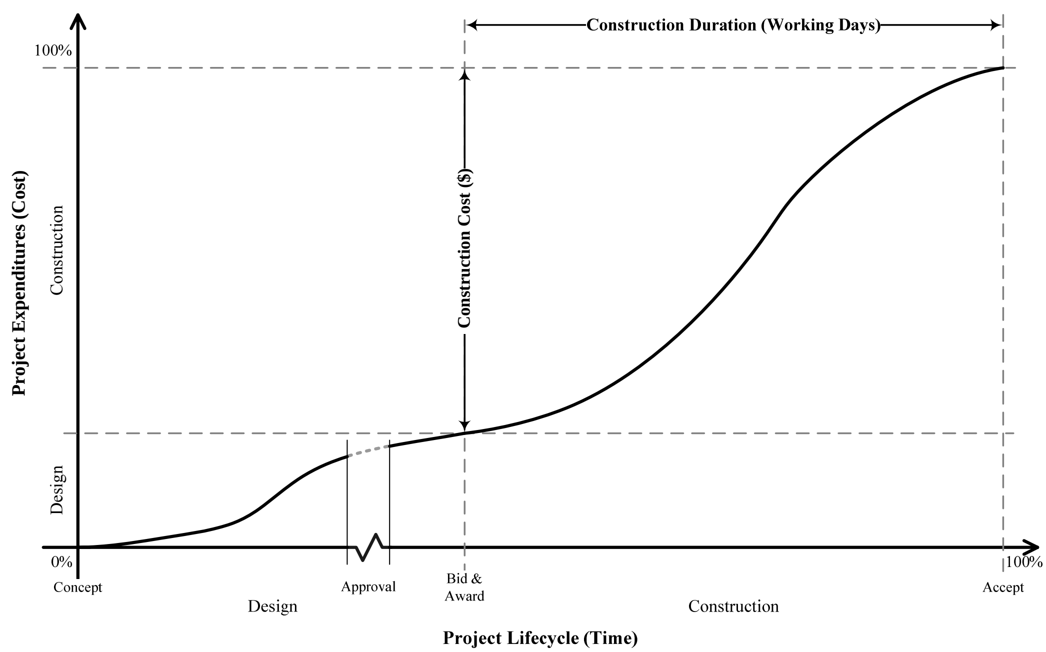

Figure 2.2 – Project Cost and Time Progression

2.3 Highway Construction Industry State of Practice

To better understand the importance of preliminary duration estimates and the existing lack of tools and procedures to assist the SHA planner, this section describes the current state of practice in the U.S. highway construction industry. First, existing tools and procedures currently used and available for estimating construction duration during the preliminary project design phase are identified. Next, the need for additional research in developing systematic, historically-based methodologies for reliably and consistently estimating highway construction duration during initial project design is highlighted. Finally, the benefits that would accrue to highway construction industry entities from the development of additional methodologies for reliably and consistently estimating construction duration using early-known project design parameters are enumerated. The benefits cited are applicable to both the public and private highway construction industry participants.

2.3.1 State of Practice Review Participants

As a recurring or habitual owner, the highway construction owner has many preconstruction responsibilities within the individual project planning area, as well as understanding construction at a program level. Such planning services may also be contracted out to private design and management professionals. These organizations assume responsibility for the construction project, and therefore, perform many of the same planning tasks as the SHA.

To provide a broad understanding of the current practice and needs for improved practices, tools, and procedures construction personnel from SHAs around the United States are described. To understand the specific needs of the SHA, the Virginia Department of Transportation (VDOT) is examined in more detail – examining Location & Design (L&D) and Construction divisions. Within VDOT, the L&D division performs high-level project programming functions. The VDOT State L&D Engineer oversees all plan development, public hearings, surveys, and preliminary planning functions, including cost and duration estimating. Therefore, the construction duration is one element of the larger project duration. The information presented adds to the understanding of the importance of accurate and reliable construction time estimates early in design, the needs of the estimate, and how the estimate changes throughout design. Meanwhile, the construction division is charged with planning the construction to take place and ensuring that work progresses as planned. Inclusion of this information adds to the understanding of the duration estimating needs, practices, and progress of the estimate as project design progresses toward advertisement and construction.

2.3.2 State of Practice Research Method

Due to the diversity of the highway construction industry, understanding the current state of practice required a personal or telephone interview approach. The interview consisted of two parts:

1. Telephone and/or personal interviews and discussions. The first objective of the discussions was to establish interest in participation, affiliation and rank within the organization surveyed, and ability or authority to speak to the topic at hand. Once this background information was established, the interviews served as the primary information collection method.
2. (*Optional*) A follow-up questionnaire forwarded via electronic mail (email) to allow the respondent to cite any additional information felt pertinent to the topic and attach any documentation regarding policies and procedures existing within their organization.

This approach was selected for a number of reasons. First, much of this information is not published. Therefore, the telephone interview of industry practitioners is the often only viable method of acquiring such information. Further, personal conversations allow for a more in-depth discussion of the state of practice, including current processes performed, why they are used, what is gained, and what could be done better. Standardized written surveys often discourage detailed explanations due to lack of space or unwillingness of the survey subject to put their thoughts into writing. Further, the personal conversations help in explaining the intent of the interview, as well as the background research for which it would be used.

2.3.3 Current State of Practice in Highway Construction

This section presents the results of the industry review by the classification of the participants involved. First, information gathered from industry design and management professionals are presented. Next, the results of the nationwide SHA review are presented. Finally, the results of the detailed look at VDOT practices and procedures are discussed.

2.3.3.1 Design and Management Professionals

Design and management professionals regularly contract with SHAs to perform project planning and management services during the project design phase. Therefore, it is necessary to understand the current state of practice, as perceived by these entities. The organizations contacted have extensive experience in multiple areas of construction around the United States. The individuals interviewed were Project Control Managers, or persons of similar company status.

Contacted organizations were first asked about their preliminary duration estimating efforts. Answers to this question were unexpectedly varied. One company interviewed described the use of schedule data collection efforts from projects across the United States. The schedules serve as a means to gather and maintain historical scheduling data for various project types, sizes, locations, etc. This information is used in conjunction with the highest available quantity estimate and assumed production rates based on industry standards. A second company described the use of an 'outline schedule' that details a single section or element of the project that can be repeated within the project. The examples given include single roadway segments or single bridge bents. These outlined schedules could be repeated as many times as necessary and joined to 'build the project' and account for the majority of the work. The durations and production rates for this process were said to come from three main sources: *(1)* practicing field engineers, *(2)* industry workbooks, and *(3)* scheduler experience. Finally, a third interviewee stated that early estimates performed within their organization were based primarily on personnel experience and similar work.

Next, the need for early duration estimates and their primary role within the organization was discussed. Numerous reasons for performing early duration estimates were cited. One of the reasons described was to enhance risk management efforts. Early and quick duration estimates at this high level of design detail allowed the organization to explore a number of options available to the construction owner – offering an optimistic schedule or duration, as well as a 'more likely' schedule or duration. Another reason cited was for environmental impact analysis purposes. It's important to note that the interviewee citing this reason frequently operates in areas having strict emissions policies governing automobiles and heavy equipment. An early duration estimate allows the organization to reasonably estimate equipment operating time. The final reason cited was enhanced long-term resource planning efforts.

The organizations contacted all expressed an interest in further studies aimed at developing a historically-based, statistically sound methodology for preliminary duration estimating. The primary reasons for interest were for the purposes of supplementing field experience and simplification of estimating efforts at the early design development stage. Design professionals commit significant time and resources to the conceptual planning and estimating processes. Therefore, the value added is the ability to more quickly and easily produce a reliable time estimate using few project details and through a system that would allow various iterations and options to be explored.

2.3.3.2 State Highway Agencies

To understand the current state of practice for preliminary duration estimating, contact was attempted with SHAs across the United States. The target interviewee was the head of the SHA's construction division, or an individual of similar rank where necessary. Of the 50 SHAs contacted, 25 were reached. While this was an initial concern, the outcomes of the review do not warrant additional efforts to contact non-responsive SHAs.

Of the 25 responsive SHAs, 20 (80%) recognized a need for estimating construction duration during the conceptual design stage. The needs recognized were varied, but generally focused around the desire to enhance project planning and scheduling efforts. Also, 5 of these 20 stated that the need seen by the SHA was relegated to certain project sizes or conditions. The circumstances cited by these 5 organizations are listed below.

- Two offered specific cost thresholds at $40 and $10 million as projects that warrant construction time estimates during conceptual design.
- Another SHA stated that the need for such an estimate was recognized for 'significant' projects, without offering a threshold cost or size.

- The remaining 2 SHAs stated that their agency recognized a need for preliminary construction time estimates for projects that were part of a series of projects on the same alignment or system.

While the majority recognized a need, a number of respondents also stated that while there is a need for such procedures or tools, the estimate developed at this stage would only be an approximation, narrowed to the number of months or seasons of construction.

In regards to currently available tools and procedures, of the 20 respondents that recognized a need for such early estimates, 10 of them (50%) replied that their SHA did not have existing procedures or tools for estimating construction time during the preliminary design phase. Three other respondents replied with detailed estimating procedures and policies that required final design information not typically available during the preliminary design stage. This leads the researcher to believe that procedures or tools for conceptual construction duration estimates do not actually exist within these SHAs. This increases the percentage of SHAs who recognize a need for conceptual construction time estimates without available procedures or tools to 13 out of 20 (65%). These results and the responses from the remaining 7 SHAs surveyed are outlined below.

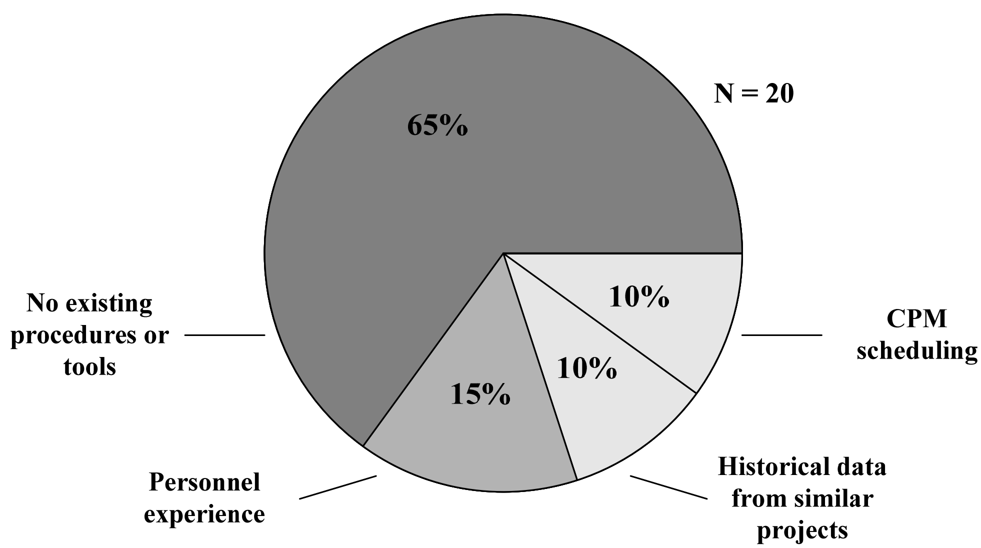

Figure 2.3 – Existing Preliminary Duration Estimating Practices

- 3 of the remaining 7 SHAs with preliminary construction duration estimating procedures state that preliminary construction duration estimates are prepared by experienced personnel and based on personal experience.

- 2 of the 7 specify the use of historical data from similar projects to develop a duration estimate for new projects. Upon further questioning, these SHAs described this process as comparing the project in question to similar projects in the area. The construction duration estimate could then be customized based on the project differences. This customization is performed using personnel experience.

- The final 2 respondents state the use of CPM schedules at this early design stage. CPM schedules were developed using personal knowledge, historical data, and assumed quantities and methods.

The interview discussion concluded by inquiring about the benefits recognized in developing a systematic, historically-based methodology for predicting construction duration during the preliminary design phase. Typically, the respondents who recognized a need for preliminary construction duration estimates also cited perceived benefits. However, one SHA who did not recognize a need for these early estimates did perceive a potential benefit arising from the methodology proposed as improving workforce planning and decision making processes. The remaining 4 respondents who did not recognize a need for preliminary construction duration estimates also did not recognize any potential benefits to the methodology proposed.

For simplicity, the benefits recognized by the 20 positive respondents are listed in order of response rate. The benefits recognized include:

1. Enhancement of project planning efforts and procedures (55%).
2. Improvement of funding knowledge or financial planning efforts (30%).
3. Improvement of public convenience and safety (30%).
4. Enhancement of resource planning measures (20%).
5. Improvement of cost estimating practices and procedures (10%).
6. Ability to consider additional design alternatives (10%).

Less explicitly cited benefits included the potential to employ innovative contracting methods, the ability to improve contractor competition, and accelerated contract completion through more reliable time estimates.

2.3.3.3 VDOT Duration Estimating Process

VDOT has been selected as an example of planning and design processes used by SHAs. This selection was based on a number of factors including the available documentation regarding planning, design, and estimating practices, the willingness to participate in the research, and the proximity to the researcher.

The VDOT project design process is guided by the Project Development Concurrent Engineering Process (PDCEP). This process provides a standard procedure for transitioning the project through pre-scoping, scoping, initial, preliminary, detailed, and final design, right of way and utilities acquisition, and construction planning [VDOT 2008a]. VDOT maintains a flowchart of the PDCEP to guide the project planning process (Available at: http://www.virginiadot.org/business/resources/LocDes/PDCE.pdf) [VDOT 2008a]. The main phases of the PDCEP are seen in the figure below. PDCEP milestones are on the bottom of Figure 2.4, while the corresponding design phases are shown between the milestones in the top portion of the figure.

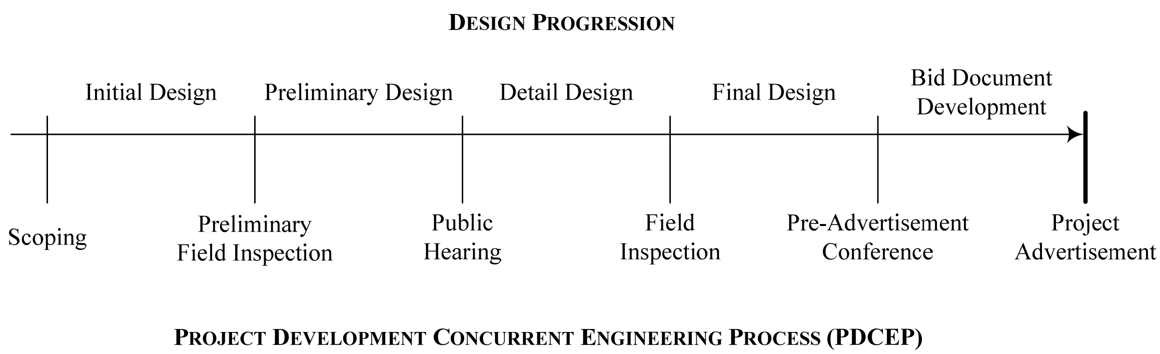

Figure 2.4 – VDOT PDCEP and Design Progression

Williams [2006] demonstrated the need for duration estimating taking place throughout the design process. Just as the construction cost estimate is continually refined throughout design, Figure 2.5 shows that the construction duration estimate should proceed similarly. To date, VDOT has developed and implemented a database of historical construction performance data for use in the pre-advertisement level duration estimating process [Williams 2006]. However, the organization does not currently have a statewide process for estimating construction duration during the initial and preliminary project design phases of the PDCEP [Williams 2006].

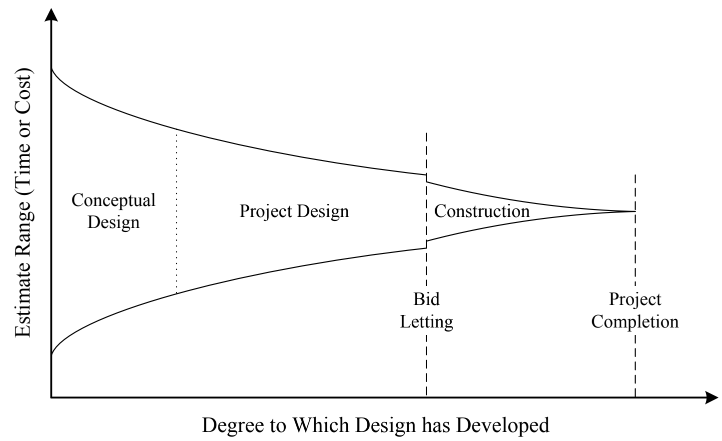

Figure 2.5 – Cost and Time Estimate Refinement with Design [Williams 2006]*

* From: Williams, R. C. (2006). "The Framework of a Multi-Level Database of Highway Construction Performance Times," Virginia Polytechnic Institute and State University, Blacksburg, VA.

This section discusses the current duration estimating processes within VDOT. Of interest here, are the needs of the organization, type of estimates performed, and the personnel involved with the various duration estimates.

Two VDOT divisions are studied herein: L&D and construction division. These divisions have the most readily applicable need for preliminary construction duration estimates. The L&D division is responsible for the construction project at the earliest conceptual design stage. Throughout design, the L&D division is also responsible for overseeing the individual project planning processes, as well as managing project development at a program level. Construction division personnel are involved during early design, but are typically responsible for more detailed planning activities and the actual post-award construction activities.

2.3.3.3.1 VDOT Location & Design Division

Since L&D division involvement spans many months, and often many years, the continual increase in project costs over time is very important to the understanding and development of the project. Of particular interest is construction cost inflation over the life of the project changes and subsequent delays on construction. Knowledge of the construction duration is critical to planning for construction costs. L&D also has a need to understand the impacts of design delays on construction. Since the time of year in which work is performed has a direct impact on the productivity and speed at which work can be performed, it is imperative to understand the construction duration when establishing the deadlines for design development.

The duration estimating practices of the L&D division are inline with that expected for program-level project management. The first duration estimate that takes place occurs upon scoping commencement. At that time, both the duration and cost estimates are "gross estimates" to offer an idea of the project duration. It is important to note here that project duration is different than construction duration. Project duration refers to the length of time between the commencement of design and the contract acceptance date. For the L&D division, the construction duration is just one element of project duration. Figure 2.6 depicts the L&D responsibilities throughout design development and 'construction'. Notice the 'construction' phase runs to project advertisement, which is essentially the end of the L&D division responsibilities [VDOT 2008b].

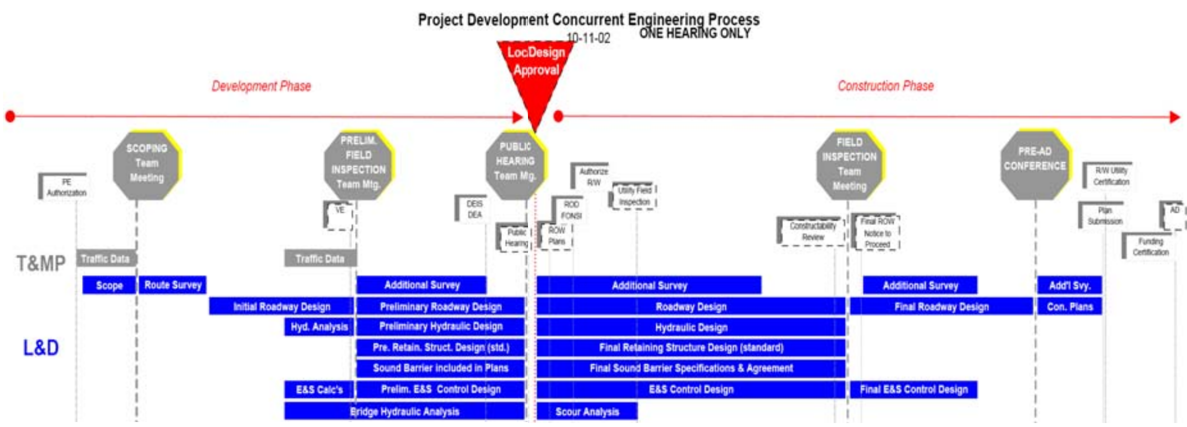

Figure 2.6 – L&D Responsibilities in PDCEP [VDOT 2008b]*

* From: Virginia Department of Transportation (2008b). "Project Development Concurrent Engineering Process." http://www.virginiadot.org/projects/resources/CE-Process.pdf

With this first gross estimate, the L&D division seeks to understand when to expect certain milestones, including project advertisement and construction completion. This estimate is based on the experience of the personnel and also used to judge the feasibility of the project.

Typically, only one more duration estimate is made by the L&D division. This occurs at the Preliminary Field Inspection stage. Since the initial project design is understood, this estimate is based on the quantity of work to be performed. Again, this is an experience-based estimate, relying on experienced and knowledgeable personnel. Typically, personnel develop a bar chart that shows project phases (i.e. design, procurement, and construction) with time being measured in months.

Unless there is a significant change in the project scope or design, this is the final estimate prepared by L&D. Later estimates serve only to validate the original estimate and ensure the project scope has not been significantly altered. These changes show up as drastic increases in either the cost or duration estimate. Typically, though, a scope or design change that results in a duration change of six months or less can be easily accommodated.

Despite the rudimentary nature of duration estimates currently performed, a historically-based methodology for estimating construction duration during preliminary design is great importance to the L&D division. While current estimates are performed using individual industry experience, a historically-based method or process is seen as a useful tool in justifying and verifying the experience-based estimate. Also, because the estimates are experience-based, there is potential in a historically-based method or process as a way to maintain construction knowledge that is typically lost when individuals resign, retire, or move between organizations.

2.3.3.3.2 VDOT Construction Division

The VDOT construction division was also studied in order to document and map-out the duration estimating process, the steps taken or processes used to prepare the estimates, and when these steps are performed, within the PDCEP process. There are a number of important outcomes to note from this work. First, there are not organization-wide standards or methods for estimating and establishing construction or contract duration within VDOT. A majority of the nine districts were surveyed through this work. While the intent of scheduling and duration estimating efforts are uniform, the process by which time estimates evolve through design varies significantly among districts.

Most districts studied perform multiple construction and contract time estimates. The majority of these districts suggested two distinct duration estimates taking place throughout design. As expected, the final time estimate takes place near the Pre-advertisement Conference (PAC), just prior to bid documents being prepared for advertisement. This final duration estimate is a detailed estimate, sufficient for establishing a fixed-date for contract completion.

Conversely, early duration estimates are not uniformly prepared throughout the state. Two districts mention estimating construction and contract duration prior to the Field Inspection (FI) stage of the PDCEP. Both Salem and Bristol districts prepare early duration estimates based on experience and comparison to similar projects in the district. These estimates are generally broad, using 'years' or 'seasons' to estimate the project duration. Salem district refines that duration estimate at the Preliminary Field Inspection (PFI) stage using a generalized sequence of construction, assumed quantities, and experience-based performance data. As projects near 75% design completion at the FI stage, 3 districts (Northern Virginia (NoVA), Richmond, and Salem) prepare a more detailed duration estimate using forthcoming design details. Finally, as project design nears the PAC stage, early estimates prepared by NoVa, Richmond, and Salem are further refined by the scheduling team. Duration estimates prepared by Bristol designers are revised by an individual in the Scheduling & Contracts division, based on his personal experience and expertise in construction. Finally, Staunton district schedulers use the nearly complete design to prepare a contract duration estimate for submittal at the PAC. These practices are visually depicted in Figure 2.7.

Another important finding of this work is that while practices vary throughout the districts, all districts noted that their scheduling and duration estimating procedures are evolving, becoming more sophisticated. Unfortunately, the districts are at various stages in this evolution. Salem, NoVa, and Richmond are currently practicing multiple duration estimates and schedule iterations throughout design. Bristol district also uses and early estimate on some projects, however, this practice is much more informal than that of other districts. Schedulers in Bristol are expecting changes to their processes in the near future. Finally, Staunton District has proposed a significant shift in duration estimating and scheduling practices to be incorporated in the near future.

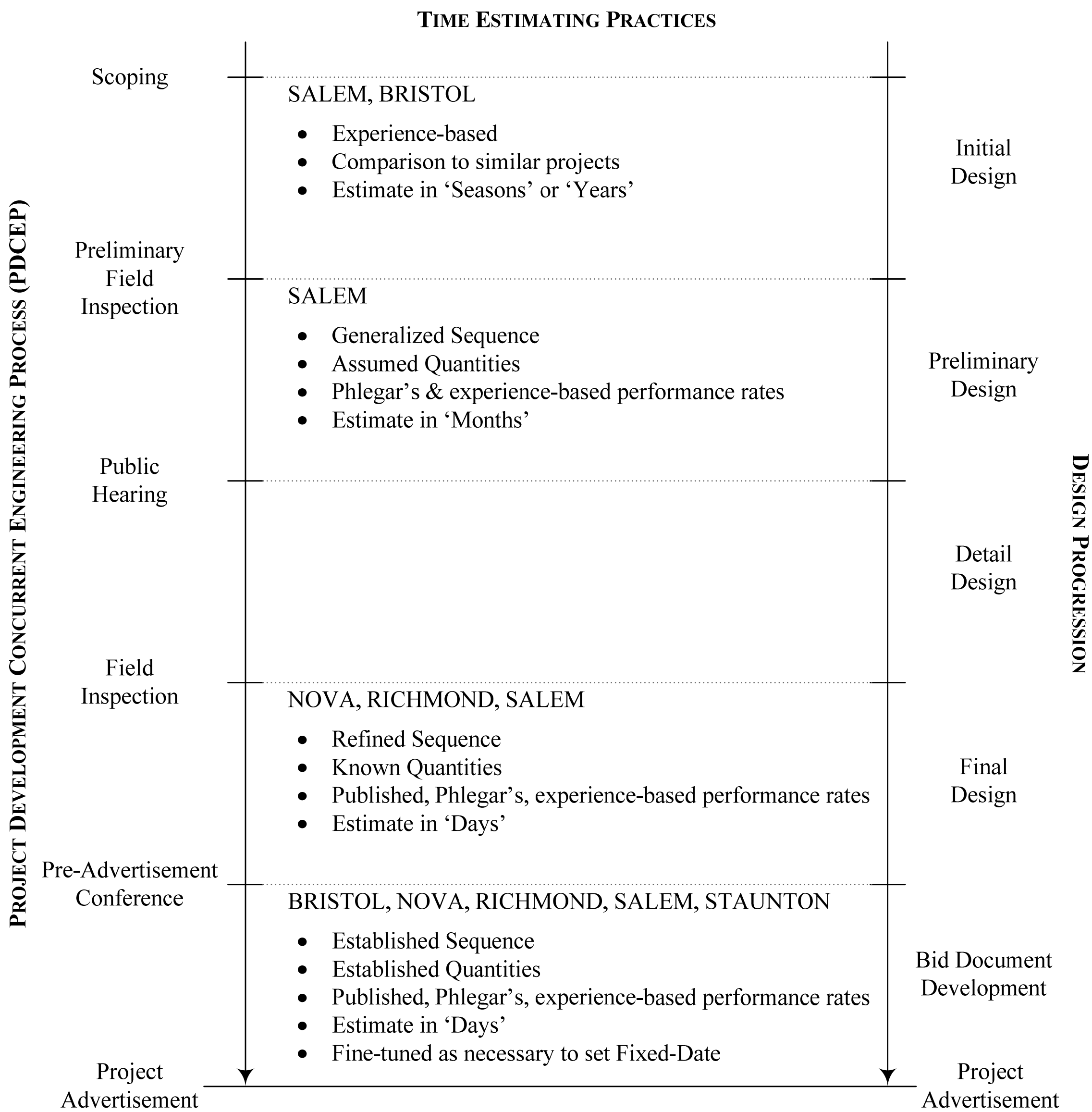

Figure 2.7 – District Duration Estimate Progression with Design

2.3.4 State of Practice Conclusions

Highway construction industry entities, both private design professionals and SHAs, recognize the significance of accurate and reliable preliminary construction duration estimates. For the design professional, the need is specifically in providing the best value to the construction owner by developing design alternatives and having the ability to determine the impacts of these design alternatives, presenting the various options to the construction owner. For the highway construction owner, the importance is placed on project management at the program and the construction division levels. At the program levels, the importance of verifying and justifying the current experience-based estimating procedure is recognized. Further, numerous project decisions occur during preliminary design and there is a need to understand how these design changes will impact the eventual construction duration. Construction division personnel recognize the need for a preliminary estimate as the starting point for their scheduling and planning work.

Despite these needs, the majority of contemporary preliminary estimating practices are often solely based on the experience of planning personnel. Such estimates have been shown to be inaccurate [Farqahl and Everett 1997]. The alternative method used is the development of a detailed CPM schedule that makes use of assumed or incomplete quantities, sequence of construction, and construction means and methods. Others, still, have no means for developing a preliminary construction duration estimate.

Commonly cited benefits to a historically-based, systematic method or procedure for estimating construction duration include the enhancement of project planning efforts and procedures, improvement of funding knowledge or financial planning efforts, improvement of public convenience and safety, and enhancement of resource planning measures.

2.4 Construction Industry Research

A review of the current state of preliminary duration estimating practices in the highway construction industry identifies a significant need for historically-based systematic methods for estimating construction duration early in project design. Other industry sectors have sought to resolve similar needs by using statistical regression analysis (SRA) to quantify and model the relationship between project parameters or factors and duration. This section presents a brief review of research performed in other construction industry sectors using SRA to reliably predict duration using early-known project details.

The works cited here are segregated into two sections *(1) Bromilow's Time-Cost Model*, and *(2) Parametric Regression Analysis to Predict Duration*. Bromilow's work stemmed a number of works in the project duration research field.

2.4.1 Bromilow's Time-Cost Model

Bromilow [1969] is the first author referenced to have studied the prediction of construction duration using early-known project details and SRA. Bromilow [1969] developed a regression model for predicting building construction duration in terms of estimated cost. Bromilow's time-cost model (BTC) has been the basis for many additional research studies relating to construction time determination [Kaka and Price 1991, Kumaraswamy and Chan 1995, Chan 1999, Ng et al. 2001, Skitmore and Ng 2001]. The regression model developed by Bromilow [1969] is shown below:

$$T = KC^{B}$$

Equation 2.1 – Bromilow's Time-Cost Model [Bromilow 1969]

where: T = actual construction time in working days,

C = final cost of building in millions of dollars,

K = constant characteristic of building time performance,

B = constant indicative of the sensitivity of time performance to cost level.

The sole use of cost as indicator for time was found only after exploring gross or net floor area, ground area, number of floors, and building volume. In fact, Bromilow [1969] also found that the construction time for building does not depend "very strongly" on the type of building or its location. Figure 3.8 below shows the non-linear relationship between construction duration and project cost found through investigation of over 300 building construction projects in Australia [Bromilow 1969].

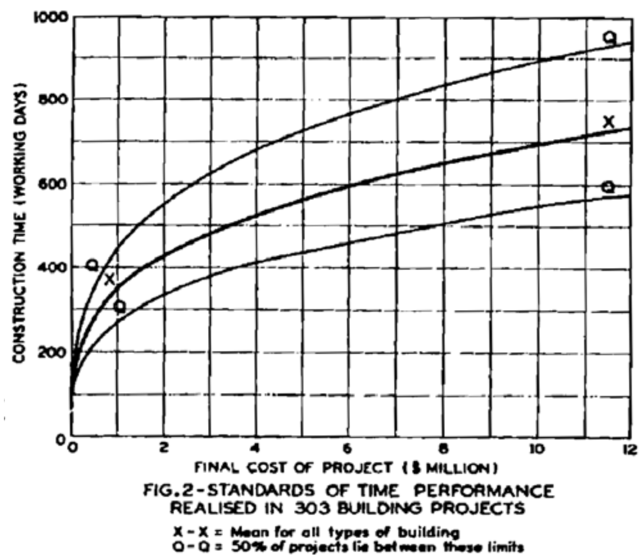

Figure 2.8 – Project Cost vs. Construction Time [Bromilow 1969]*

* From: Bromilow, F. J. (1969). "Contract time performance expectations and the reality." Building Forum, 1(3).

While Bromilow's model did not incorporate additional factors in estimating construction time, other authors have not agreed. Ng et al. [2001] modified the BTC model and concluded that different parameter estimates are needed for different project types. To accommodate the differences between projects, Ng et al. modified the BTC's components to accommodate two different project types: one for industrial projects and one for non-industrial projects.

The work performed by Ng et al. [2001] was followed up by Skitmore and Ng [2001]. This study set out to determine whether a simpler ratio measure of the relationship between time and cost could be found [Skitmore and Ng 2001]. The authors developed the following ratio:

$$\mathrm{B} = \frac{\ln T}{\ln c}$$

Equation 2.2 – Modified Time-Cost Model [Skitmore and Ng 2001]

where: T = construction time in working days,

C = project cost in dollars,

B = constant indicative of the sensitivity of time performance to cost level.

Kaka and Price [1991] extended beyond modeling construction time using only cost. Instead, Kaka and Price [1991] set out to demonstrate first, the difference between modeling public and private buildings, as well as civil engineering projects. Citing a work by the National Economic Development Office (NEDO) in 1988, the authors also supposed and tested the influence of the type of client (public or private), type of tender, and the form of tender [Kaka and Price 1991]. Through this study, Kaka and Price [1991] reached the following conclusions: *(1)* the type of bid competition did not affect the reliability of the BTC model; *(2)* the type of client (public or private) does influence the time-cost relationship with public building works generally taking longer than private works; *(3)* the type of project affected the relationship considerably with civil engineering works taking less time to complete than buildings of the same value; and *(4)* the form of the contract significantly influenced the time-cost relationship with adjusted price contracts being the largest (in dollars) and longest (in working days) [Kaka and Price 1991].

Kumaraswamy and Chan [1995] set out to determine the influences of building parameters on construction time in the Hong Kong construction market. Kumaraswamy and Chan [1995] first confirmed the BTC model. The authors also considered other expected influential factors, particularly, the influence of building floor area on construction time [Kumaraswamy and Chan 1995]. To explore this relationship, the authors developed a model in the form:

$$T = LA^M$$

Equation 2.3 – Floor Area and Construction Duration [Kumaraswamy and Chan 1995]

where: A = floor area in square meters (m^2).

The other model factors (T, L, and M) are the same as those in the BTC model [Kumarasway and Chan 1995]. The authors found this model to be statistically significant, but recommended the consideration of numerous other project factors in additional phases of research.

Chan and Kumaraswamy [1995] investigated relationships between building project variables such as the construction duration, construction cost, gross floor area, and the number of building stories [Chan and Kumaraswamy 1995]. This research confirmed the earlier discovered difference between the public and private construction industries. Public sector buildings and civil engineering works performed much more consistently while private building works were much more unpredictable [Chan and

Kumaraswamy 1995]. This research phase also confirmed a statistically significant relationship between the construction time and the number of building stories [Chan and Kumaraswamy 1995].

Finally, Chan [1999] confirmed the original BTC model and established the factors generally applicable in Hong Kong [Chan 1999]. Chan [1999] also confirmed the findings by Kaka and Price [1991] that publicly funded building construction projects generally take longer than their similarly valued private counterparts.

2.4.2 Parametric Regression Analysis to Predict Duration

Collection of research relating to parametric regression analysis for predicting construction duration demonstrated a large detachment between the work taking place in the United States, and that being performed internationally in the United Kingdom, Australia, and Hong Kong.

Orczyk [1989] proposed identifying and modeling milestone dates for construction projects using conceptual design parameters. To do so, Orczyk [1989] first surveyed construction industry professionals to determine those project milestones and parameters most crucial to construction duration for a small office building (1 to 4 story). The survey results indicated that the five most crucial factors to construction schedules were type of frame, owner's schedule requirements, subsurface conditions, type of cladding, and number of floors [Orczyk 1989]. Orczyk [1989] next developed a survey requesting actual historical data on these parameters from across the United States. Through model development and schedule simulation, Orczyk [1989] was able to explain 72% of the variation in the parameters and the timing or occurrence of events. The total area of the building accounted for 45% of variation explained [Orczyk 1989]. Finally, Orczyk [1989] reviewed the results of surveys sent to highway and bridge constructors. While there was not a sufficient amount of data returned for analysis, Orczyk [1989] cites the most commonly reported influential parameters for highway and bridge construction as weather and the volume of earthwork. Meanwhile, completion of sub-grade, paving, bridge substructure, and bridge superstructure as significant milestones in the construction process [Orczyk 1989].

A study performed by the United States Army Corps of Engineers states describes the prevalence of underestimating construction duration within in military and civil works [East et al. 1992]:

> *"In fiscal year 1988, actual duration of military construction projects took an average of 17 percent longer than estimated. Similarly, actual duration of civil construction projects averaged 19 percent longer than estimated."*

To remedy this problem, the study incorporates three factors found to unexpectedly extend construction activities: work delays, weather delays, and productivity delays. From this information, a contract scheduling system is developed that requires the input of project parameters and specific activity information and durations [East et al. 1992]. While this study was focused on building construction, a matrix of influential activity factors on building construction activities are useful to the study at hand. A number of factors found to affect building construction activity are also expected to impact highway construction.

Nkado [1992] found eight statistically significant building construction time variables: gross floor area, height, type of cladding, number of stories, location, predominant frame, storey height, and approximate building volume. To incorporate the large amount of variables into a model, Nkado [1992] developed regression models to predict durations of individual project components, then modeled the relationship between the duration or size of those components and the lag time between successive component starts. Once developed, the models developed were tested against the estimates from nine planners from three separate project offices and found to be within the distribution of planner estimates [Nkado 1992].

Next, Khosrowshahi and Kaka [1996] sought to determine regression models for both building project cost and duration. The authors considered such factors as the project type, operation, scope, form, structural properties, cost, ground condition, building height (in stories), site access, and build-ability [Khosrowshahi and Kaka 1996]. Regression analysis determined a duration estimation model that incorporates cost, access, build-ability, scope, operation, framing, start month, and building height [Khosrowshahi and Kaka 1996]. The authors note that a log transform of cost within the equation produces a better prediction of duration [Khosrowshahi and Kaka 1996].

In the Hong Kong public housing sector, Chan and Kumaraswamy [1999a, 1999b] found that prediction equations typically included the number of stories, gross floor area, ratio of gross floor area to ground floor plan area, external cladding area, type of foundation, information exchange between architect/contractors, ground conditions, and labor productivity [Chan and Kumaraswamy 1999a, Chan and Kumaraswamy 1999b]. Notice several of these same factors exist in their earlier building industry prediction models.

Chan and Chan [2004] developed a benchmark model for project time performance in Hong Kong. Of particular interest to Chan and Chan [2004] was the identification of critical factors influencing construction durations of high-rise public housing buildings in Hong Kong. Through their analysis, Chan and Chan [2004] determined five statistically significant regression variables: *(1)* total construction cost, *(2)* type of housing scheme, *(3)* use of pre-cast facades, *(4)* building volume, and *(5)* ground floor area per floor.

Burrows et al. [2005] studied the relationship between project sector, procurement route, contractor selection method, client type, building function, and location to the building construction duration using data from more than 1,500 new building construction projects in the U.K. between 1998 and 2002. Each of these categories were subdivided into a number of more detailed classifications, each of which analyzed for their relationship to building construction duration [Burrows et al. 2005]. The results provided a number of important insights regarding the relationship between the aforementioned factors and the construction duration. First, there is a significant relationship between project cost and duration [Burrows et al. 2005]. This relationship is more pronounced as the cost increases [Burrows et al. 2005]. The study also revealed that housing projects tend to take the longest to complete, all other parameters being equal [Burrows et al. 2005]. Finally trends between the location or region and the project duration were noticed. Small projects tended to have similar durations through the country while larger projects, on the other hand, tended to be completed in less time in urban regions as opposed to rural regions [Burrows et al. 2005].

As seen, other construction industry sectors have investigated the relationships between early-known design parameters and project duration (both contract and construction duration). SRA is the most commonly cited method for identifying and quantifying the relationships between early-known project details and duration. It is expected that the highway construction industry would realize similar results.

2.5 Summary

The current state of practice within the highway construction industry, including findings from both design professionals and SHAs, is presented. Because SHAs are the most prominent highway construction owner, a more detailed investigation was undertaken to explore construction planning processes and procedures used by SHAs across the United States. The need for additional or enhanced methods, procedures, and tools was established, and the potential benefits of such improvements were documented. Finally, VDOT was studied to determine the current practices, needs, and personnel involved in the duration estimating process, as it proceeds throughout the design development process.

While organizations varied, there are commonalities across the industry. First, SHAs and design professionals alike recognize the importance and need for accurate and reliable highway construction duration early in the project design development lifecycle. Next, despite this need, SHAs generally do not possess a standardized, historically-based, reliable methodology, process, or procedure for estimating construction duration during preliminary design. Such estimates are often based on comparisons to similar projects, the development of an overcomplicated CPM schedule with multiple assumptions, or the opinion of experienced personnel and can result in inaccurate estimates. Design professionals perform such estimates in a variety of methods, mirroring the current practices of SHAs. Finally, the industry recognizes numerous potential benefits, including enhanced project resource planning efforts and procedures, improved funding knowledge or financial planning efforts, improved public convenience and safety, and enhanced resource planning measures.

This document also presents a synopsis of preliminary construction duration research from the United States and abroad. As stated, there exists a discontinuity between the attention and research granted to construction industry sectors. While building and residential construction sector researchers have presented an understanding of the early-known project design details that impact construction duration [Bromilow 1980, Khosrowshahi and Kaka 1996, Chan and Kumaraswamy 1995, Skitmore and Ng 2001, Chan and Chan 2004], highway construction has not. Further, the building construction industry has investigated the relationships between these early-known factors and construction duration [Nkado 1992, Chan and Kumaraswamy 1999a, Chan and Kumaraswamy 1999b, Burrows et al. 2005]. Highway construction has no such understanding or quantification of the early-known parameters or factors that influence construction duration. Therefore, there is a need for additional research in the highway construction industry sector.

2.6 References

Bromilow, F. J. (1969). "Contract time performance expectations and the reality." *Building Forum*, 1(3).

Bromilow, F. J., Hinds, M. F., and Moody, N. F. (1980). "AIQS Survey of Building Contract Time Performance." *The Building Economist*, 19(2), 79-82.

Burrows, T. K., Pegg, I., and Martin, J. (2005). "Predicting Building Construction Duration." *AACE International Transactions*, PS.14.1-PS.14-9.

Chan, A. P. C. (1999). "Modelling building durations in Hong Kong." *Construction Management and Economics*, 17, 189-196.

Chan, A. P. C., and Chan, D. W. M. (2004). "Developing a benchmark model for project construction time performance in Hong Kong." *Building and Environment*, 39, 339-349.

Chan, D. W. M., and Kumaraswamy, M., M. (1995). "A study of the factors affecting construction durations in Hong Kong." *Construction Management and Economics*, 13, 319-333.

Chan, D. W. M., and Kumaraswamy, M., M. (1999a). "Forecasting construction durations for public housing projects: a Hong Kong perpective." *Building and Environment*, 34, 633-646.

Chan, D. W. M., and Kumaraswamy, M., M. (1999b). "Modelling and predicting construction durations in Hong Kong public housing." *Construction Management and Economics*, 17, 351-362

East, E. W., Hassanein, A., and Melin, J. (1992). "Evaluation Factors that Affect Construction Project Duration." US Army Corps of Engineers.

Farqahl, S. H., and Everett, J. G. (1997). "Learning Curves: Accuracy in Predicting Future Performance." *ASCE Journal of Construction Engineering and Management*, 123(1), 41-45.

Hendrickson, C. (2000). *Project Management for Construction*, 2nd Edition. Available at: http://pmbook.ce.cmu.edu.

Kaka, A. P., and Price, A. D. F. (1991). "Relationship between value and duration of construction projects." *Construction Management and Economics*, 9, 383-400

Khosrowshahi, F., and Kaka, A. P. (1996). "Estimation of Project Total Cost and Duration for Housing Projects in the U.K." *Building and Environment*, 31(4), 373-383.

Krokowski, A. W. (1992). "Estimating from Process Flow Diagrams: A Conceptual Approach." *AACE International Transactions*, 1, A.3.1-A.3.6.

Kumaraswamy, M., M., and Chan, D. W. M. (1995). "Determinants of construction duration." *Construction Management and Economics*, 13, 209-217.

Nkado, R. N. (1992). "Construction time information system for the building industry." *Construction Management and Economics*, 10, 489-509.

Ng, S. T., Mak, M. M. Y., Skitmore, R. M., Lam, K. C., and Varnam, M. (2001). "The predictive ability of Bromilow's time-cost model." *Construction Management and Economics*, 19, 165-173.

Orczyk, J. J. (1989). "Parametric Construction Scheduling," Dissertation, Purdue University, West Lafayette, Indiana.

Skitmore, R. M., and Ng, S. T. (2001). "Australian project time-cost analysis: statistical analysis of intertemporal trends." *Construction Management and Economics*, 19, 455-458.

Virginia Department of Transportation (2008a). "Project Development Concurrent Engineering Process." Available at: http://www.virginiadot.org/business/resources/LocDes/PDCE.pdf

Virginia Department of Transportation (2008b). "Project Development Concurrent Engineering Process." Available at: http://www.virginiadot.org/projects/resources/CE-Process.pdf

Williams, R. C. (2006). "The Framework of a Multi-Level Database of Highway Construction Performance Times," Virginia Polytechnic Institute and State University, Blacksburg, VA.

Chapter 3 – Highway Construction Duration Data Collection and Treatment in Preparation for Statistical Regression Analysis

3.1 Introduction

A study of state highway agencies (SHAs) across the United States identified the lack of historically-based conceptual level construction time estimating practices and procedures [Williams et al. 2008]. In fact, while 80 percent of SHAs stated the need for improved construction duration estimating procedures, most had no systematic method for estimating construction duration early in the project development process [Williams et al. 2008].

Meanwhile, other construction industry sectors (including building, residential, and industrial construction) have witnessed success in using statistical regression analysis (SRA) to model the relationship between project factors or parameters and project duration [Bromilow 1980, Nkado 1992, Chan and Kumaraswamy 1995, Chan 1999, Skitmore and Ng 2001, Burrows 2005]. Despite this, there are few studies aimed at using a historically-based modeling methodology, like SRA, to determine duration-influential project factors and parameters, quantify their relationship to duration, or develop a mathematical model for estimating construction duration using these early-known project design details [Williams et al. 2008]. Jiang and Hongbo [2007] is the only source cited that investigates highway construction duration estimating using SRA. Jiang and Hongbo [2007] focus on the relationship between contract duration and project cost, developing a series of mathematical models for numerous project types. Jiang and Hongbo [2007] also analyze contract level data, such as project size, location, and construction management characteristics to adjust the estimated contract duration calculated, however, these “adjustment factors” are not included in the regression equations.

It is hypothesized that numerous other project parameters and factors influence highway construction duration. Such information is more detailed than the contract level information investigated by Jiang and Hongbo [2007], but known early enough in the project development process to incorporate into a conceptual level duration estimating model.

Here, the data needed for such a conceptual level duration analysis is described. This discussion includes the type of data needed, as extrapolated from similar studies in other construction industry sectors. The data source selection process is discussed. Finally, the data treatment process is described. The data treatment process describes the processes involved in readying the construction project data for SRA. This includes data compilation, pre-treatment to remove erroneous data, and seasonal duration adjustment procedures.

3.2 Data Acquisition

Because construction duration estimating in highway construction has not seen as much attention as other industry sectors, the data needed for such a study is largely extrapolated based on previously completed works from other construction industry sectors. This process includes identifying the type of data needed for such a study, selecting a data source, and data collection and compilation.

3.2.1 Identification of Statistical Regression Analysis Data Parameters

While there is little research offered on the parameters or data fields that are thought to relate to highway construction duration, numerous studies have taken place in the residential and building construction sectors [Bromilow 1980, Nkado 1992, Chan and Kumaraswamy 1995, Chan 1999, Skitmore and Ng 2001, Burrows 2005]. These studies inform the search for the type of data needed for this research. Here, the parallel studies are cited to assist in identifying the data required for a SRA study in the highway construction duration arena.

Jiang and Hongbo [2007] used SRA to model the relationship between highway contract duration and project cost and developed adjustment factors to consider the influence of other contract level project factors on the contract duration. These adjustment factors include type of highway, traffic volume, weather conditions, and location [Jiang and Hongbo 2007].

Orczyk [1989] presented results of a small survey of highway construction professionals to identify project milestones and parameters felt influential on construction time. Limited responses to the survey lead Orczyk to pursue analysis in the building construction sector. However, the responsive highway construction personnel relayed that the most influential highway and bridge project parameters were weather and volume of earthwork [Orczyk 1989]. The most influential or important milestones to meet, according to Orczyk [1989] were completion of sub-grade, paving, bridge superstructure, and bridge substructure.

Earlier studies in the building and residential construction industries have been performed in a similar fashion to that proposed herein. Bromilow [1969] is the first study cited in this regard and determined that the total project cost was directly related to the construction duration. A more recent study used the same simple model as Bromilow [1969] but used the building floor area (m^2) to predict construction duration [Kumaraswamy and Chan 1995]. Khosrowshahi and Kaka [1995] developed more complex models for estimating building construction duration using operation, scope, start month, structural properties, cost, building height, and site access, and buildability. Finally, in the residential arena, Chan and Chan [2004] developed a model for predicting construction duration using construction cost, type of housing scheme, use of pre-cast facades, building volume, and ground floor area.

From the studies cited, it is clear that a variety of project factors influence the construction duration. These factors pertain to both the contract level details (location, type, cost) and the project level details which quantify the complexity and size of the project being constructed. As such, it is expected that a study into the factors which influence highway construction duration should make use of both contract and project level factors to adequately model highway construction duration.

3.2.2 Highway Construction Data Parameters

As seen the parameters studied include both high-level contractual data fields and more detailed, parametric project details such as the floor area, site access, buildability properties, and structural properties. Therefore, it is expected that a study similar to these in the highway construction realm requires historical data at these same levels.

In highway construction, contractual data includes the project type, location, terrain, traffic volume, cost, and overall length. These fields serve to describe the conditions under which the work is performed. Such data is regularly maintained by SHAs for administrator-level program oversight.

Parametric data for highway construction projects varies by project type. For instance, roadway construction projects are dissimilar from bridge construction projects at the parametric level. Roadway construction parameters likely include number of lanes or lane miles, earthwork volumes, number of signalized intersections, number and length of turning lanes, etc. Bridge construction project parameters of interest likely include bridge type (steel or concrete), number of spans, length, height, width, and number of lanes.

These parameters are similar to those pursued by researchers using SRA in the building or residential construction industries in that they describe the size and complexity of the project and are known early in the design process. Comments from industry contacts confirm that many of these parameters are critical to the construction duration. Therefore, the availability of data fields mentioned above will serve as the basis for identifying and selecting a potential data source.

3.2.3 Identification and Selection of a Historical Project Data Source

The highway construction owner is in the best position to provide the two distinct types of data required for this research. The individual SHA is the most common highway construction owner, requiring large annual construction volume and having responsibilities over large geographical areas. Also, information collected and maintained is largely public and not proprietary.

Most state highway agencies regularly maintain contract level data. This information is pertinent to agency officials and serves as the basis for high-level project monitoring. Conversely, few SHAs maintain parametric level project data needed. Of those that do, Texas and Virginia stand out among the existing literature and maintain both the contractual and project parametric data such as that required by this study [Chou 2005].

A conceptual cost estimating study [Chou 2005] exhibits Texas' parametric level data storage. Chou [2005] analyzed the relationship between estimated project cost and preliminary (or parametric) project factors using Texas Department of Transportation project data. The relationship was established indirectly by first using project geometry factors such as length, width, number of lanes, should width, location, and traffic volumes to develop a regression model for predicting earthwork, landscaping, sub-grade, pavement, and traffic control quantities [Chou 2005].

The Virginia Department of Transportation (VDOT) maintains parametric level data in the *Project Cost Estimating System (PCES)*. *PCES* (see Figure 3.1) uses user-input cost items, as well as input project parameter quantities (x_i, x_{i+1}, x_{i+2}, …, x_n) and historically-based cost coefficient values (β_i, β_{i+1}, β_{i+2}, …, β_n) to estimate construction cost [Kyte et al. 2004]. Estimators use *PCES* as project design proceeds from scoping through the field inspection stages of project development.

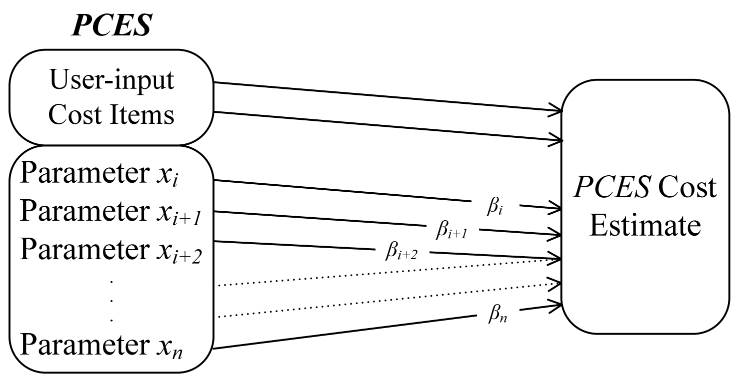

Figure 3.1 – *PCES* Cost Estimating

VDOT is an optimal data supplier option for a number of reasons. VDOT has a long-standing relationship with the academic community and has demonstrated a strong desire to improve practices and procedures through the development of construction knowledge. VDOT has expressed a significant interest in the research, as well as a willingness to participate. Finally, the proximity of VDOT to the research makes them the optimal candidate for data supply.

3.2.4 Data Collection and Compilation

Data necessary to perform a statistical analysis and prepare a set of mathematical models to estimate construction time is extracted from two sources within VDOT: *Data Warehouse Management Information Portal* (*DWMIP*) and *PCES*. *DWMIP* serves as the data source for high-level project and contract information while *PCES* provides actual project parameter values.

Figure 3.2 shows how a system using both *PCES* and *DWMIP* works. The same parameter quantities used in *PCES* are used in the construction duration estimating models. However, time-coefficient values (α_i, α_{i+1}, α_{i+2}, …, α_n) are quantified, applied, and incorporated into the SRA model. High-level, contract parameters in *DWMIP* (y_i, y_{i+1}, y_{i+2}, …, y_n) are also incorporated into the construction duration estimating model using the quantified coefficients (θ_i, θ_{i+1}, θ_{i+2}, …, θ_n). These parameters are known early in the project development lifecycle, and are thought to be influential on the overall construction duration.

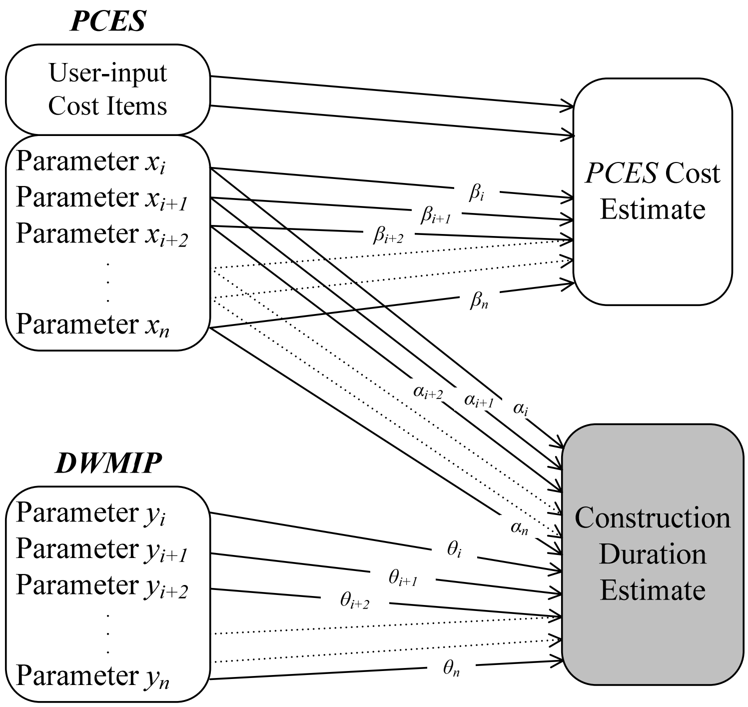

Figure 3.2 – *DWMIP* & *PCES* Data Interaction

3.2.4.1 DWMIP Data Collection

DWMIP is the mainframe contract data management system within VDOT. This system maintains administrative contract and project data that is used for numerous functions. Data from *PPMS* (Program/Project Management System), *TRNS*PORT* (an AASHTO cost estimating software suite), and *RUMS* (Right-of-Way and Utilities Management System) can be retrieved through *DWMIP*.

In order to obtain an adequate population of project data and to ensure that the data analyzed reflects current construction conditions, *DWMIP* was queried for all construction projects, funding types, roadway systems, locations, and budget conditions for projects awarded <u>and</u> completed between January 1, 2000 and January 1, 2007. This query produced 2,225 contract records. However, each record represents an individual project, not contract. Often, a single contract will be broken down into numerous projects. Upon further analysis, 1,506 unique contracts were identified. While there are 39 data fields available, many fields are repetitive or unrelated to duration. Table 3.1 contains a list of the data fields from *DWMIP* that are pertinent to construction duration.

Table 3.1 – *DWMIP* Data Fields

• Uniform Project Code (UPC)	• Contract Acceptance Date
• District	• Highway Description (State, Primary, Federal, etc.)
• Scope of Work (Work Type)	• Contract Setting (Urban, Rural, etc.)
• Contract Execution Date	

3.2.4.2 PCES Data Collection

PCES is a cost estimating spreadsheet tool that makes use of project details that become available during the initial design phase. These details are continually refined throughout the design process, until a more detailed estimate is initiated at 75% design completion. VDOT does not use *PCES* to estimate costs of every project advertised. Instead, *PCES* estimates are prepared for approximately 100 projects each year. When a cost estimate is prepared in *PCES*, the project factors and quantities entered by the estimator are stored within the system.

It is expected that many of the same parameters used by *PCES* to estimate construction cost also influence construction duration. Therefore, *PCES* parameters will be analyzed as potential explanatory variables in the regression analysis process. Table 3.2 shows the *PCES* parameters that are expected to influence the construction duration.

Table 3.2 – *PCES* Parameters

• Geometric Design Standard	• Annual Daily Traffic	• Project Length
• Two-lane Length (mi)	• Four-lane Length (mi)	• Crossover Count
• Bridge Type	• Bridge Length (ft)	• Bridge Width (ft)
• Shoulder or Curb & Gutter	• Length of Curb & Gutter (ft)	• Length of Loops / Ramps (mi)
• Median Type	• Median Length (ft)	• Number of New Signals
• ROW Length (ft)	• ROW Width (ft)	• % ROW Agriculture Use
• % ROW Commercial Use	• % ROW Industrial Use	• % ROW Residential Use

As noted in Figure 3.2, the construction duration estimate does not make use of the construction cost estimate. While this seems contrary to earlier authors who evidenced a significant relationship between construction cost and duration [Bromilow 1969, Chan and Chan 2004, Jiang and Hongbo 2007], it is not possible here to use the construction cost estimate to help explain the development of construction duration. At the initial project design development phase, VDOT depends on the *PCES* system to provide a reliable and accurate cost estimate based on the parameters available at the time, known to influence cost. It is too early to predict the engineer's construction estimate, much less the contract award amount. Therefore, the *PCES* cost estimate is the only construction cost estimate available at this stage of project design. As seen in Figure 3.1, the *PCES* estimate is derived from a set of project design parameters and historically-based coefficients which depict the relationship between the parameter quantity value and construction cost. Therefore, the *PCES* cost estimate is a summation of the individual relationships between project parameters and construction cost. Using both the *PCES* parameter quantities <u>and</u> the *PCES* cost estimate would likely induce multicollinearity issues between the regressors [Ott and Longnecker 2001]. Instead, the *PCES* parameters are analyzed in the SRA. Duration coefficients explain the relationship between these *PCES* parameters and their influence on construction duration. The final model incorporates both *DWMIP* and *PCES* parameter coefficients to explain the development of construction duration.

There are 3,960 estimate entries in *PCES*. These entries have been compiled from multiple data sources from data collected over the past 50 years. Therefore, *PCES* contains many incomplete and superfluous data entries. More recent *PCES* data are inputs into the *PCES* system, and therefore, more reliable as the *PCES* system has been in common use across Virginia over the past several years.

The UPC is the unique identifier for each record, found in both *DWMIP* and *PCES*. A 'select query' is used to compile a complete set of regression data containing both contract and project level information. Two limiting criterion exists in the *DWMIP* and *PCES* data sets. *DWMIP* data was collected from projects which have been advertised and completed between January 1, 2000 and January 1, 2007. Also, *PCES* is only used on approximately 100 projects per year, and the most reliable data is that entered in the past several years. With these limitations, the data compilation process still yields 331 unique records having both *DWMIP* and *PCES* data (see Figure 3.3).

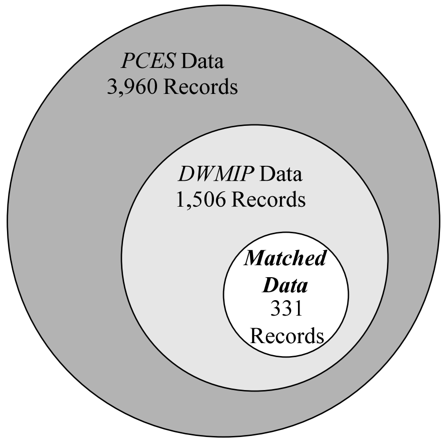

Figure 3.3 – Data Compilation

3.3 Data Treatment

The data treatment process prepares the data from the 331 projects collected for regression analysis. The data treatment process involves data pretreatment and seasonal adjustment. These processes will ensure the data acquired is suitable for SRA.

3.3.1 Data Pretreatment

Once compiled, the data is preliminarily analyzed to ensure the records are complete and accurate. An inherent limitation of retrospective data analysis is that it is often difficult, if not impossible, to guarantee accuracy for every record. Therefore, the pretreatment process focuses on identifying topical errors and ambiguities, correcting data records where possible and justifiable, and removing clearly erroneous data entries that would unfairly bias the analysis. The following paragraphs describe the processes of identifying the estimate entry, initial parameters analysis, and paring of the data.

3.3.1.1 Multiple Estimates

The 331 records referenced represent unique project records. As *PCES* is used throughout initial design as details become available, there are multiple cost estimate entries for each of the 331 records. By sorting the data set by UPC, then estimate date, each set of estimates is identified. In reviewing the data, it is clear that the most recent (nearest to advertisement date) estimate is the most complete record. While such an estimate would not take place until the latter phases of initial design, the information collected is necessary for regression analysis and to establish an accurate understanding of construction duration. Therefore, the most recent estimate record is extracted for each project in the data set to produce 331 unique records with both *DWMIP* and *PCES* data.

3.3.1.2 Erroneous or Incomplete Records

While the most recent estimates are typically the most complete, there are instances in which data fields have been left blank or appear to contradict earlier quantity measurements. This occurred in 4 of the 331 projects. Analysis of three specific projects revealed that the most recent estimate from which data was selected was performed immediately after the previous estimate. These estimates had no quantity entries for any field. Therefore, it is expected that these records were generated by opening the system (*PCES*), beginning an estimate, and then closing the system without completing the estimate. This would generate a record in the *PCES* data records without pertinent project information. In these instances, the next most recent estimate substituted for the record of choice.

Another, less common, occurrence is ambiguity between the most recent estimate quantities and earlier used estimate quantities. However, because the cost and duration estimates are an iterative and progressive process, the most recent parameter quantity values should be the most accurate. Further, the cost estimate data and contract level data entries are based on the most recent, detailed estimate. Therefore, the data fields of all 331 records were checked to ensure that the most recent quantity was used and that these quantities were in general agreement with earlier used quantities. This ensures that the most recently used quantity was not erroneously entered.

3.3.1.3 Data Focusing

Road and bridge construction involve different operations, materials, and components that impact construction duration. As such, SRA to identify and model construction duration influential project factors and parameters must be specific to, at least, the type of work performed: road or bridge construction. Initial data analysis reveals that, of the 331 records, 23 are related to bridges. These projects vary from bridge replacement (19 projects), to bridge rehabilitation (3 projects), to new bridges (1 project). Among the guidelines for SRA, is a minimum sample size of 25 records [Ott and Longnecker 2001]. While SRA has been employed on fewer than 23 data records in the past, it is felt that so few data points could not adequately seek to explain the relationship between early-known project parameters and bridge construction duration.

The data set is further focused by removing projects with ambiguous work descriptions (such as "Environmentally Related" or "Demolition"). Finally, those project records without a specified work description are removed from the data set. After this focusing, the data set contains 294 individual project records containing both *PCES* and *DWMIP* parameter records.

3.3.2 Seasonal Duration Adjustment

Construction time is the duration required for construction of all work items within the contract. This duration is typically measured in working-days. Meanwhile, contract time is the duration required for completion of the contract. This duration includes non-working days such as weekends, holidays, and adverse weather days and is usually measured in calendar-days and defined with a fixed completion date. In the data sampled, this duration is established using the 'Contract Execution Date' and the 'Contract Acceptance Date'. Subtracting the acceptance date from the execution date yields calendar-day contract duration.

Due to the nature of the durations, contract duration is a more subjective estimate than the construction duration. The allowances included in contract duration estimates often include variability from a number of sources. The contract duration estimation process is very involved. While project details may be nearly identical, two scheduling professionals would likely arrive at different durations. Further, there is an amount of variability existing in the construction startup process. While agencies have moved to a fixed completion date contract, there is not a fixed commencement date on which the contractor must begin work.

Further, bias enters into the construction duration by the time of year in which work begins or is completed. According to Jervis and Levin [1987], weather is "the most common cause of excusable delay." Virginia, like much of the United States, experiences seasonal changes throughout the year that bring differing weather patterns and events. These events can delay construction activities, and prolong the contract duration. NCHRP [1978] states that 45% of construction activities are subject to delay due to adverse weather. In planning for such delays, McDonald [2000] recommends consideration of the project type, location, calendar months, and phase of construction.

The impact of adverse weather is different for each project and its specific details. Therefore, it is unfair to compare contract durations from projects of varying start and end dates, types, locations, and durations without considering the weather events that may have impacted said duration.

While both bias and variability are detrimental to the SRA process, it is felt that the bias induced due to adverse weather can and should be removed from the contract duration. This converts the contract duration to a seasonally adjusted construction duration for regression analysis.

3.3.2.1 Seasonal Adjustment Options

A retrospective analysis involves using historical data without the possibility to influence the data collection and maintenance processes. Often, the analyst is presented with limitations that must be addressed. Data specificity, feasibility of collection, and feasibility of eventual use must be considered. This study presents four seasonal adjustment options (see Table 3.3) ranging from generic to specific. First, data can be segregated and examined as individual projects or generalized into project or work types. Project-by-Project examination requires studying individual project records to obtain weather events and impacts. Examination by work type groups projects of a similar work type and quantifies weather impacts for each type of work, rather than for individual projects. Second, weather data may be examined on a day-to-day basis for each project or by monthly average event occurrence. Daily weather occurrences are specific to individual project locations and environments. Monthly average occurrences are a historical average of weather event occurrences in each location or group of locations (i.e. district or region).

Table 3.3 – Seasonal Adjustment Options

	Daily Occurrence	Monthly Average
Project-by-Project	**I** Manual, High Accuracy, Project & Weather Record Intensive	**II** Moderate Accuracy, Project Record Intensive
Work Type	**III** Moderate Accuracy, Weather Record Intensive, Generalized Work-type Impacts	**IV** Conceptual-level Accuracy, Automatable, Generalized Weather & Work-type Impacts

The flow of the matrix from quadrant one through four proceeds from most accurate, labor intensive, and manual (Quadrant I) to the more generic, least labor intensive, and readily automated (Quadrant IV). Each quadrant is described here, followed by an overview of the attributes for each quadrant.

Quadrant I involves a manual examination of each project, from beginning to end, and the daily occurrence of weather and the impacts of those weather events. Daily historical weather data (see Figures 3.4, 3.5, and 3.6) is available from the National Oceanographic and Atmospheric Association (NOAA) [NOAA 2004]. Also, daily project records or diaries for each project are maintained by SHAs. This quadrant is the most detailed and accurate method available. This accuracy is due to the amount of data that would be examined in order to extract individual weather events from individual projects. This process is labor intensive, manual, and time consuming. Such an adjustment assumes the need for and ability to forecast daily weather occurrences over the life of the construction project.

Quadrant II involves the examination of each project, from beginning to end, and the average monthly occurrence of adverse weather and the influences thereof. Again, daily project diaries are available to obtain project specific information. Detailed weather data would then be summarized by the average number of adverse weather occurrences per month (ex: the number of days each month on which ½" of rainfall occurred). The adjustment process in quadrant II is less detailed than that of quadrant I but still requires a large amount of project data. The reduction in time and labor required by averaging weather events by month is more than offset by the need to examine each project and the daily event records. Ultimately, a less specific weather calendar would be applied to a very specific location and condition.

Quadrant III involves the summation of projects by work type to generalize the impacts of weather on different types of work. After grouped into work types, the daily occurrence of weather events would be analyzed using historical daily diaries. Quadrant III is less labor intensive than quadrants I and II because it does not involve manual analysis of individual project records. Instead, the impacts of weather events on various types of work would be developed from a literature review of current practice. These impacts would be applied for all projects of that type. However, in quadrant III, weather events would be analyzed daily and the generalized impacts applied for each individual weather event.

Finally, quadrant IV involves the generalization of weather impacts on various types of work and summation of historically average monthly weather event days. This process is the least manual and labor intensive of the four methods. Essentially, this process develops a weather calendar for each type of work and month during which work is performed. These weather calendars can be specified for each district using NOAA data. The calendars are then applied to projects on a work type and location (district / region) basis. The application of these calendars is easily automated.

The seasonal adjustment method described in quadrant IV has been selected. Collection and examination of individual project records and daily weather events is unnecessary. Such detailed information is beyond the level of detail available or needed to develop a conceptual duration estimate. Instead, generalized weather impact factors and historically average monthly weather occurrences can be more readily defined from existing literature and historical data. Finally, the ability to automate this process also makes it an excellent choice for adjusting a large historical project data set.

U.S. Department of Commerce
National Oceanic & Atmospheric Administration
National Environmental Satellite, Data,
and Information Service

Climatography of the United States No. 20 1971-2000

National Climatic Data Center
Federal Building
151 Patton Avenue
Asheville, North Carolina 28801
www.ncdc.noaa.gov

Station: BLACKSBURG, VA **COOP ID: 440766**

Climate Division: VA 6 **NWS Call Sign: RNK** **Elevation: 2,100 Feet** **Lat: 37°12N** **Lon: 80°25W**

Temperature (°F)																						
	Mean (1)			Extremes										Degree Days (1) Base Temp 65		Mean Number of Days (3)						
Month	Daily Max	Daily Min	Mean	Highest Daily(2)	Year	Day	Highest Month(1) Mean	Year	Lowest Daily(2)	Year	Day	Lowest Month(1) Mean	Year	Heating	Cooling	Max >= 100	Max >= 90	Max >= 50	Max <= 32	Min <= 32	Min <= 0
Jan	41.1	20.6	30.9	73	1975	30	41.6	1974	-18	1985	21	18.9	1977	1059	0	.0	.0	7.9	7.1	27.6	1.4
Feb	44.8	22.1	33.5	75	1977	27	41.5	1976	-12	1970	4	23.0	1978	883	0	.0	.0	10.9	4.8	23.6	.4
Mar	53.3	29.4	41.4	85	1985	31	48.0	1976	2	1969	11	35.3	1996	733	0	.0	.0	20.0	1.3	19.0	.0
Apr	62.9	37.0	50.0	95+	1957	29	55.0	1994	14	1982	7	46.2	1997	452	0	.0	.0	26.3	.1	9.1	.0
May	71.5	46.3	58.9	89+	1996	20	64.3	1991	23	1973	5	53.6	1973	219	31	.0	.0	30.6	.0	1.5	.0
Jun	78.6	55.2	66.9	94	1988	27	70.7	1994	30	1972	11	61.6	1974	47	104	.0	.4	30.0	.0	@	.0
Jul	82.5	59.7	71.1	99	1954	15	75.4	1993	41+	1972	7	68.0	1984	7	196	.0	3.0	31.0	.0	.0	.0
Aug	81.3	57.8	69.6	99	1983	23	73.0	1975	36	1986	29	67.2	1985	11	151	.0	2.1	31.0	.0	.0	.0
Sep	75.3	50.7	63.0	96	1953	2	67.8	1978	22	1983	24	59.6	1976	106	46	.0	.6	29.9	.0	.3	.0
Oct	65.3	38.0	51.7	88	1954	6	59.7	1984	13	1962	27	44.4	1988	418	5	.0	.0	29.3	.0	8.9	.0
Nov	55.0	30.6	42.8	79	1961	6	49.0	1985	0	1970	25	35.9	1976	665	0	.0	.0	20.5	.5	18.2	.0
Dec	44.9	23.2	34.1	75+	2001	6	43.4	1984	-10	1989	23	24.0	1989	959	0	.0	.0	11.1	4.7	26.0	.5
Ann	63.0	39.2	51.2	99+	Aug 1983	23	75.4	Jul 1993	-18	Jan 1985	21	18.9	Jan 1977	5559	533	.0	6.1	278.5	18.5	134.2	2.3

+ Also occurred on an earlier date(s)

@ Denotes mean number of days greater than 0 but less than .05

Complete documentation available from: www.ncdc.noaa.gov/oa/climate/normals/usnormals.html

Issue Date: February 2004 008-A

(1) From the 1971-2000 Monthly Normals

(2) Derived from station's available digital record: 1952-2001

(3) Derived from 1971-2000 serially complete daily data

Figure 3.4 – NOAA Temperature Data for Blacksburg, VA [NOAA 2004]*

*From: National Oceanic and Atmospheric Administration (NOAA). (2004). Climatography of the United Sates No. 20. Virginia Monthly Station Climate Summaries, 1971-2000. Washington, D.C.

U.S. Department of Commerce
National Oceanic & Atmospheric Administration
National Environmental Satellite, Data,
and Information Service

Climatography of the United States No. 20 1971-2000

National Climatic Data Center
Federal Building
151 Patton Avenue
Asheville, North Carolina 28801
www.ncdc.noaa.gov

Station: BLACKSBURG, VA **COOP ID: 440766**

Climate Division: VA 6 **NWS Call Sign: RNK** **Elevation: 2,100 Feet** **Lat: 37°12N** **Lon: 80°25W**

Precipitation (inches)																								
	Precipitation Totals									Mean Number of Days (3)				Precipitation Probabilities (1) Probability that the monthly/annual precipitation will be equal to or less than the indicated amount										
	Means/ Medians(1)		Extremes							Daily Precipitation				Monthly/Annual Precipitation vs Probability Levels These values were determined from the incomplete gamma distribution										
Month	Mean	Med-ian	Highest Daily(2)	Year	Day	Highest Monthly(1)	Year	Lowest Monthly(1)	Year	>= 0.01	>= 0.10	>= 0.50	>= 1.00	.05	.10	.20	.30	.40	.50	.60	.70	.80	.90	.95
Jan	3.37	3.15	2.26	1992	4	7.41	1998	.60	1981	11.0	7.2	2.4	.7	1.13	1.45	1.92	2.32	2.69	3.09	3.51	4.01	4.65	5.64	6.56
Feb	3.02	2.81	2.32	1994	11	6.28	1982	.38	1980	10.3	7.0	2.2	.3	.92	1.21	1.64	2.01	2.36	2.73	3.14	3.62	4.24	5.21	6.10
Mar	3.83	3.32	2.00	1975	14	8.33	1975	1.33	1988	11.2	8.1	3.0	.7	1.48	1.84	2.34	2.76	3.16	3.57	4.01	4.52	5.17	6.16	7.07
Apr	3.83	3.58	3.42	1992	22	7.75	1987	1.21	1976	11.5	7.7	2.2	.8	1.40	1.76	2.28	2.71	3.13	3.55	4.01	4.54	5.22	6.27	7.23
May	4.39	4.38	2.92	1975	26	9.75	1971	1.41	1994	12.9	9.0	3.3	.9	1.57	1.98	2.58	3.08	3.56	4.05	4.59	5.21	6.01	7.23	8.36
Jun	3.93	3.64	3.80	1976	17	8.66	1976	.91	1990	10.6	7.5	2.5	1.0	.96	1.33	1.90	2.41	2.92	3.45	4.05	4.76	5.68	7.15	8.53
Jul	4.17	3.97	3.33	1996	9	9.35	1992	1.56	1998	12.7	8.9	3.0	.7	1.74	2.12	2.65	3.09	3.50	3.92	4.37	4.89	5.55	6.55	7.46
Aug	3.68	3.50	3.44	1992	27	7.46	1985	1.64	1987	10.2	6.6	2.6	.8	1.63	1.96	2.42	2.79	3.14	3.49	3.86	4.29	4.84	5.66	6.40
Sep	3.39	3.36	3.42	1959	30	8.09	1989	.49	1991	9.9	6.2	2.7	.9	.62	.92	1.42	1.88	2.35	2.85	3.43	4.13	5.05	6.54	7.96
Oct	3.19	2.92	2.94	1966	19	9.14	1976	.02	2000	8.6	6.0	2.4	.7	.38	.63	1.07	1.52	1.99	2.52	3.14	3.90	4.93	6.64	8.29
Nov	2.96	2.71	2.55+	1989	16	7.41	1977	.49	1981	9.4	6.3	2.1	.5	.82	1.10	1.52	1.89	2.26	2.64	3.06	3.56	4.21	5.23	6.18
Dec	2.87	2.59	3.18	1958	29	6.03	1973	.32	1989	10.3	6.7	2.0	.3	.70	.97	1.39	1.76	2.13	2.52	2.96	3.48	4.16	5.24	6.26
Ann	42.63	42.35	3.80	Jun 1976	17	9.75	May 1971	.02	Oct 2000	128.6	87.2	30.4	8.3	32.47	34.49	37.05	38.97	40.67	42.30	43.97	45.80	48.02	51.20	53.93

+ Also occurred on an earlier date(s)
# Denotes amounts of a trace
@ Denotes mean number of days greater than 0 but less than .05
** Statistics not computed because less than six years out of thirty had measurable precipitation

(1) From the 1971-2000 Monthly Normals
(2) Derived from station's available digital record: 1952-2001
(3) Derived from 1971-2000 serially complete daily data
Complete documentation available from:
www.ncdc.noaa.gov/oa/climate/normals/usnormals.html

Figure 3.5 – NOAA Precipitation Data for Blacksburg, VA [NOAA 2004]*

*From: National Oceanic and Atmospheric Administration (NOAA). (2004). Climatography of the United Sates No. 20. Virginia Monthly Station Climate Summaries, 1971-2000. Washington, D.C.

U.S. Department of Commerce
National Oceanic & Atmospheric Administration
National Environmental Satellite, Data,
and Information Services

Climatography of the United States No. 20 1971-2000

National Climatic Data Center
Federal Building
151 Patton Avenue
Asheville, North Carolina 28801
www.ncdc.noaa.gov

Station: BLACKSBURG, VA **COOP ID: 440766**

Climate Division: VA 6 **NWS Call Sign: RNK** **Elevation: 2,100 Feet** **Lat: 37°12N** **Lon: 80°25W**

Snow (inches)																							
Snow Totals															Mean Number of Days (1)								
Means/Medians (1)					Extremes (2)										Snow Fall >= Thresholds					Snow Depth >= Thresholds			
Month	Snow Fall Mean	Snow Fall Median	Snow Depth Mean	Snow Depth Median	Highest Daily Snow Fall	Year	Day	Highest Monthly Snow Fall	Year	Highest Daily Snow Depth	Year	Day	Highest Monthly Mean Snow Depth	Year	0.1	1.0	3.0	5.0	10.0	1	3	5	10
Jan	7.2	4.0	1	0	20.4	1998	28	39.2	1996	27	1996	8	6	1996	2.9	1.9	.7	.3	.1	7.8	4.0	2.4	.8
Feb	5.7	4.0	1	1	8.4	1996	2	21.5	1996	15	1982	28	4	1983	2.8	1.6	.7	.5	.0	7.1	3.2	1.6	.4
Mar	3.8	1.2	#	0	13.0	1981	23	18.0	1993	18+	1993	15	3	1993	1.3	.9	.4	.2	.1	2.2	1.1	.7	.2
Apr	1.1	.0	#	0	10.0	1971	7	10.0	1971	10	1971	7	#	1993	.5	.3	.1	.1	@	.4	.2	.1	@
May	#	.0	#	0	#	1989	7	#	1989	#	1987	2	#	2000	.0	.0	.0	.0	.0	.0	.0	.0	.0
Jun	.0	.0	0	0	.0	0	0	.0	0	0	0	0	0	0	.0	.0	.0	.0	.0	.0	.0	.0	.0
Jul	.0	.0	0	0	.0	0	0	.0	0	0	0	0	0	0	.0	.0	.0	.0	.0	.0	.0	.0	.0
Aug	.0	.0	0	0	.0	0	0	.0	0	0	0	0	0	0	.0	.0	.0	.0	.0	.0	.0	.0	.0
Sep	.0	.0	0	0	.0	0	0	.0	0	0	0	0	0	0	.0	.0	.0	.0	.0	.0	.0	.0	.0
Oct	#	.0	0	0	#	1999	25	#+	1999	#+	1993	31	0	0	.0	.0	.0	.0	.0	.0	.0	.0	.0
Nov	.7	.0	#	0	3.2	1995	14	6.0	1971	6	1971	27	1	1971	.6	.3	.1	.0	.0	.7	.1	.1	.0
Dec	3.6	1.3	#	0	11.2	1997	30	16.4	1997	10+	1997	30	2	1989	1.5	1.1	.4	.2	.1	4.1	1.9	.7	.1
Ann	22.1	10.5	N/A	N/A	20.4	Jan 1998	28	39.2	Jan 1996	27	Jan 1996	8	6	Jan 1996	9.6	6.1	2.4	1.3	.3	22.3	10.5	5.6	1.5

+ Also occurred on an earlier date(s) #Denotes trace amounts

@ Denotes mean number of days greater than 0 but less than .05

-9/-9.9 represents missing values
Annual statistics for Mean/Median snow depths are not appropriate

(1) Derived from Snow Climatology and 1971-2000 daily data

(2) Derived from 1971-2000 daily data

Complete documentation available from:
www.ncdc.noaa.gov/oa/climate/normals/usnormals.html

Figure 3.6 – NOAA Snowfall Data for Blacksburg, VA [NOAA 2004]*

*From: National Oceanic and Atmospheric Administration (NOAA). (2004). Climatography of the United Sates No. 20. Virginia Monthly Station Climate Summaries, 1971-2000. Washington, D.C.

3.3.2.2 Seasonal Adjustment Overview

The contract duration collected is a calendar-day duration that includes both working and non-working days. Because the seasonal adjustment option selected (Quadrant IV) does not rely on individual daily project records, the actual days that were not worked are not known for each project. These non-working days could have occurred due to the occurrence of a weekend, holiday, adverse weather, or contractor prerogative. The seasonal adjustment process estimates the typical number of non-working days, based on historical weather data and the impacts of adverse weather occurrences on construction. Once the historically typically working and non-working days are developed for each month, the Seasonally Adjusted Construction Duration may be totaled over the length of the project.

3.3.2.3 Historical Weather Impacts and Data Collection

Temperature, wind, snow, and precipitation can all negatively impact the project duration, and cause substantial delays to project completion. In considering the adverse effects of weather on duration, the scheduler considers historic weather patterns to forecast the number of working days for the project. Historical data is available for numerous weather events including temperature, precipitation, and snowfall [NOAA 2004]. A review of the literature reveals guidance regarding the threshold event amounts that impact construction operations.

NCHRP [1978] established that there is little effect of wind on horizontal highway construction work. Therefore, the effects of wind are not considered in this study.

The effects of temperature on worker productivity were documented by Koehn and Brown [1985] who suggested work taking place at temperatures above 100 degrees and below 20 degrees is significantly impacted. Unfortunately, 20 degrees is not a threshold data collection point for NOAA [2004]. Instead, 100 and 32º F are used as temperature extremes that impact the work.

NOAA [2004] provides precipitation data for 0.01”, 0.1”, 0.5”, and 1.0”. The significance of 0.1” was established by a number of authors previously [Smith and Hancher 1989, Finke 1990]. El-Rayes and Moselhi [2001] used 0.5” as a transition point.

This research will also use 0.5”. Finally, a 1.0” rainfall event in a day is seen as a significant event in Virginia. Therefore, 1.0” rainfall events will also be considered through this work.

Little guidance is offered on the effects of snowfall on construction. Many construction companies significantly reduce workloads and workforces during winter months. However, highway construction work can still proceed in many instances. Therefore, winter weather, in the form of snow, is considered in this study. NOAA [2004] provides snowfall depths at 0.1”, 1.0”, 3.0”, 5.0”, and 10”. Using a 10:1 snow to rain ratio [Roebber et al. 2003], the 1”, 5”, and 10” events coincide with the amount of precipitation necessary to produce 0.1”, 0.5”, and 1.0” rainfall events, respectively. Therefore, this study will consider these three snow events as significant thresholds.

Weather data from 68 stations was collected to account for climatic differences across the state of Virginia. These weather stations were then classified by their respective VDOT district. There is an average of 8 weather stations in each district, offering coverage and collection of weather across the entire district. Monthly weather quantities are totaled by station and averaged by month across the district. This will minimize the effects of extreme or rare weather events or district outliers and facilitate the use of the district as the lowest level of location identification.

3.3.2.4 Weather Impact Factors

Aforementioned authors documented the need to consider the type of work being performed, the frequency, and the severity of weather events in determining their impacts on construction [NCHRP 1978, Koehn and Brown 1985, Finke 1990, Smith and Hancher 1989, McDonald 2000, El-Rayes and Mosehli 2001]. Despite its importance, there is little guidance offered as to what constitutes a weather event or the impacts of such events.

An impact factor is used to describe the effect (in number of days) of a weather event on construction operations. For example, an impact factor of 0.5 represents the loss of one half day of work due to weather. For severe weather events, impact factors may be greater than one. For instance, an impact factor of 2.0 represents the loss of the day of occurrence and the following day due to the residual effects of the weather event. Here, a review of the literature reveals existing thresholds for various weather events.

The authors referenced are primarily focused on rainfall events. In that regard, the most commonly accepted rain event is one-tenth of an inch (0.1") in a day. Smith and Hancher [1989] define a day when less than 0.1" precipitation falls as a 'dry-day'. Finke [1990] states that "daily rainfall is considered to be meaningfully and reliably measurable when it equals or exceeds one-tenth of one inch." Finke [1990] goes further to suggest that rainfall greater than one-half inch to be 'moderately heavy', but does not define the impacts of such an event. Finally, El-Rayes and Moselhi [2001] offer threshold rain events that were used in the development of *WEATHER* – a decision support system for quantifying the impacts of rain on productivity and duration of highway operations. *WEATHER* quantifies rainfall impacts based on the type of operation, the amount of precipitation, the duration of the event, and the time of occurrence. Rainfall thresholds for earthmoving operations are 6-13 mm (0.25 – 0.5") and 13-25 mm (0.5 – 1.0"). To account for the impact of these events, El-Rayes and Moselhi [2001] use 0.5 – 1 day for 6-13mm of rain, and 1.5 – 2 days for 13-25mm of rain.

While the amount of rainfall for numerous areas can be easily quantified, the impact of rainfall events is much more difficult to ascertain. Likewise, there is little precedent for quantifying temperature and snow impacts. McDonald [2000] captures this uncertainty by saying that when deciding on compensable weather delays, one must use "a combination of arbitrary numbers, experience in the type of work being affected, and common sense."

Using the literature reviewed and engineering judgment, impact factors were developed for each significant weather event. Table 3.4 depicts the impact factors chosen for this study.

Table 3.4 – Weather Impact Factors

Contract Type	**Rain**			**Snow**			**Temperature**	
	> 1/10"	**> 1/2"**	**> 1"**	**> 1.0"**	**> 5.0"**	**> 10"**	**< 32º F**	**> 100º F**
Full-Depth Section	0.5	1.5	2	0.5	1.5	2	1	0.5
Highway Improvements	1	2	3	1	2	3	1	0.5

Impact factors are categorized first by work type. Two work types are shown in Table 3.4: Full-Depth Section and Highway Improvements. Full-Depth Section contracts generally include bulk movement of soil and are typically less susceptible to rain, snow, and temperature events than Highway Improvement contracts. If small rainfall or snow amounts (0.1" or 1.0" respectively) affect work at all, it is likely these effects would not cause the loss of an entire work day. Rather, site work may have to be rededicated, temporarily, to sealing off the borrow pile or recent lift, or to scarifying and drying the material placed. Larger rainfall and snow events (0.5 – 1.0" of rain or 5 – 10" of snow) have greater effects on site work. These events would saturate the soil and make site conditions unsuitable for work to proceed. Extreme temperature events impact Full-Depth Section and Highway Improvement work similarly. Temperatures below freezing (32º F) may prevent compaction efforts as well as make the site unsafe due to ice. Meanwhile, temperatures over 100º F would not have such a profound effect. While worker productivity would likely slow, such a high temperature would not be reached until after midday, or completion of 0.5 working days [Koehn and Brown 1985]. The occurrence of a day with greater than 100º F temperatures in Virginia is uncommon. Therefore, there is little affect on the yearly working days due to these events.

Because of the nature of work undertaken, Highway Improvement work types are more susceptible to weather events. The bulk of these projects include asphalt or concrete pavement activities which have stringent specifications that are influenced by even minimal amounts of rain, snow, and temperature. Whereas a small rainfall event on a Full-Depth Section project could cause redirected work activities, this same minimal rainfall event could potentially cause the contractor on a Highway Improvement project to cancel a full day of activities. Therefore, the impact factors associated with these work types are generally higher than those of Full-Depth Section contracts.

3.3.2.5 Working Day Calendar Development

The impact factor is multiplied by the average number of event occurrences to calculate the average adverse weather days by month for each station. Averaging adverse weather days for each district removes extreme outliers that can occur within a district.

Historical weather data is collected for every day of the year. Therefore, weekends are not extracted from the data and it must be assumed that some portion of these weather events happened during the weekend. Because state highway work is typically performed during the work week, with the exception of critical or emergency projects, it is assumed that weekends will be extracted from the contract duration as non-working days. Therefore, the number of days lost per month can be multiplied by 5/7 (the likelihood that one of the days lost falls on a weekday) to get a final number of days lost due to weather. Finally, holidays, which are scheduled to occur during the work-week, are removed, producing an average number of working days per month for each district and work type. Table 3.5 summarizes the work day calendars developed.

Table 3.5 – Average Monthly Working Days

Contract Type	Month	Statewide Average			
		Rain Days Lost	Snow Days Lost	Temp. Days Lost	Working Days
Full-Depth Sections	January	7	3	5	9
	February	6	3	3	10
	March	7	2	1	15
	April	6	1	1	15
	May	7	0	0	16
	June	6	0	1	16
	July	7	0	1	15
	August	6	0	1	17
	September	6	0	1	15
	October	6	0	0	16
	November	6	2	1	13
	December	6	2	3	13
	Total Working Days per Year:				**171**
Highway Improvements	January	10	4	5	7
	February	9	4	3	8
	March	10	3	1	11
	April	9	1	1	13
	May	11	0	0	13
	June	9	0	1	14
	July	11	0	1	12
	August	9	0	1	15
	September	9	0	1	13
	October	8	0	0	14
	November	8	2	1	11
	December	9	2	3	11
	Total Working Days per Year:				**143**

3.3.2.6 Calendar Application

The calendars developed are applied through a visual basic application that considers the project start and end dates. The application begins by prorating the potential days worked in the beginning and end months based on the start and end dates and the total calendar days in the month. The application then cycles through interim months keeping a count of the total potential work days that accrue during each month the project takes place. The output of the application is a seasonally adjusted, historically typical construction duration.

3.4 Summary

The need for additional tools and procedures for reliably estimating highway construction duration has been cited [Williams et al. 2008]. Also, the potential for statistical regression analysis (SRA) has been demonstrated in other construction industry sectors [Bromilow 1980, Nkado 1992, Chan and Kumaraswamy 1995, Chan 1999, Skitmore and Ng 2001, Burrows 2005]. This paper presents a review of the type of data needed for such a study, as well as the process involved with obtaining, compiling, and pre-treating the data.

The Virginia Department of Transportation (VDOT) is the selected data source. A compilation of historical data from their *DWMIP* and *PCES* systems provide 294 highway construction projects for statistical regression analysis. These records represent highway construction projects across Virginia, started and completed between January 1, 2000 and January 1, 2007.

Because construction duration is the dependent variable of SRA, the variability and bias must be addressed. While variability is specific to the schedule and cannot be adequately assessed, bias due to adverse weather is removed using historical event occurrences.

Adverse weather days are removed from the contract duration by developing historically typical weather calendars for Full-Depth Section and Highway Improvement work types using NOAA weather data and a review of the literature regarding weather impacts on construction. Consideration of weather impacts at this level of detail is inline with the early design stage at which the SRA model is applied. Finally, the construction duration is calculated using the historically typical weather calendars, contract execution and acceptance dates (contract duration), work type, and location.

3.5 References

Bromilow, F. J. (1969). "Contract time performance expectations and the reality." *Building Forum*, 1(3).

Burrows, T. K., Pegg, I., and Martin, J. (2005). "Predicting Building Construction Duration." *AACE International Transactions*, PS.14.1-PS.14-9.

Chan, A. P. C. (1999). "Modeling building durations in Hong Kong." *Construction Management and Economics*, 17, 189-196.

Chan, A. P. C., and Chan, D. W. M. (2004). "Developing a benchmark model for project construction time performance in Hong Kong." *Building and Environment*, 39, 339-349.

Chan, D. W. M., and Kumaraswamy, M., M. (1995). "A study of the factors affecting construction durations in Hong Kong." *Construction Management and Economics*, 13, 319-333.

Chou, J.-S. (2005). "Item-Level Quantity-Based Preliminary Cost Estimating System for Highway Earthwork, Landscape, Subgrade Treatments, Base, Surface Courses, Pavement and Traffic Control," University of Texas at Austin, Austin, TX.

El-Rayes, K., and Moselhi, O. (2001). "Impact of Rainfall on the Productivity of Highway Construction." *Journal of Construction Engineering and Management*, 127(2), 125-131

Finke, M. R. (1990). "Weather-Related Delays on Government Contracts." *Transactions of the American Association of Cost Engineers*, F.3.1-F.3.8.

Jervis, B. M., and Levin, P. (1987). *Construction Law - Principles and Practice*, McGraw-Hill, New York.

Jiang, Y., and Wu, H. (2007). "A Method for Highway Agency to Estimate Construction Durations and Set Contract Times." *International Journal of Construction Education and Research*, 3, 199-216.

Khosrowshahi, F., and Kaka, A. P. (1996). "Estimation of Project Total Cost and Duration for Housing Projects in the U.K." *Building and Environment*, 31(4), 373-383.

Koehn, E., and Brown, G. (1985). "Climatic Effects on Construction." *Journal of Construction Engineering and Management*, 111(2), 129-137.

Kumaraswamy, M., M., and Chan, D. W. M. (1995). "Determinants of construction duration." *Construction Management and Economics*, 13, 209-217.

Kyte, C. A., M. A. Perfater, et al. (2004). Developing and Validating a Highway Construction Project Cost Estimation Tool. Charlottesville, VA, Virginia Transportation Research Council.

McDonald, D. F. (2000). "Weather Delays and Impacts." *Construction Engineering*, 42(5), 34-39.

Moselhi, O., Gong, D., and El-Rayes, K. (1997). "Estimating Weather Impact on the Duration of Construction Activities." *Canadian Journal of Civil Engineering*, 24, 359-366.

National Cooperative Highway Research Program (NCHRP). (1978). "Effect of Weather on Highway Construction." *Synthesis of Highway Practice 47*, Transportation Research Board, Washington, D.C.

National Oceanic and Atmospheric Administration (NOAA). (2004). Climatography of the United Sates No. 20. Virginia Monthly Station Climate Summaries, 1971-2000. Washington, D.C.

Nkado, R. N. (1992). "Construction time information system for the building industry." *Construction Management and Economics*, 10, 489-509.

Orczyk, J. J. (1989). "Parametric Construction Scheduling," Dissertation, Purdue University, West Lafayette, Indiana.

Ott, R. L., and Longnecker, M. (2001). *An Introduction to Statistical Methods and Data Analysis*, Duxbury, Pacific Grove, CA.

Roebber, P. J., Bruening, S. L., Schultz, D. M., and Jr., J. V. C. (2003). "Improving Snowfall Forecasting by Diagnosing Snow Density." *Weather and Forecasting*, 18(2), 264-287.

Skitmore, R. M., and Ng, S. T. (2001). "Australian project time-cost analysis: statistical analysis of intertemporal trends." *Construction Management and Economics*, 19, 455-458.

Smith, G. R., and Hancher, D. E. (1989). "Estimating Precipitation Impacts for Scheduling." *Journal of Construction Engineering and Management*, 115(4), 552-566.

Williams, R. C., Vorster, M. C., and Hildreth, J. C. (2008). "The Need for Highway Construction Duration Estimating Using Early-known Design Details." *Transportation Research Board*, Submitted for review.

Chapter 4 – Statistical Regression Analysis to Identify and Model Relationships between Conceptual Design Parameters and Highway Construction Duration

4.1 Introduction

An accurate, reliable estimate of highway construction duration is needed for numerous design and financing decisions early in the project lifecycle. By the nature of design, however, the project details that influence the construction duration are not yet fully known until the project nears advertisement for bids. To compensate for this, the highway construction industry has relied on critical path schedules with multiple assumptions, comparison of the current project to previously completed projects, or the knowledge of experienced personnel [Williams et al. 2008a].

There is not currently an understanding of the early-known project factors that influence highway construction duration. Further, the relationship between these factors and highway construction duration is not known. Finally, there is not an understanding of how these duration-influential project factors combine to predict construction duration.

This document presents a statistical regression analysis study aimed at *1)* identifying the early-known construction duration-influential factors, *2)* empirically quantifying the coefficient or relationship of these factors with highway construction duration, and *3)* modeling the build-up of construction duration from these factors.

4.2 Study Background

This study relies on historical data from Virginia Department of Transportation (VDOT) projects awarded and completed between January 1, 2000 and January 1, 2007. It was not possible to control the project factors, for which data was collected, nor the levels or values of these variables. As a retrospective regression analysis, a limitation of this study is that cause and effect cannot be determined. This section provides background on linear and non-linear regression models and the reasons behind selecting multiple linear regression (MLR) as the analysis technique for this study.

4.2.1 Linear Regression Models

The purpose of regression analysis is to use data to determine empirical relationships and to use those relationships to predict an outcome [Kahane 2001]. The relationship described and trend discovered can take many different forms, including a simple or multiple linear form, or an exponential expression.

MLR models are proposed to explain the relationship between preliminary project factors and construction duration. The viability of linear models is stated by Montgomery et al. 2006:

> *"Linear regression models provide a rich and flexible framework that suits the needs of many analysts."*

Equation 4.1 shows a general MLR model [Montgomery et al. 2006].

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k + \varepsilon$$

Equation 4.1 – General Multiple Linear Regression Model

where: y = dependent variable (response),

β_0 = intercept,

$\beta_1 - \beta_k$ = model parameters (coefficients),

$x_1 - x_k$ = independent or explanatory variables (regressors), and

ε = error.

Using this general format, with additional explanatory variables, the use of transformations or exponential explanatory variables where necessary, most data trends can be modeled. As additional explanatory data becomes available and its relationship with the response variable determined, larger multivariate regression models can be formed.

MLR models are generally able to describe the data better than a simple linear model which is based on a single explanatory variable. However, it is usually the intent of the researcher to reduce the model to its most efficient form [Ott and Longnecker 2001]. Doing so creates a model that is more robust and readily applicable.

4.2.2 Nonlinear Regression Models

Nonlinear regression models are nonlinear in the regression coefficients [Montgomery et al. 2006]. These models should be selected only when knowledge of the underlying relationships between regressor and response variables are known to be nonlinear. Such knowledge does not exist in the highway construction duration estimating realm.

Nonlinear regression models also have a "strong sensitivity to outliers" and there are "fewer model validation tools for the detection of outliers in nonlinear regression than there are for linear regression" [NIST 2004]. Finally, MLR often provides a suitable approximation to complex, nonlinear relationships over limited ranges of explanatory variables [Myers et al. 2002]. Therefore, multiple linear regression models are used for this statistical analysis.

4.3 Data Collection

In preparation for statistical regression analysis, historical data has been collected from two VDOT sources: *Data Warehouse Management Information Portal* (*DWMIP*) and *Project Cost Estimating System* (*PCES*). Matching *DWMIP* to *PCES* data and extracting erroneous data points produces 294 projects for this study having both *DWMIP* and *PCES* data and falling within the aforementioned date guidelines.

4.3.1 *DWMIP* Data

DWMIP is the mainframe contract data management system within VDOT. This system maintains administrative contract and project data that is used for numerous functions. Data from *PPMS* (Program/Project Management System), *TRNS*PORT* (an AASHTO cost estimating software suite), and *RUMS* (Right-of-Way and Utilities Management System) can be retrieved through *DWMIP*. There are 39 data fields available in *DWMIP*, however, many fields are repetitive or unrelated to duration. The data fields from *DWMIP* pertinent to construction duration are listed below with a brief description:

- *District* – the project location; VDOT has 9 Districts.
- *Scope of Work* – the type of work being performed; used to categorize projects as "Full-Depth Section" or "Highway Improvement"
- *Contract Execution Date* – the date on which the contract (between VDOT and contractor) begins.
- *Contract Acceptance Date* – the date on which the construction work is accepted as complete and the contract (between VDOT and contractor) ends.

- *State Highway System* – the roadway system on which work is taking place.
- *Area Location Name* – the area or setting in which work is taking place (Rural, Small Urban, Medium Urban, etc.).

4.3.2 *PCES* Data

PCES is a cost estimating tool that makes use of project details that become available during the initial design phase [Kyte et al. 2004]. These details are refined throughout the design process, until a more detailed estimate is initiated at 75% design completion [Williams et al. 2008a]. *PCES* estimates are prepared for approximately 100 projects each year. When a cost estimate is prepared in *PCES*, the project factors and quantities entered by the estimator are stored within the database.

It is expected that many of the same variables used by *PCES* to estimate construction cost also influence construction duration. Therefore, *PCES* variables are analyzed as potential explanatory variables in the regression analysis process.

The *PCES* estimate begins as a historical average cost per mile ($/mile) [Kyte et al. 2004]. This value is then adjusted using VDOT developed adjustment factors for location, terrain, speed, and traffic volume which vary by geometric design standard. Other project factors, such as additional lanes and signals are additive factors. The data fields of interest from *PCES* are:

- *Geometric Design Standard* – type of roadway under construction.
- *Traffic Volume (ADT)* – average daily traffic count on the roadway.
- *Length of Loops/Ramps* – total length (in miles) of loops or ramps being constructed.
- *Curb & Gutter* – whether the roadway has curb & gutter or shoulder.
- *Length of Curb & Gutter* – total length (in miles) of curb & gutter being constructed.
- *Crossover Count* – number of crossovers being constructed.
- *Median* – whether the roadway is separated by a median.
- *New Signal Count* – number of new signal lamps installed.
- *Project Length* – total construction length (in miles) of the project.

Using the average cost per mile and the project length, the cost estimate is developed through a series of adjustments and additions to reach a “relatively reasonable” [Kyte et al. 2004] base estimate. A cost contingency (equal to 10% of base estimate) and construction engineering budget are added to the base estimate to form the construction cost estimate.

4.3.3 Data Application

Other authors have noted a significant relationship between construction cost and duration [Bromilow 1969, Chan and Chan 2004, Jiang and Hongbo 2007]. However, this study is primarily aimed at identifying the specific factors that influence construction duration. Because the *PCES* estimate is derived from a set of project design factors and historically-based coefficients which depict the relationship between the factor quantity value and construction cost, the *PCES* construction cost estimate is a summation of the individual relationships between project factors and construction cost. Using both the *PCES* factor quantities <u>and</u> the *PCES* cost estimate would likely induce multicollinearity issues between the regressors [Ott and Longnecker 2001]. Therefore, it is not desirable here to use both the *PCES* factors <u>and</u> the construction cost estimate to explain the development of construction duration. Instead, two models are analyzed and discussed on their individual merits: *1)* a condensed model which uses the *PCES* construction cost estimate as a regressor for predicting construction duration and *2)* a full model which incorporates statistically significant *PCES* factors (without the construction cost estimate) into a model for predicting construction duration (see Figures 4.1 and 4.2). Duration coefficients explain the relationship between these *PCES* factors and their influence on construction duration.

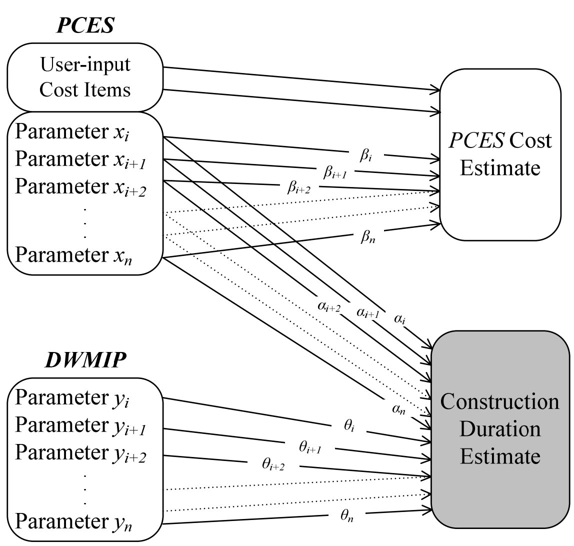

Figure 4.1 – Full Model Data Source

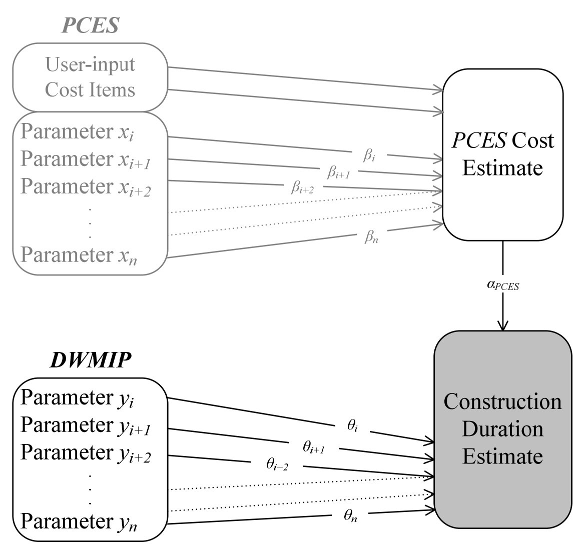

Figure 4.2 – Condensed Model Data Source

4.4 Data Preparation

The data preparation process includes seasonal adjustment, indicator variable establishment, and a discussion of the sample sizes for Full-Depth and Highway Improvement project regression analyses.

4.4.1 Seasonal Adjustment

The seasonal adjustment process converts the historical contract duration to historically-typical construction duration. Unlike the contract duration, the construction duration does not include non-working days attributed to weather, weekends, or holidays. Historically-typical work calendars for both Full-Depth Section and Highway Improvement projects were developed previously [Williams et al. 2008b]. Williams et al. [2008b] used National Oceanographic and Atmospheric Association (NOAA) precipitation and temperature data from Virginia weather stations 1971 - 2000 and a review of the existing literature to develop weather impact factors for various rain, snow, and temperature events. These impact factors were applied to the contract duration, removing weather related non-working days from the historical project record. Additionally, it was assumed that project work did not take place on the weekends or federal holidays. Therefore, these non-working days were also removed from the historic contract duration. This process was automated using *Microsoft Excel*™ and the historically-typical construction duration was calculated for each of the 294 project records collected. This historically-typical construction duration is used as the response variable (y) of the regression analysis.

4.4.2 Indicator Variables

Indicator variables are used to incorporate qualitative or categorical data into statistical regression analysis. A base value is selected and indicator variables for all other values of the field are assigned a binary value (0 or 1). This prevents an unwarranted hierarchical structure amongst regressor levels. When multiple qualitative levels exist, indicator variables are created to properly code the qualitative regressor value. The base value is generally selected as the value suspected to have the minimum construction duration. This expectation is tested through regression analysis.

In this study, several qualitative variables with multiple levels exist. Below is a list of the qualitative variables being considered in this study. The base indicator value and number of levels for each variable is denoted in the parentheses.

- District (Bristol, 9),
- Area Location Name (Rural, 4),
- State Highway System (Primary Arterial Network, 5),
- Geometric Design Standard (GS 1 – Rural Principal Arterial, 9),
- Curb & Gutter (Has Curb & Gutter, 2), and
- Median (Has Median, 2).

The number of indicator variables for each of these qualitative variables is equal to $k - 1$ (where k is the number of variable levels) [Montgomery et al. 2006]. The list above has 25 potential indicator variables created using these qualitative data fields ($\Sigma(k - 1)$ for the 6 variables).

4.4.3 Data Population

As noted, the 294 project records gathered are segregated into Full-Depth Section projects and Highway Improvement projects due to the nature of the work undertaken in each instance. Full-Depth Section projects include new construction, reconstruction, and widening. Highway Improvement projects include restoration and rehabilitation, resurfacing, and safety upgrade projects.

Of the 294 project records collected, 113 records have a scope of work that classifies as a 'Full-Depth Section' project. Inspection of these 113 projects, however, revealed a deficiency in the data collected. It appears that 38 of the 113 projects do not have a traffic volume entry. Further, these same deficient projects do not have a geometric design standard associated with them. As such, these projects will not be considered in the regression analysis. This reduces the Full-Depth Section population size to 75 projects.

Highway Improvement projects represent 181 of the 294 project records collected. As these are smaller projects, it is logical that they are more heavily represented in the data population. However, like the Full-Depth Section projects, many of the Highway Improvement project records are incomplete in the traffic volume and geometric design standard fields. Removal of those records without these two important data entries reduces the Highway Improvement project population size to 61 projects.

While authors have offered threshold sample sizes for regression analysis, there is not a universal rule on the required sample size [Stevens 2001]. Common recommendations include 10 or 15 records for every explanatory or regressor variable in the model [Stevens 2001]. Unfortunately, retrospective studies, such as this, are often limited by the availability of data. Therefore, it is thought that the 75 Full-Depth Section and 61 Highway Improvement project records represent a sufficiently large data set for conducting statistically regression analysis.

4.5 Regression Analysis

A traditional statistical regression analysis methodology is adopted and adapted to suit the needs of this research. The benefits of using statistical regression analysis for this type of research have been established [Williams et al. 2008b] and used in similar studies in other construction industry sectors. This section discusses the regression assumptions, variable selection processes, and model development processes used herein.

4.5.1 Regression Assumptions

The use of regression analysis as a modeling technique relies on the fulfillment of a number of assumptions [Montgomery et al. 2006]:

1. The relationship between the response y and the regressors is linear, at least approximately.
2. The error term ε has zero mean.
3. The error term ε has constant variance σ^2.
4. The errors are uncorrelated.
5. The errors are normally distributed.

These assumptions will be tested primarily with residual analysis throughout the regression analysis process.

4.5.2 Variable Analysis and Selection

This work develops four (4) regression models for predicting highway construction duration. These models have numerous quantitative and qualitative variables that explain the variation in construction duration. Therefore, the variable analysis and selection processes are described in generic terms here. The variables selected and the models developed are defined and described in section 4.5.3.

4.5.2.1 Variable Selection and Model Building

The potential of multiple linear regression analysis among modeling techniques was discussed previously. Despite the benefits of regression analysis, the process is based on empirical relationships in the data, and hence there rarely exists a single 'best' equation that outperforms all others. Instead, several plausible competing models usually exist, and the art of regression analysis is to use engineering judgment to select the optimal model for the purposes of the research.

There also exists, two basic modeling techniques: *1)* model and compare all possible regression equations or *2)* develop an optimal equation using a stepwise approach to evaluating regressors. Because of the large number of indicator variables and combinations therein, developing all possible models is not a viable alternative for this research. Therefore, a stepwise approach to regressor evaluation and modeling is selected.

4.5.2.2 Stepwise Analysis

Stepwise regression procedures generally fall into three categories [Montgomery et al. 2006]: *1)* forward selection, *2)* backward elimination, and *3)* stepwise regression. Stepwise regression combines both the forward selection and backward elimination processes in which, as each regressor is considered for addition to the model, its significance is evaluated in light of their inclusion. The backward elimination phase of stepwise regression looks to exclude redundant explanatory variables from the model by reevaluating the significance of the regressors in the model. The partial correlation, or the addition of explanatory power to the model, is the primary metric considered. The partial F-test used is as follows [Ott and Longnecker 2001]:

$$F = \frac{\left(SSE_R - SSE_F\right)/(k - L)}{SSE_F / (n - k - 1)}$$

Equation 4.2 – Partial F Test for Significance of Regression

where: SSE_R = error sum of squares in the reduced model,

SSE_F = error sum of squares in the larger model,

k = number of variables in the larger model,

L = number of variables in the reduced model, and

n = sample size.

A second metric considered in adding variables to the model is the adjusted coefficient of determination (R^2_{adj}). Equation 4.3 shows the calculation of R^2_{adj}. Because R^2_{adj} considers the mean sum of squares, the value increases if the regression model is actually improved by the addition of an explanatory variable whereas the R^2 increases with each additional regressor variable [Montgomery et al. 2006].

$$R^2_{adj} = \frac{MS_{Res}}{MS_{Total}} = \frac{(SS_{Res})/(n - k)}{(SS_{Total})/(n - 1)}$$

Equation 4.3 – Adjusted Coefficient of Determination

While dependent on the true underlying relationship, the R^2_{adj} for the models created here are expected to be in the 50 – 75% range. This projection is based on findings from an earlier pilot study, and generally accepted thresholds for retrospective engineering studies. Lower R^2_{adj} values do not indicate a failure of the regression analysis. Conversely, higher R^2_{adj} values do not indicate a true fit to reality, only the data analyzed. As such, the models are evaluated on additional criteria, explained later.

4.5.2.3 Logical Considerations

The regression model must also meet criteria for being practical and logical in nature, at least in the parameters. As such, transformations and interaction effects modeled should be used only where significant improvement in the model is demonstrated. Therefore, improvement must be statistically significant and practically important.

An effort was made to minimize the degree of the model, testing for quadratic functions and below. Generally, higher than quadratic functions should be avoided unless subject matter knowledge, beyond the data set, warrants [Montgomery et al. 2006]. Interaction effects were tested where practical knowledge justified this possibility.

4.5.3 Regression Models

Two project types are considered in this study. Within each project type, two regression models were developed to attempt to identify and model the relationship between early-known project factors and construction duration. The various models, variables, parameter estimates, statistical significance, and practical importance are discussed here.

For Full-Depth Section and Highway Improvement projects, the "Full Model" is the model which incorporates *DWMIP* and *PCES* data (see Figure 4.1). Because the *PCES* construction cost estimate is built from these variables, the cost estimate is not used in the "Full Model" regressions. The "Condensed Model" uses *DWMIP* data and the summation of *PCES* data via the construction cost estimate. This is done to provide an alternative measure or relationship between these early known factors that is likely simpler, but may not be as accurate in modeling construction duration.

4.5.3.1 Full-Depth Section Models

A stepwise regression approach is used for both the full and condensed versions of the model. This approach produced statistically significant and practically important hierarchical regression models.

The models developed (Full and Condensed versions) are compared on a number of statistical criteria including:

- number of independent variables,
- Adjusted Coefficient of Variation (R^2_{adj}),
- R^2 for Prediction (R^2_{pred}),
- Fisher's F,
- Mallow's C_p, and
- Akaike Information Criterion (*AIC*).

The number of independent variables is indicative of the complexity of the model developed. It is generally preferable to minimize the number of variables in the model while maintaining the ability to predict the main features of the underlying relationship. R^2_{pred} gives an indication of the predictive capability of the regression model using new observations [Montgomery et al. 2006]. Fisher's F (or F ratio) is a measure of the significance of regression, performed by comparing the mean squares for regression and residuals. Higher F ratios indicate greater significance of regression. Mallow's C_p is used to determine whether the model has an appropriate number of regressor variables or whether it has been "over-fit" with too many regressors [NIST 2004]. When an ideal model is found, C_p will be equivalent to the number of variables in the model. Finally, Akaike Information Criterion (*AIC*) is a model selection criterion that measures the goodness of fit of a model. Ideally, the model with the lowest *AIC* value should be selected.

All variables will be considered in the model selection process. In model selection, the preferable model is that with the lower number of independent variables, and *AIC*, while maintaining a higher R^2_{adj}, R^2_{pred}, and F ratio. Finally, the optimal model should have a Mallow's C_p which is near to the number of variables in the model.

The final criterion of consideration is the practicality or ease of use of each model. Cumbersome models are difficult to incorporate in practice. The increased number of independent variables adversely effects the interpretability of the model and generally increases the chance of not having the appropriate information to input into the model. While a condensed model is considered, this research is focused on identifying and modeling the early-known construction factors that influence construction duration. Due to the unique nature of the data, the condensed model is presented primarily to show the impact of having detailed variable information versus only limited factors or a conglomerate construction cost estimate variable. However, the condensed model will be considered for prediction purposes as the model will likely contain fewer independent variables and be easier to use than the full models.

4.5.3.1.1 Full Model

The full model developed regresses early-known construction factors against the construction duration to determine which, if any, factors influence construction duration. Due to the nature of the data collected, the full model does not consider construction cost as this is a product of the individual *PCES* variables.

Stepwise regression analysis was used to determine the final model. Again, the stepwise approach uses a mixture of forward selection and backward elimination to produce the optimal model. The significance level for entry and removal (α_{enter} and α_{remove}) was set at 10% ($\alpha = 0.1$). In other words, regressors in the models have only a 10% chance of not being statistically significant. This is a commonly used threshold and is narrow enough to restrict insignificant variables from entering the model and broad enough to allow the identification of duration-influential project factors.

The full model includes 15 independent variables and is both statistically significant ($F = 15.11$ and $p\text{-}value < 0.0001$) and practically important as it explains 80%

of the variability in the response (R^2_{adj} = 0.80). The variables found to statistically significant are listed in Table 4.1. As seen, the terms included in the full model include both continuous and indicator variables. The continuous variables are Traffic Volume (Average Daily Traffic (ADT)), and New Signal Count. Indicator variables include a Geometric Design Standard, District, whether the project roadway has a Median, and whether the project roadway has Curb & Gutter. Staunton District was found to have a statistically significant influence on construction duration. All Geometric Design Standards were found to be statistically significant (at α = 0.1) and are, therefore, included as indicator variables. Estimates are in units of Working Days. Further explanation of the District and Geometric Design Standard indicator variables is necessary here.

VDOT is segregated into 9 Districts, all of which were included in the regression analysis. However, only one District (Staunton) was found to influence construction duration. As Bristol is the base District (i.e. indicators $d_1 - d_8$ equal 0), this is interpreted as a statistically significant difference from the influence of Bristol District on the construction duration. In essence, if all other factors are unchanged, the construction duration in Staunton is reduced by *29.1* working days. The remaining 7 Districts do not influence the construction duration in a statistically significant manner.

VDOT classifies projects by 9 Geometric Design Standards, all of which were represented and analyzed here. GS-1 Rural Principal Arterial is the base value for the Geometric Design Standard variable and the remaining 8 standards are represented using indicators $g_1 - g_8$. All 8 Geometric Design Standards demonstrate statistically significant influence on construction duration. The effect on the construction duration is linear in the amount of the coefficient *Estimate* for the project's Geometric Design Standard. These effects are relative to the base Geometric Design Standard (GS-1 Rural Principal Arterial). As an indicator variable, the applicable coefficient *Estimate* is multiplied by 1, and the remaining coefficient *Estimates* are multiplied by 0, cancelling them out.

All potential two-way interactions between project factors were considered in the regression analysis. These effects can explain a relationship between project factors and cause an otherwise seemingly insignificant variable to be significant in the hierarchical model. In this model, a statistically significant interaction effect exists between the

Traffic Volume (*ADT*) and Geometric Design Standard g_3 ($g_{3\text{ – GS 4 Rural Local Road}}$). This implies that construction duration increases as traffic volume increases on rural local road projects, at more than an additive rate for each factor alone.

Table 4.1 – Full Model Parameter Estimates

Term	Estimate	Std Error	t Ratio	P-Value
Intercept (β_0)	271.35	53.42	5.08	<.0001
ADT	0.003	0.002	1.84	0.0736
New Signal Count	141.47	34.20	4.14	0.0002
$CG_{1\text{ - Has Curb \& Gutter}}$	84.03	34.70	2.42	0.0206
$M_{1\text{ - Has Median}}$	-130.59	48.04	-2.72	0.010
$d_{8\text{ - Staunton District}}$	-29.10	15.98	-1.82	0.077
$g_{1\text{ - GS 2 (Rural Minor Arterial)}}$	-250.38	57.36	-4.37	0.0001
$g_{2\text{ - GS 3 (Rural Collector Road)}}$	-133.10	54.12	-2.46	0.0188
$g_{3\text{ - GS 4 (Rural Local Road)}}$	-201.06	53.91	-3.73	0.0007
$g_{4\text{ - GS 5 (Urban Principal Arterial)}}$	-172.25	43.91	-3.92	0.0004
$g_{5\text{ - GS 6 (Urban Minor Arterial)}}$	-347.86	90.24	-3.86	0.0005
$g_{6\text{ - GS 7 (Urban Collector Street)}}$	-97.91	48.83	-2.00	0.0525
$g_{7\text{ - GS 8 (Urban Local Street)}}$	-228.90	62.03	-3.69	0.0007
$g_{8\text{ - Interstate}}$	-298.47	86.49	-3.45	0.0014
g_3 x ADT	0.12	0.02	5.40	<.0001

4.5.3.1.2 Condensed Model

While the full model serves the main purposes of this research, the condensed model is presented here to offer an alternative to the larger model and to verify the ability of the full model to predict construction duration using multiple variables. By contrast, the condensed model is based on the construction cost estimate, which is predicted using the aforementioned *PCES* factors and a prediction model developed by VDOT. As such, the error and variability of the construction variables are included in the construction cost estimate. The condensed model also analyzes *DWMIP* data fields (Area Location Name and Highway System) to identify influence on construction duration. Based on these facts, it is suggested that the condensed model will not be able to explain variability in the data as well as the full model. However, the condensed model is expected to have fewer independent variables and, thus, be simpler and easier to use for prediction purposes.

A statistically significant condensed model was obtained using stepwise regression ($\alpha_{enter} = \alpha_{remove} = 0.1$, $F = 50.96$, $p\text{-}value < 0.0001$) with many fewer independent variables (4 in the condensed model versus 15 in the full model). Also, as expected, the condensed model has less ability to explain construction duration variability than the full model (*Condensed* $R^2_{adj} = 0.75$ versus *Full* $R^2_{adj} = 0.80$). Details of the model estimates are given in Table 4.2.

The condensed model uses the natural logarithm of the Construction Estimate from *PCES*, a Highway System indicator variable for interstate roadway projects, and the interaction between the terms to predict construction duration. The logarithmic transform is used to linearize the model terms and improve the explanatory power of the model. Also, the *PCES* Construction Estimate is in millions of dollars, escalated to January 1, 2008. This adjustment was done using the same escalation factors VDOT uses in *PCES* to account for the yearly increase in construction costs. Using millions of dollars makes the parameter estimate simpler and decreases the potential for error in model application.

While not trivial, the 5 percent decrease in R^2_{adj} using the condensed model is less than expected. Further analysis of the indicator variable provides some explanation for this. As expected, the *PCES* Construction Estimate by itself is less capable ($R^2_{adj} = 0.69$) of explaining the variability in the construction duration then the cumulative full model. A Partial-F test is used to evidence the significance of $h_{4 - Interstate}$ and the interaction term h_4 *x ln Construction Estimate* and the addition of explanatory power to the model through their use. Hence, the *PCES* Construction Estimate as a surrogate for the individual *PCES* variables is not as capable and the additional explanatory variables ($h_{4-Interstate}$ and h_4 *x ln Construction Estimate*) from *DWMIP* significantly improve the explanatory power of the model.

Table 4.2 – Condensed Model Parameter Estimates

Term	Estimate	Std Error	t Ratio	P-Value
Intercept (β_0)	118.20	5.02	23.53	<.0001
$h_{4\text{ - Interstate}}$	140.13	38.87	3.60	0.0008
ln Construction Estimate (2008 $M)	54.78	5.24	10.45	<.0001
h_4 x ln Construction Estimate	-40.62	20.55	-1.98	0.0541

Consideration of individual project factors produces very different linear regression models. The statistics listed in Table 4.3 are those commonly used to analytically judge and select regression models. As seen, the full model incorporated 11 more independent variables, and explains an additional 5 percent of data variability. Each model has adequate Mallow C_p values (less than or equal to the number of independent variables in the model). However, the condensed model has a lower *AIC*. The R^2_{pred} for the full and condensed models are 77% and 74%, respectively, meaning either model is expected to be able to explain approximately three-quarters of the variability in new observations. Because the two models use different regressor variables, the significance of adding the additional variables cannot be tested using the Partial-*F* test. Instead, the remaining summary statistics for the model are the metrics for comparison.

Table 4.3 – Full-Depth Section Model Comparison Statistics

	I.V.	R^2_{adj}	R^2_{pred}	*F*	C_p	*AIC*
Full Model	15	0.80	0.77	15.11	13	395.15
Condensed Model	4	0.75	0.74	50.96	4	358.62
Δ (Full – Cond.)	11	0.05	0.03		9	36.53

Given these facts, it appears that the full model is superior to the condensed model in explaining the variability in construction duration. Further, the full model, due to its inclusion of individual construction project variables, is far superior to the condensed model in identifying the early-known project factors having a statistically significant linear relationship with construction duration (Geometric Design Standard, Traffic Volume, Curb & Gutter, and New Signal Count). However, the condensed model has similar explanatory power, lower *AIC*, and fewer regressor variables. Therefore, the condensed model is expected to be the better model for prediction of Full-Depth Section construction duration.

4.5.3.2 Highway Improvement Models

The same process for analyzing the Full-Depth Section projects is used for Highway Improvement Models. Here, two models are also created (full and condensed) to better understand the relationships between early-known project factors and construction duration.

4.5.3.2.1 Full Model

The Highway Improvement full model investigates the early-known project factors and their relationship with construction duration. As such, the full model incorporates these factors separately but does not include the *PCES* construction cost estimate.

The same potential regressors used in Full-Depth Section project analysis are used for Highway Improvement projects, however their relationships and potential transforms on the regressors are considered independently. None of the statistically significant regressor variables demonstrated a non-linear relationship with construction duration.

Using stepwise regression, the full model identified is statistically significant (F = 14.66 and $p\text{-}value$ < 0.0001). However, the model does not explain as much of the variation in construction duration (R^2_{adj} = 0.64) as the Full-Depth Section full model. The model developed incorporates fewer than expected explanatory variables (6) compared to the Full-Depth Section project model with 15 variables.

The statistically significant project factors included in the model are Traffic Volume (*ADT*), Geometric Design Standard, presence of Median ($M_{1\ -\ Has\ Median}$), and New Signal Count. Again, potential interaction effects were considered, but none were found to be statistically significant (at α = 0.10 significance level). The variables included and their coefficients are listed in Table 4.4 below.

Table 4.4 – Full Model Parameter Estimates

Term	Estimate	Std Error	t Ratio	P-Value
Intercept (β_0)	47.45	3.98	11.93	<.0001
ADT	0.001	0.000	6.15	<.0001
$M_{1\text{ - Has Median}}$	-55.19	12.38	-4.46	<.0001
New Signal Count	18.53	5.79	3.20	0.0030
$g_{5\text{ - GS 6 (Urban Minor Arterial)}}$	42.30	11.63	3.64	0.0009
$g_{6\text{ - GS 7 (Urban Collector Street)}}$	59.66	14.81	4.03	0.0003

Traffic Volume and New Signal Count are continuous variables and appear in both the Full-Depth Section and Highway Improvement models. Unlike the Full-Depth Section full model, only two (2) Geometric Design Standards influence construction duration for Highway Improvement projects. The two standards (*GS 6* and *GS 7*) are both related to urban roadways, and add to the construction duration. The model tells us that other Geometric Design Standards do not significantly impact or alter construction duration. The other indicator variable included in the model ($M_{1\,-\,Has\ Median}$) shows a decrease in construction duration when the roadway under construction has a median. The validity of the model and parameter estimates is discussed later.

4.5.3.2.2 Condensed Model

Recall the Full-Depth Section models demonstrated improved explanatory power by incorporating the individual project factors in place of the conglomerate construction cost estimate (see Table 4.3). The same improvement is expected in the Highway Improvement projects.

Unlike the Full-Depth Section projects, the construction cost estimate did not warrant a transform to linearize the relationship with construction duration. Numerous transforms were tested (including logarithmic, quadratic, and reciprocal) however, the linear form was the most statistically significant form ($F = 21.08$ and $p\text{-}value < 0.0001$).

Using stepwise regression, four factors were considered in the condensed model regression analysis: *PCES* Construction Cost (2008 $M), Construction Length (miles), Area Location Name (indicator with 3 levels), and Highway System (indicator with 4 levels). The model selected incorporated just two variables, in addition to the intercept, and proved statistically significant ($F = 23.11$ and $p\text{-}value < 0.0001$). However, once again, the model does not explain as much of the variability in construction duration ($R^2_{adj} = 0.55$) as its Full-Depth Section model counterpart. The R^2_{adj} for the full model is higher than that of the condensed model (R^2_{adj} = 0.64 and 0.55, respectively). This finding is inline with the hypothesis that individual project variables are better able to explain the variability in construction duration than the *PCES* construction cost estimate alone.

Table 4.5 – Condensed Model Parameter Estimates

Term	Estimate	Std Error	t Ratio	P-Value
Intercept (β_0)	37.40	5.61	6.670	<.0001
$a_{2\text{ - Medium Urban (50,000 - 199,999)}}$	69.00	22.83	3.020	0.005
Construction Estimate (2008 $M)	40.18	6.76	5.940	<.0001

Table 4.6 shows the direct comparison of the full and condensed Highway Improvement models.

Table 4.6 – Highway Improvement Model Comparison Statistics

	I.V.	R^2_{adj}	R^2_{pred}	F	C_p	*AIC*
Full Model	6	0.64	0.52	14.66	6	260.07
Condensed Model	3	0.55	0.49	23.11	3	252.52
Δ (Full – Cond.)	3	0.09	0.03		3	7.55

As seen, the full model is the better of the two models across all metrics considered, except *AIC* which was lower in the condensed model than the full model. Given these facts, it appears that the full model is superior to the condensed model in explaining the variability in construction duration for Highway Improvement projects. Further, the full model is superior to the condensed model in identifying and modeling the early-known project factors having a statistically significant linear relationship with construction duration (Geometric Design Standard, Traffic Volume, New Signal Count, and Median). Also, like the Full-Depth Section models, the Highway Improvement condensed model has a similar R^2_{adj}, lower *AIC*, and fewer regressor variables. Once again, this leads to the expectation that the condensed model will be the better model for predicting Highway Improvement construction duration.

4.5.4 Residual Analysis and Outliers

Retrospective studies are particularly prone to outliers. Outliers are extreme observations in the data that are not typical of the rest of the data set [Montgomery et al. 2006]. Outliers can be identified by preparing a scatter-plot of the residual by predicted values, or numerically from the residual values for the estimation data set.

Residual analysis was completed by analyzing the scatter-plot of residual by predicted values to ensure that the residuals were evenly distributed in a general horizontal band about zero. Also, a normal probability plot of the studentized residuals was constructed for each model. The plot of the studentized residuals should form a near straight line. Deviations (heavy or light tails) indicate potential outliers or unstable variance.

Studentized residuals (r_i) are scaled to a constant variance and show deviation of data points from the model developed. If $|r_i| \geq 3$, the point is a potential outlier and warrants additional analysis. Few outliers were identified in this study. One record from each of the Full-Depth Section and Highway Improvement analyses was identified as outliers. Records demonstrating outlier traits were those that had an abnormally high or low construction cost relative to the construction duration. These points have been documented and removed from the regression analysis. Further, these values have not been included in calculating the model parameter estimates and statistics in tables 4.1, 4.2, 4.4, and 4.5.

4.6 Model Validation

The general model validation procedure consists of stability analysis and coefficient validation [Montgomery et al. 2006]. Stability analysis and validation ensures the model is sufficient for predicting new or future values. This can be done by either collecting and analyzing new data (data not included in the estimation set) or through splitting the existing data set into estimation and prediction data sets. Due to the retrospective nature of this study, the latter was chosen and the splitting of the data sets was previously discussed.

4.6.1 Data Splitting

Data splitting refers to the process by which the data population is divided into estimation and prediction data sets for regression analysis and validation. The estimation data set is used to complete the statistical regression analysis while the prediction data set is used to validate the model and ensure that the models prepared reflect reality. Once the regression model is constructed and validated, the full data set is used to establish the parameter estimates. Note that the parameter values shown in tables 4.1, 4.2, 4.4, and 4.5 are calculated using the full data set.

Splitting the data set can be done in various proportions [Ott and Longnecker 2001]. Halving the data is perhaps the most common split. To split the data into equal halves, the following inequality must be satisfied:

$$n \geq 2p + 20$$

Equation 4.4 – Required Sample Size for Data Splitting

where: n = sample size, and

p = number of parameters in the model [Ott and Longnecker].

Initial estimates for the number of potential regressors suggested that as many as 30 parameters could enter the model. As such, the sample size needed to halve the data set into estimation and prediction sets is 80. Halving the data set would therefore require that 40 projects be used for estimation and 40 projects for prediction. However, both populations (Full-Depth Section and Highway Improvement) are smaller than the required 80. As such, the data sets are not split into equal-size estimation and prediction sets. Instead, the estimation sets will be comprised of 2/3 of the data, while the prediction data sets will be made up of the remaining 1/3 of the data sets. The estimation data set receives more records than the prediction set to enhance the modeling process. Generally, the potential for explaining variance in the data is improved through increasing the estimating sample size. This splitting is done through *JMP 6.0* using the random sampling and exclusion of rows from the regression analysis.

4.6.2 Stability Validation

Stability validation is used to test the ability of the model to predict the response using actual data. Here, the prediction data set records are used, along with the models developed, to calculate construction duration. The durations developed are examined for correlation or similarity to the actual construction duration. The primary metric for stability validation is the Pearson coefficient of correlation (R_{corr}). R_{corr} measures the correlation present between actual (y) and predicted ($\hat{y}$) response (construction duration). Generally, this value is expected to be close to the value of R^2_{adj} to yield a stable model. Pearson's coefficient of correlation is also tested to ensure a statistically significant correlation (ρ) is exhibited by the model. This test is performed using a t-test and the R_{corr} derived from comparison of the predicted construction duration and the actual construction duration of the cross-validation data set.

$$H_o : \rho = 0$$
$$H_a : \rho \neq 0$$
$$t_{obs} = R_{corr} \times \sqrt{\frac{n - 2}{1 - R_{corr}^2}}$$

Equation 4.5 – Test for Significance of Correlation

where: R_{corr} = Pearson coefficient of correlation, and

n = prediction set sample size.

If $|t_{obs}| \leq t_{1-\alpha/2, n-2}$ then fail to reject the null hypothesis (H_o) and conclude that R_{corr} and correlation (ρ) is significantly different than zero. The results of the stability validation test for all models are included in Table 4.7.

Table 4.7 – Stability Validation Using R_{corr}

	Model	R_{corr}	$\lvert t_{obs}\rvert$	$t_{1-\alpha/2, n-2}$	Conclusion
Full-Depth	Full	0.304	1.393	1.717	Accept the null ($\rho = 0$)
Section	Condensed	0.784	5.930	1.717	Reject the null ($\rho \neq 0$)
Highway	Full	0.354	1.609	1.734	Accept the null ($\rho = 0$)
Improvement	Condensed	0.463	2.215	1.734	Reject the null ($\rho \neq 0$)

As seen, the condensed models show significantly higher correlation than the full model counterparts. In the Full-Depth Section models, this was expected due to the higher number of variables in the full model compared to the condensed model. The high number of independent variables and the low R_{corr} suggests that the full models are over-parameterized.

With the lower number of independent variables (6), such a problem would not be expected in the Highway Improvement models. The full model has three additional independent variables. However, testing of the condensed models shows a significant correlation ($R_{corr} = 0.463, |t_{obs}| = 2.215 > t_{1-\alpha/2, n-2} = 1.734$) between predicted and actual construction duration values whereas the full model does not ($R_{corr} = 0.354$, $|t_{obs}| = 1.609 < t_{1-\alpha/2, n-2} = 1.734$).

4.6.3 Coefficient Validation

Coefficient validation relies on subject matter knowledge and understanding of the independent and dependent variables. The primary focus of coefficient validation is to ensure that coefficients return the expected sign and that they are of proper magnitude.

The coefficient validation process is performed after the stepwise regression and model development processes. The initial phase of this process is to examine the implied effects of the parameter coefficients on the cumulative construction duration. Duration is mostly additive in that project features (such as adding crossovers or signals) tend to add time to the project (working days). The addition of work or to quantities of work should generally increase project duration. However, this is not always the case as larger work quantities may warrant changes in construction equipment or sequence and could in turn reduce overall construction duration. An example of this is the indicator $M_{1\text{ – Has Median}}$ in the full models. The parameter estimate suggests that the presence of median leads to a decrease in duration, and does not indicate that additional work is created in building the median. This seems logical as the presence of a median generally expands the available work area and separates the work crews from oncoming traffic. Such conditions could improve the productivity, thereby reducing the number of working days required.

While a number of the parameters are additive, the signs associated with indicator variable parameters are much more open to interpretation. As mentioned, indicator variables evidence an effect on construction duration from the base variable (base indicators include: District – Bristol, Area – Rural, Highway System – Primary Arterial Network, and Geometric Design Standard – GS-1 Rural Principal Arterial). Therefore, the signs on these regressors may be positive or negative indicating a higher or lower construction duration from the base. While subject matter knowledge is helpful in this evaluation, the indicator variables are not excluded based on sign alone. The magnitude must also be considered to ensure a plausible relationship.

The magnitude of the coefficients is evaluated to ensure that a realistic function exists between both qualitative and quantitative variables with construction duration. Also, examination of the parameter estimates depict potential scaling transforms that should be performed to make the model more useful. One example of this is the *PCES* Construction Cost Estimate which has been transformed to millions of dollars. This change alters the parameter coefficient by the same factor (10^6), removing zeros after the decimal which could cause errors for the user if missed. There does not appear to be any issues with the magnitude of the parameter estimates in the models developed.

4.7 Results

The results of the statistical regression analysis are separated into discussions regarding the Full-Depth Section and Highway Improvement models. The models developed are reviewed and explained.

4.7.1 Full-Depth Section Results

The Full-Depth Section statistical regression analysis developed two statistically significant and practically important regression models to predict construction duration. As expected, the full model was better able to explain variability in construction duration than the condensed model (R^2_{adj} = 0.80 versus 0.75, respectively). However, in the validation process, the full model was not as stable as the condensed model, (R_{corr} = 0.304 versus 0.784, respectively).

The full and condensed models developed are presented below with an explanation of the independent variables. The coefficient values are found in Tables 4.1 and 4.2.

$$DURATION = y = \beta_0 + \beta_1 \bullet ADT + \beta_2 \bullet SigCnt + \beta_3 \bullet CG_1 + \beta_4 \bullet M_1 + \beta_5 \bullet Dist + \beta_{6i} \bullet GeoStd + \beta_7 \bullet (g_3 \times ADT)$$

Equation 4.6 – Full-Depth Section Full Model

where: DURATION = Construction Duration (working-days) = y,

β_o = Intercept,

β_{6i} = g_1 if *GeoStd* is GS-2, g_2 if *GeoStd* is GS-3, …, g_7 if *GeoStd* is GS-8, 0 otherwise,

ADT = Traffic Volume (vehicles/day),

SigCnt = New Signal Count,

CG_1 = 1 if roadway has curb & gutter, 0 otherwise,

GeoStd = 1 if *GeoStd* is GS-2 through GS-8, 0 otherwise,

M_1 = 1 if roadway has median, 0 otherwise,

$Dist$ = 1 if District is Staunton, 0 otherwise, and

g_3 = 1 if $GeoStd$ is GS-4, 0 otherwise.

$$DURATION = y = \beta_0 + \beta_1 \bullet CostEst + \beta_2 \bullet HWSystem + \beta_3 HWSystem \times CostEst$$

Equation 4.7 – Full-Depth Section Condensed Model

where: DURATION = Construction Duration (working-days) = y,

β_o = Intercept,

$\beta_2 = h_4$ if $HW\ System$ is Interstate,

$CostEst$ = ln($PCES$ Construction Cost Estimate (2008 \$M)), and

$HWSystem$ = 1 if Interstate, 0 otherwise.

4.7.2 Highway Improvement Results

For the Highway Improvement models, the full model generally outperformed the condensed model. The full model was better able to explain variability in construction duration than the condensed model (R^2_{adj} = 0.64 versus 0.55, respectively). However, the condensed model proved more stable than the full model (R_{corr} = 0.463 versus 0.354, respectively). This is counter-intuitive as the number of independent variables in the models differed by just 3 variables.

The Highway Improvement models are presented below. Coefficient estimates can be found in Tables 4.4 and 4.5.

$$DURATION = y = \beta_0 + \beta_1 \bullet ADT + \beta_2 \bullet SigCnt + \beta_3 \bullet M_1 + \beta_{4i} GeoStd$$

Equation 4.8 – Highway Improvement Full Model

where: DURATION = Construction Duration (working-days) = y,

β_o = Intercept,

$\beta_3 = M_1$ if roadway has median, 0 otherwise,

$\beta_{4i} = g_5$ if $GeoStd$ is GS-6, g_6 if $GeoStd$ is GS-7, 0 otherwise,

ADT = Traffic Volume (vehicles/day),

$SigCnt$ = New Signal Count,

M_1 = 1 if roadway has median, 0 otherwise, and

$GeoStd$ = 1 if $GeoStd$ is GS-6 or GS-7, 0 otherwise.

$$DURATION = y = \beta_0 + \beta_1 \bullet CostEst + \beta_2 \bullet Area$$

Equation 4.9 – Highway Improvement Condensed Model

where: DURATION = Construction Duration (working-days) = y,

β_o = Intercept,

CostEst = *PCES* Construction Cost Estimate (2008 $M), and

Area = 1 if Medium Urban (50,000 – 199,999), 0 otherwise.

4.8 Conclusions

This document presents a statistical regression analysis study aimed at *1)* identifying the early-known construction duration-influential factors, *2)* determining the coefficient or relationship of these factors with highway construction duration, and *3)* modeling the build-up of construction duration from these factors. The conclusions reached in this work are discussed under the heading of these three objectives.

First, early-known highway construction duration influential factors were identified. This work developed full and condensed models for Full-Depth Section and Highway Improvement project types and identified the duration-influential project factors in each data set. The primary explanatory variable in the condensed models is the *PCES* Construction Cost Estimate. This estimate incorporates the numerous construction factors considered by *PCES*. Full models identify the individual factors that influence both the construction duration that are incorporated into the *PCES* construction cost estimate. These factors include District, Geometric Design Standard, presence of Median, presence of Curb & Gutter, New Signal Count, and Traffic Volume. The factors that are statistically significant in both the Full-Depth Section and Highway Improvement analyses are Traffic Volume and New Signal Count. All Geometric Design Standards are influential in the Full-Depth Section data set but not in the Highway Improvement set. Only GS 6 (Urban Minor Arterial) and GS 7 (Urban Collector Street) demonstrated a statistically significant relationship with Highway Improvement construction duration. Also, the only VDOT District which appears to effect construction duration is Staunton. Though tested in both full models, none of the remaining VDOT District showed evidence of a relationship with construction duration.

MLR is an excellent technique for identifying construction duration-influential factors, whether in the form of continuous, quantifiable variables or as qualitative indicator variables. Most of the duration-influential factors are *PCES* or project-level factors. As such, it appears that the project size and complexity are better indicators of the project duration than the conditions (i.e. setting and location).

Second, the relationships between these duration-influential factors and construction duration are identified using statistical regression analysis. Again, MLR facilitates the quantification of these relationships and the intuitive understanding of the magnitude, form, and shape of the relationships between factors and with construction duration. These relationships, listed as coefficient or parameter estimates are consistent with subject matter knowledge and general principles of duration estimating. These coefficient values are found in Tables 4.1, 4.2, 4.4, and 4.5. The intercept is identified through this process as well. If the principle were to hold that the addition of work items or work quantities correlates to additional construction duration, the intercept term would represent the minimal construction duration achievable. However, in the case of this analysis, indicator variables (which demonstrate the difference in effect of the variable from the base indicator or category) invalidate this assumption. For instance, for the full Full-Depth Section model, the intercept of 271.35 working days is not the minimum construction duration because the indicator variables included in the model generally reduce this duration. Therefore, the intercept cannot be taken as the minimum construction duration, but is included due to its importance to the model.

Finally, the statistically significant relationships determined are modeled to predict highway construction duration for both Full-Depth Section and Highway Improvement models. In both the Full-Depth Section and Highway Improvement data sets, the model including individual construction project variables is better able to explain variability in construction duration than the model using the *PCES* construction cost estimate as a surrogate. However, the condensed models developed are simpler and easier to use than the full models, because they have fewer variables. The condensed models also demonstrated similar abilities in explaining the variability in construction duration (R^2_{adj} = 0.80 and 0.75 for Full-Depth Section full and condensed models, respectively; R^2_{adj} = 0.64 and 0.55 for Highway Improvement full and condensed models, respectively). The additional independent variables in the full models cause the appearance of instability and over-parameterization in the models and make the validation process more difficult, here, due to the high ratio of variables to data records. Such instability could lead to inaccurate construction duration estimates. This research has successfully developed a methodology for identifying and modeling early-known project factors that influence highway construction duration. However, expansion of the current data set would improve the validation analysis for the full models.

It is felt that the condensed models are an adequate representation and tool for predicting the construction duration using early-known project factors. The condensed model uses the conglomerate *PCES* construction cost estimate to describe the size and complexity of the project and account for individual project factor values. Just as the project-level factors were the predominant explainers of construction duration, the *PCES* cost estimate explains much of the variation in construction duration. The models produce construction (working-day) duration. This duration may be used by the scheduler to assist in the estimation of contract duration early in the design process.

4.9 Recommendations

Several recommendations for future research are derived from this study. These recommendations are detailed here.

- Expand the data set by collecting more project data. This study made use of all data available from a historical database of projects whose cost has been estimated using VDOT's *PCES* system. As a retrospective study, the analysis is limited to the data available and must exclude those records deemed erroneous or invalid. While the analysis began with a large data set (294 projects) the regression analysis data sets were reduced to 75 records for Full-Depth Section projects and 61 records for Highway Improvement projects. These population sizes, in conjunction with the large number of potential explanatory variables, may be the cause of the full models appearing unstable.
- Since this study was a retrospective analysis, the quantity and quality of the data was largely beyond the control of the researcher. Designed studies offer the ability to monitor and systematically improve the data collected. However, as

this study made use of several years of construction data, a designed study would not have been practical. It is thought, though, that a designed study lasting 2-3 years, on a larger spectrum of project data (i.e. not just projects with *DWMIP* and *PCES* data) could yield a sufficient data set for analysis.

- Again, as a retrospective study, the project factors that could be analyzed were somewhat limited. *PCES* maintains the project factors, known during early design, thought to contribute to the construction cost. As such, these factors may not be the only variables which impact construction duration. There is potential for future research in expanding the factors collected for the purposes of construction duration analysis.
- Finally, the study presented uses historical data and is a 'snapshot' of highway construction duration information. This study did not seek to review individual project records, instead using statistical regression analysis to identify potential outliers or projects whose relationship between construction duration and project factors did not fit the general relationship. Because of this, the construction duration prediction models developed could still be influenced by poorly measured contract durations. It is felt that a designed study would, again, assist in insuring the data collected and analyzed is of the highest quality.

4.10 Acknowledgements

The authors would like to acknowledge the Virginia Tech Laboratory for Interdisciplinary Statistical Analysis. Particularly, consultants Abdo G. Abdel-Salam and Yi Chao are thanked for their consultation services and guidance in completing the statistical regression analysis portion of this work.

4.11 References

Bromilow, F. J. (1969). "Contract time performance expectations and the reality." *Building Forum*, 1(3).

Chan, A. P. C., and Chan, D. W. M. (2004). "Developing a benchmark model for project construction time performance in Hong Kong." *Building and Environment*, 39, 339-349.

Jiang, Y., and Wu, H. (2007). "A Method for Highway Agency to Estimate Construction Durations and Set Contract Times." *International Journal of Construction Education and Research*, 3, 199-216.

Kahane, L. H. (2001). *Regression Basics*, Sage Publications, Inc., Thousand Oaks, CA.

Kyte, C. A., Perfater, M. A., Haynes, S., and Lee, H. W. (2004). "Developing and Validating a Highway Construction Project Cost Estimation Tool." Virginia Transportation Research Council, Charlottesville, VA.

Montgomery, D. C., Peck, E. A., and Vining, G. G. (2006). *Introduction to Linear Regression Analysis*, 4th Ed., John Wiley & Sons, Inc., Hoboken, NJ.

Myers, R. H. (1986). *Classical and Modern Regression with Applications*, Duxbury Press, Boston, MA.

Myers, R. H., Montgomery, D. C., and Vining, G. G. (2002). *Generalized linear models: with applications in engineering and the sciences*, Wiley-Interscience, New York.

National Institute of Standards and Technology (NIST) / SEMATECH, I. (2004). "NIST/SEMATECH e-handbook of statistical methods." <http://www.itl.nist.gov/div898/handbook>.

Ott, R. L., and Longnecker, M. (2001). *An Introduction to Statistical Methods and Data Analysis*, Duxbury, Pacific Grove, CA.

Snee, R. (1977). "Validation of Regression Models: Methods and Examples." *Technometrics*, 19(4), 415-428.

Stevens, J. P. (2001). *Applied Multivariate Statistics for the Social Sciences*, 4th Ed., Lawrence Erlbaum Associates, Mahwah, New Jersey.

Williams, R. C., Vorster, M. C., and Hildreth, J. C. (2008a). "The Need for Highway Construction Duration Estimating Using Early-known Design Details." *Transportation Research Record: Journal of the Transportation Research Board*, Submitted for review.

Williams, R. C., Hildreth, J. C., and Vorster, M. C. (2008b). "Highway Construction Data Collection and Treatment for Statistical Regression Analysis." *Journal of Construction Engineering and Management*, Submitted for review.

Chapter 5 – Prediction of Highway Construction Duration using Conceptual Design Parameters

5.1 Introduction

Knowledge of construction duration is pertinent to a number of project planning functions prior to detailed design development. Important project funding and financing decisions are made during the preliminary design stage using estimated construction duration. These decisions, such as the duration of financing, date of occupancy, and the beginning of positive cash flows arising from the construction investment, are imperative to project success. Resource allocation decisions are also critical to project, program, and organizational success. The construction owner must be able to adequately plan and allocate management and oversight personnel and funding. Further, resources must be allocated to multiple projects concurrently. This presents challenges to entire organization, and requires an understanding of construction duration early in project design.

Williams [2008a] developed a series of regression models for predicting construction duration using early-known project factors. The statistical regression analysis identified duration-influential project factors, quantified their relationship with construction duration, and created models for predicting construction duration using these duration-influential factors [Williams 2008a].

These models were developed through analysis of Virginia Department of Transportation (VDOT) data from construction projects starting and completing between January 1, 2000 and January 1, 2007. Two project types were considered in the analysis: Full-Depth Section and Highway Improvement. For each project type, full and condensed models were created. Full models analyzed project parameters gathered from *Data Warehouse Management Information Portal* (*DWMIP*) and *Project Cost Estimating System* (*PCES*). Condensed models used *DWMIP* data and the construction cost estimation – the output of *PCES*. This hierarchy is presented in Figure 5.1.

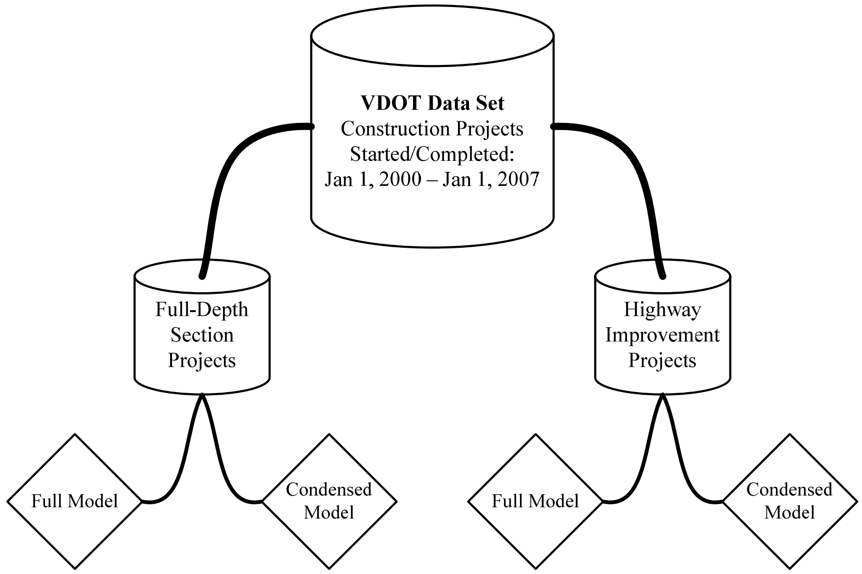

Figure 5.1 – Model Hierarchy

Statistical regression analysis showed that the full models were better able to explain variability in construction duration. Also, the full models served as the mechanism for identifying duration-influential project factors known early in design development. However, these models incorporated numerous variables and appeared to be over-parameterized and less stable than the condensed models. The condensed models incorporated fewer project parameters and adequately explained variability in construction duration. Hence, the condensed models are better suited to prediction of construction duration.

The purpose of this work is to discuss the application of the models and the larger research model. While specific and immediate application of the research is attributed to VDOT, the larger highway and general construction industries can apply the research to their purposes as well. This document discusses the potential for such application, including the application of the models and the research methodology.

5.2 Highway Construction Duration Estimating Models

Full and condensed models were developed for both Full-Depth Section and Highway Improvement project types. Figures 5.2 and 5.3, below, depict the difference in the full and condensed models from the data collection and analysis perspective.

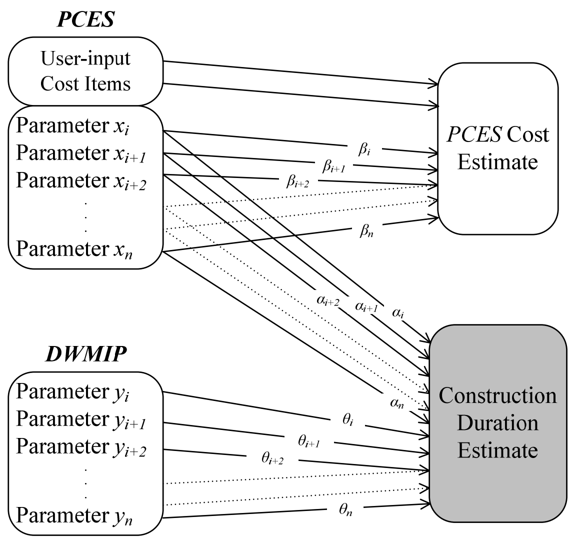

Figure 5.2 – Full Model Data Source

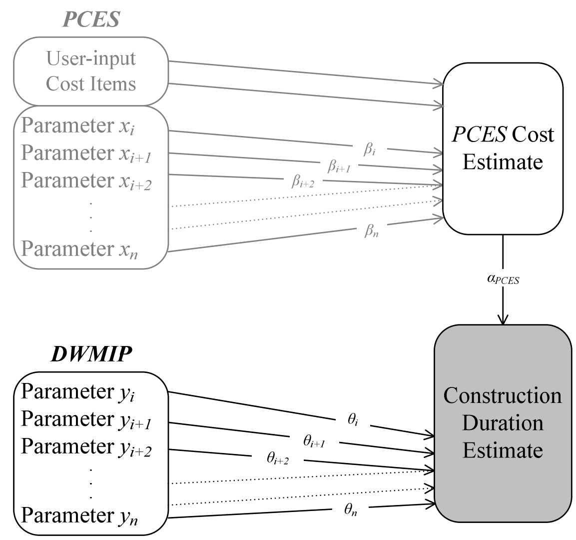

Figure 5.3 – Condensed Model Data Source

As seen in the figures above, the chief difference in the full and condensed models is the inclusion of either all *PCES* project variables, or the calculated *PCES* Construction Cost Estimate. The *PCES* Construction Cost Estimate is developed using the *PCES* variables and historical relationships with cost for VDOT projects [Kyte et al. 2004]. Hence, the regression models cannot use <u>both</u> *PCES* Construction Cost Estimate <u>and</u> the *PCES* factors. The full and condensed models were developed to test the hypothesis that a better (in terms of explanatory power) model could be developed using the individual factors, and to provide an alternative model that incorporated fewer variables and was easier to use. Here, the models developed for each project type are described.

5.2.1 Full-Depth Section Models

Full-Depth Section projects include new construction, reconstruction, and widening. To simplify the application process and provide more robust models, these project types have been considered together as Full-Depth Section projects. The models developed are described below.

5.2.1.1 Full-Depth Section, Full Model

A statistically significant (at α = 0.1) and practically important (R^2_{adj} = 0.80) multiple linear regression model was developed using both *DWMIP* and *PCES* project parameters. The analysis process identified Traffic Volume (ADT), New Signal Count, Curb & Gutter, Median, District, and Geometric Design Standard as duration-influential factors. These factors are incorporated into the construction duration (working day) prediction model below.

$$DURATION = y = \beta_0 + \beta_1 \bullet ADT + \beta_2 \bullet SigCnt + \beta_3 \bullet CG_1 + \beta_4 \bullet M_1 + \beta_5 \bullet Dist + \beta_{6i} \bullet GeoStd + \beta_7 \bullet (g_3 \times ADT)$$

Equation 5.1 – Full-Depth Section Full Model

where: DURATION = Construction Duration (working-days) = y

β_o = Intercept

β_{6i} = g_1 if *GeoStd* is GS-2, g_2 if *GeoStd* is GS-3, …, g_7 if *GeoStd* is GS-8, 0 otherwise

ADT = Traffic Volume (vehicles/day)

SigCnt = New Signal Count

CG_1 = 1 if roadway has curb & gutter, 0 otherwise

GeoStd = 1 if *GeoStd* is GS-2 through GS-8, 0 otherwise

M_1 = 1 if roadway has median, 0 otherwise

Dist = 1 if District is Staunton, 0 otherwise

g_3 = 1 if *GeoStd* is GS-4, 0 otherwise

Table 5.1 shows the coefficient estimates for the variables in the model as well as the 95% Confidence Interval for each coefficient estimate and the limits (minimum and maximum) of the variables included in the model. A confidence interval is a forecast interval for the mean coefficient estimate [Ott and Longnecker 2001]. In other words, a "95% confidence interval" means that there is 95% confidence that the parameter or coefficient estimate is within the bounds of the interval specified. As such, confidence intervals are indicative of the overall quality of the regression model [Montgomery et al. 2006]. Generally, the wider the confidence interval, the less certain the model is of the coefficient estimate.

The factor limits stated are the minimum and maximum data values entered in the analysis data set. The confidence interval is narrowest at the mean factor value. As the extents of the factor data is reached, the confidence interval widens, increasing the potential for an inaccurate estimate of the coefficient. The danger of extrapolation is noted here. Extrapolation beyond the factor limits is not recommended and will likely lead to an inaccurate estimate of construction duration.

Table 5.1 – Full-Depth Section, Full Model Summary

Term	Estimate	Estimate Confidence Interval			Factor Limits	
		Lower 95%	Upper 95%	Width	Min	Max
Intercept (β_0)	271.35	163.01	379.69	216.68	NA	
ADT	0.003	0.000	0.006	0.01	50	44,000
New Signal Count	141.47	72.10	210.84	138.74	0	2
CG_1 - Has Curb & Gutter	84.03	13.65	154.41	140.75	0	1
M_1 - Has Median	-130.59	-228.03	-33.15	194.88	0	1
d_8 - Staunton District	-29.10	-61.51	3.31	64.82	0	1
g_1 - GS 2 (Rural Minor Arterial)	-250.38	-366.71	-134.05	232.66	0	1
g_2 - GS 3 (Rural Collector Road)	-133.10	-242.85	-23.35	219.50	0	1
g_3 - GS 4 (Rural Local Road)	-201.06	-310.39	-91.73	218.66	0	1
g_4 - GS 5 (Urban Principal Arterial)	-172.25	-261.31	-83.20	178.11	0	1
g_5 - GS 6 (Urban Minor Arterial)	-347.86	-530.87	-164.85	366.01	0	1
g_6 - GS 7 (Urban Collector Street)	-97.91	-196.95	1.13	198.08	0	1
g_7 - GS 8 (Urban Local Street)	-228.90	-354.70	-103.11	251.59	0	1
g_8 - Interstate	-298.47	-473.87	-123.06	350.80	0	1
g_3 x ADT	0.12	0.07	0.16	0.09	0	2,200

5.2.1.2 Full-Depth Section, Condensed Model

The condensed model used *DWMIP* data and the calculated *PCES* Construction Cost Estimate to model construction duration. As expected, the model used far fewer variables (4 versus 15 in the full model). Also, as expected, the condensed model was less able to explain variability in construction duration (R^2_{adj} = 0.75 and 0.80 for condensed and full models, respectively). The Full-Depth Section condensed model is described below. The coefficient estimate and variable limits table follows.

$$DURATION = y = \beta_0 + \beta_1 \bullet CostEst + \beta_2 \bullet HWSystem + \beta_3 HWSystem \times CostEst$$

Equation 5.2 – Full-Depth Section Condensed Model

where: DURATION = Construction Duration (working-days) = y

β_o = Intercept

$\beta_2 = h_4$ if *HW System* is Interstate

CostEst = ln(*PCES* Construction Cost Estimate (2008 \$M))

HWSystem = 1 if Interstate, 0 otherwise

Table 5.2 – Full-Depth Section, Condensed Model Summary

Term	Estimate	Estimate Confidence Interval			Factor Limits	
		Lower 95%	Upper 95%	Width	Min	Max
Intercept (β_0)	118.20	108.09	128.31	20.22	NA	
$h_{4\text{ - Interstate}}$	140.13	61.88	218.38	156.50	0	1
ln Construction Estimate (2008 $M)	54.78	44.23	65.33	21.10	-2.194 $111,484	2.731 $15,342,746
h_4 x ln Construction Estimate	-40.62	-82.00	0.75	82.75	0.000 $0	2.731 $15,342,746

Notice here that the confidence intervals for the coefficient estimates in the condensed models are generally narrower than the full model. This is consistent with the findings regarding the instability of the full models developed and that the condensed models are better suited for prediction of construction duration.

5.2.2 Highway Improvement Models

Highway Improvement projects include restoration and rehabilitation and resurfacing projects. As with Full-Depth Section projects, both a full and condensed model were developed for Highway Improvement projects. The results are discussed below.

5.2.2.1 Highway Improvement, Full Model

A statistically significant (at α = 0.1) and practically important (R^2_{adj} = 0.64) multiple linear regression model was developed using both *DWMIP* and *PCES* project factors. The duration-influential factors include Traffic Volume (ADT), New Signal Count, Median, and Geometric Design Standard. The full model developed and the coefficient estimates are depicted below.

$$DURATION = y = \beta_0 + \beta_1 \bullet ADT + \beta_2 \bullet SigCnt + \beta_3 \bullet M_1 + \beta_{4i} GeoStd$$

Equation 5.3 – Highway Improvement Full Model

where: DURATION = Construction Duration (working-days) = y

β_o = Intercept

$\beta_3 = M_1$ if roadway has median, 0 otherwise

$\beta_{4i} = g_5$ if *GeoStd* is GS-6, g_6 if *GeoStd* is GS-7, 0 otherwise

ADT = Traffic Volume (vehicles/day)

SigCnt = New Signal Count

M_1 = 1 if roadway has median, 0 otherwise

GeoStd = 1 if *GeoStd* is GS-6 or GS-7, 0 otherwise

Table 5.3 – Highway Improvement, Full Model Summary

Term	Estimate	Estimate Confidence Interval			Factor Limits	
		Lower 95%	Upper 95%	Width	Min	Max
Intercept (β_0)	47.45	39.37	55.54	16.17	NA	
ADT	0.001	0.001	0.001	0.001	43	56,000
M_{1} - Has Median	-55.19	-80.35	-30.03	50.31	0	1
New Signal Count	18.53	6.76	30.30	23.54	0	1
g_5 - GS 6 (Urban Minor Arterial)	42.30	18.67	65.94	47.27	0	1
g_6 - GS 7 (Urban Collector Street)	59.66	29.56	89.76	60.20	0	1

5.2.2.2 Highway Improvement, Condensed Model

Again, the condensed model developed was less capable of explaining variability in construction duration (R^2_{adj} = 0.55 and 0.64 for the condensed and full models, respectively). Also, the condensed model contained fewer variables (3 versus 6 in the full model), making the model simpler and easier to use. The model developed and summary statistics are shown below.

$$DURATION = y = \beta_0 + \beta_1 \bullet CostEst + \beta_2 \bullet Area$$

Equation 5.4 – Highway Improvement Condensed Model

where: DURATION = Construction Duration (working-days) = *y*

β_o = Intercept

CostEst = *PCES* Construction Cost Estimate (2008 \$M)

Area = 1 if Medium Urban (50,000 – 199,999), 0 otherwise

Table 5.4 – Highway Improvement, Condensed Model Summary

Term	Estimate	Estimate Confidence Interval			Factor Limits	
		Lower 95%	Upper 95%	Width	Min	Max
Intercept (β_0)	37.40	26.02	48.77	22.75	NA	
a_2 - Medium Urban (50,000 - 199,999)	69.00	22.70	115.29	92.59	0	1
Construction Estimate (2008 \$M)	40.18	26.47	53.90	27.43	0.03 \$33,489	2.22 \$2,224,933

5.3 Duration-Influential Project Factors

While this research used VDOT specific data, the findings regarding the early-known project factors that influence construction duration are pertinent to the entire highway construction industry. The models developed were introduced previously. This section details each project factor, determined through this study, to have a statistically-significant relationship with highway construction duration. As a retrospective statistical regression analysis, one may only speculate about the meaning or cause behind the relationships. Here, the factors are explained and insight into sensitivity noted. Logical explanation of the findings is offered where warranted.

Overall, nine project factors were shown to influence construction duration. Three of these factors are continuous or quantitative variables and are discussed first here. The remaining six factors are categorical or qualitative.

5.3.1 Traffic Volume (ADT)

The highway traffic volume is the average daily traffic (ADT) count on the roadway under construction. The traffic volume is indicative of the size of the roadway, the incidental work that must be completed, as well as the conditions under which work is being performed. Incidental work includes managing traffic and constructing detours. Traffic volume also indicates the conditions under which work is performed, in that heavier traffic flows generally require increased safety measures and can influence the speed of work.

Traffic volume was recognized as a statistically-significant project factor in both full models prepared. The relationship with construction duration is additive, indicating that as traffic volume increases, construction duration also increases. The coefficient estimates for Traffic Volume ranged from 0.001 (for Highway Improvement projects) to 0.003 (for Full-Depth Section projects). This indicates that, all other variables constant, a 10,000 ADT increase in traffic volume equates to 10 additional days of construction duration for Highway Improvement projects, and 30 days for Full-Depth Section projects.

5.3.2 New Signal Count

The new signal count is the number of new traffic signals being installed on the project under construction. The new signal count is a primarily a measure of work added to the construction project. This work could be in the form of additional lanes, turning lanes, or be an indicator of increasing the dimensions of the roadway or median. The number of new signals installed would indirectly relate to the type of roadway under construction.

New signal count was found to be an early-known duration-influential project factor in both Full-Depth Section and Highway Improvement models. Its impact on Full-Depth Section project is much larger (parameter estimate of 141.5) than Highway Improvement projects (parameter estimate of 18.5).

5.3.3 Construction Cost Estimate

The construction cost estimate is the output of *PCES* and is, therefore, only analyzed in the condensed models. This estimate is based on historical VDOT data and in-house developed influence factors and base values for additional work [Williams et al. 2008a]. The construction cost is escalated to the projected construction completion date. Therefore, the records were inflated to January 1, 2008 values, so as to normalize the data values.

For both project types, the construction cost estimate is found to be statistically significant. This conclusion is justified as the full models indicate that multiple *PCES* factors influence construction duration. Therefore, it is logical that the culmination of those factors, in the form of the *PCES* cost estimate, would have a relationship with construction duration.

However, the relationship between duration and cost is not linear in both project types. Highway Improvement projects demonstrated a linear relationship with construction duration. Meanwhile, Full-Depth Section projects had a log-linear relationship with construction duration. The magnitude and sign of the construction cost estimate coefficients were validated previously [Williams et al. 2008a].

5.3.4 District

District is the first of the qualitative or categorical variables analyzed for a relationship with construction duration. VDOT is divided, geographically, into nine districts. As such, the district variable is an indication of where the construction project takes place within the state of Virginia. As a qualitative field, the district was handled as an indicator variable, where the base district is Bristol. The remaining eight districts were assigned a 'dummy' variable $d_1 - d_8$. Therefore, when a district dummy variable appears statistically-significant, the coefficient indicates the difference in construction duration over/under Bristol district. In other words, a positive coefficient for d_1 indicates that projects in district d_1 take longer to complete (statistically speaking). The converse is true for negative coefficient values.

Though all districts were studied in development of the full models, only $d_{8\ -\ Staunton\ District}$ was found to be statistically significant in Full-Depth Section projects. The negative coefficient (-29.1 working days) indicates that projects in Staunton district take less time (in working days) to complete than projects in Bristol district. This finding also indicates that other VDOT districts (or other project locations) do not have a statistically-significant relationship with construction duration different from that demonstrated by Bristol district.

5.3.5 Curb & Gutter

The curb and gutter indicator variable (CG_1) is a binary yes/no to indicate whether the project roadway incorporates curb and gutter or a shoulder. If the roadway has a curb and gutter, the coefficient value is multiplied by 1 and added to the total construction duration. If the roadway does not have a curb and gutter, $CG_1 = 0$ and the coefficient is cancelled out of the model. While studied in both project types, presence of curb and gutter was only statistically-significant in Full-Depth Section projects. It is important to note here that this does not indicate that the presence of curb and gutter for Highway Improvement projects does not influence construction duration, only that the data analyzed does not prove a statistically-significant relationship. Subject matter knowledge suggests that the presence of curb and gutter for Highway Improvement projects (such as resurfacing projects) would require additional milling and prep work that roadways with only a shoulder would not. Another possibility is that the increased work associated with building and reconstructing curbs after resurfacing operations balance out the impact of curb and gutter. As such, there is not sufficient evidence to include/reject parameters beyond statistical significance.

In Full-Depth Section projects, the presence of curb and gutter has a positive (84 working day) impact on construction duration. This increase in construction duration over projects with only a shoulder can be intuitively explained. Construction of curb and gutter systems generally include additional earthwork, entryway, and concrete placement work that would be expected to increase construction duration. Also, such activities would be expected to take longer than shoulder construction activities (which can generally proceed with construction of the roadway itself) and therefore, would not increase project duration in the same fashion.

5.3.6 Median

Like curb and gutter, the median indicator variable (M_I) is a binary yes/no to indicate whether the project roadway incorporates a median. Again, if the roadway has a median, the coefficient is multiplied by 1 and added to the total construction duration. Also, the lack of median cancels the coefficient.

The presence of median is influential on construction duration in both Full-Depth Section and Highway Improvement projects. Also, the coefficient value is negative in each project type, indicating that the presence of median tends to reduce the overall construction duration. While less intuitive, this finding can be justified. It is expected that the presence of median in the roadway, for either project type, could have the effect of expanding the allowable work area, decongesting the project work site, and allow for traffic to be shifted away from construction activities, improving safety and reducing time lost due to managing traffic. Combined with the positive influences in the models, it is expected that this finding is valid.

5.3.7 Area Location Name

The area location name is a qualitative variable used to indicate the setting under which the construction project is completed. VDOT classifies projects by one of four areas: Rural, Small Urban (5,000 – 49,999), Medium Urban (50,000 – 199,999), and Large Urban (200,000 +). Rural is used as the base area. Therefore, statistically-significant coefficients indicate influence of urbanized settings over rural settings.

While analyzed in both condensed models, the area was only found to be statistically-significant in Highway Improvement projects. Also, only one of the three indicator areas ($a_{2\text{ – Medium Urban (50,000 – 199,999)}}$) demonstrated a difference on construction duration over the rural setting. According to this analysis, medium urban settings tend to take longer than projects in rural settings, all other variables held constant. The lack of influence demonstrated by small and large urban projects is not immediately logical, but can be justified using knowledge of the data set and subject matter knowledge.

The lack of influence demonstrated by small urban projects is most readily explained as the Highway Improvement regression analysis data did not have an entry from a small urban setting. As such, a limitation of the condensed model is that it is not suited to predicting construction duration for projects classifying as small urban settings. The lack of influence exhibited by large urban projects is less obvious. It is expected that projects taking place in large urban or heavily populated areas are able to make use of alternate traffic routes and detours which could negate the impact of traffic maintenance and safety. Additionally, it is likely that special measures would be taken to shorten construction duration in congested or urbanized areas.

5.3.8 State Highway System

The state highway system is a qualitative description of the roadway primacy. VDOT uses five categories: Primary Arterial Network, Primary System, Secondary System, Urban System, and Interstate System. The state highway system is analyzed in the condensed model analysis as a similar metric is not accounted for in the *PCES* Construction Cost data.

One state highway system level was found to influence construction duration: Interstate System. This occurred in the Full-Depth Section condensed model and demonstrated a positive impact on construction duration, above the base system category (Primary). Also included in the Full-Depth Section condensed model is an interaction effect between interstate projects and the construction cost estimate.

In essence, indicator variables increase or decrease the intercept of the regression line. This is due to the nature of the dummy variables which add the coefficient value (or 1 times the coefficient value) when the indicator category is met. In instances where an interaction term appears between indicator and nonlinear continuous variables, though, the slope of the second regression line is altered. As the interaction between highway system (interstate) and construction estimate is negative, the two lines will converge and possibly intersect. This indicates that as size (in terms of cost) increases, the impact on construction duration continually decreases.

5.3.9 Geometric Design Standard

The final statistically-significant, early-known project factor determined to have a relationship with construction duration is geometric design standard. VDOT classifies projects as one of nine standards, indicative of the type of roadway being constructed. Geometric design standard values include principal arterial, minor arterial, collector, and local roads for both urban and rural settings. Again, a base standard is selected and dummy variables generated for each of the remaining design standards.

The geometric design standard was found to be influential on construction duration. In fact, all eight indicator geometric design standards were found to influence construction duration in Full-Depth Section projects. All standards evidenced a reduction of construction duration from the base (Rural Principal Arterial). There was also an interaction effect between rural local road and traffic volume deemed statistically-significant.

For the Highway Improvement full model, two geometric design standards (GS 6 - Urban Principal Arterials and GS 7 – Urban Minor Arterials) were found to positively affect construction duration. GS 8 – Urban Local Street was not represented in the data analyzed. Therefore, these projects are not adequately modeled using the full Highway Improvement model. However, these projects are not outside the prediction capabilities of the condensed model as it uses a generalized construction cost estimate relationship with construction duration.

This difference in findings cannot be readily explained outside of the data set used for analysis. Given the higher number of variables in the Full-Depth Section full model, it is likely that the coefficients serve to reduce the impact demonstrated by other indicators like new signal count. The Highway Improvement findings appear more logical as the urban projects represented were shown to positively influence construction duration.

5.4 Application

The results and underlying methodology of this work and those documented previously [Williams et al. 2008a] are applicable to numerous construction industry entities. Most specifically, the research applies to VDOT. However, the highway construction industry, and general construction industry also see benefits and applicability of the research.

This section describes how the outcomes of this work (concept, methodology, duration-influential factors, and prediction equations) can be applied by various members of the construction industry. While the concept behind the work is applicable to all construction owners, the research methodology, duration-influential factors, and prediction equations can be applied in either principle or detail by various entities.

The work is directly applicable to construction owners. Public and private owner applications are discussed herein. The table below summarizes the entities to which the work applies. The remainder of this section describes how each of these entities may apply this work.

Table 5.5 – Application Mapping

Construction Entity	Research Concept	Research Methodology		Duration-Influential Factors		Prediction Equations	
		Principle	*Detail*	*Principle*	*Detail*	*Principle*	*Detail*
VDOT	✓	✓	✓	✓	✓	✓	✓
State Highway Agencies	✓	✓	✓	✓	✗	✓	✗
Other Public Owners	✓	✓	✗	✗	✗	✗	✗
Private Owners	✓	✓	✗	✗	✗	✗	✗

5.4.1 VDOT Application

The statistical regression analysis conducted made use of historical VDOT data. As such, the coefficient estimates and models (equations) developed are specific to projects constructed in Virginia. As seen in the above table, VDOT can directly apply the concept, methodology, information regarding the duration-influential project factors, and the prediction equations developed to model construction duration.

The importance of accurately predicting construction duration early in the project design process cannot be overstated. As a habitual owner, VDOT has a need for systematically predicting construction duration and estimating contract duration. The models developed are tailored to use by VDOT for understanding the early-known project factors that influence construction duration and predicting construction duration.

VDOT recognizes the need for such research and the value of accurately and reliably predicting highway construction duration early in the design lifecycle. The Location & Design (L&D) division of VDOT manages the construction design and programming across the state of Virginia. As such, L&D sees potential in using the models to predict construction duration for the purposes of improving financing decisions (accounting for cost inflation), understanding and mitigating the impacts of design delays on construction, and capturing and supplementing experienced personnel knowledge in the area of duration estimating [Williams et al. 2008c]. Meanwhile, the VDOT Construction Division manages the more detailed construction planning, estimating, and scheduling functions, once the project has entered the design process. As such, the division recognizes the potential for using the models developed to analyze various design alternatives for construction and to predict construction duration during the initial phases of design [Williams 2008c]. This remainder of this section describes the appropriate use of and outcomes from the models developed as applied by VDOT.

5.4.1.1 Model Use

Both *DWMIP* and *PCES* data is required to use the models developed. However, this information is known early in the project design lifecycle. The models produce construction duration in working-days, indicative of the number of days required by the contractor to complete the construction activities. This duration does not include an allowance for adverse weather days, weekends, or holidays. Further, the duration does not include an allowance for the time of year the project is begun or during which work is performed.

Therefore, a calendar must be applied to the construction duration calculated using the models. At this early stage of project design, a critical path method (CPM) scheduling tool is not necessary. Instead, the calendar or contract duration may be calculated by considering the estimated start date, potential adverse weather, and non-working days in each month during which construction work will be performed. VDOT currently uses the *Duralator* for such purposes [Hildreth 2005].

Also, the scheduler must exercise engineering and professional judgment in using the calculated construction duration. The duration predicted by the model is based on historical VDOT data and the relationships between project factors and construction duration. As such, unique projects require special attention beyond what is permitted by the models. Further, all projects should be considered individually and the construction duration prediction assessed by the user.

5.4.1.2 Model Data Source

The data required for the regression models are obtained from *DWMIP* and *PCES*. Fields include both quantitative or continuous fields, such as Traffic Count (ADT), as well as qualitative or categorical fields like Geometric Design Standard. The tables below list the data fields in each model, the source of each data field, and transforms necessary to appropriately use the data in the model.

Table 5.6 – Full-Depth Section, Full Model Data Source

Term	Source (System - Field Name)	Transform
ADT	*PCES* - ci_current_adt	None
New Signal Count	*PCES* - ci_new_signal_count	None
CG_{1} - Has Curb & Gutter	*PCES* - ci_shoulder_or_cg	1 if has C&G, 0 otherwise
M_{1} - Has Median	*PCES* - ci_median_type	1 if has Median, 0 otherwise
d_{8} - Staunton District	*DWMIP* - District	1 if Staunton, 0 otherwise
g_{1} - GS 2 (Rural Minor Arterial)	*PCES* - ci_geo_std	1 if GS 2, 0 otherwise
g_{2} - GS 3 (Rural Collector Road)	*PCES* - ci_geo_std	1 if GS 3, 0 otherwise
g_{3} - GS 4 (Rural Local Road)	*PCES* - ci_geo_std	1 if GS 4, 0 otherwise
g_{4} - GS 5 (Urban Principal Arterial)	*PCES* - ci_geo_std	1 if GS 5, 0 otherwise
g_{5} - GS 6 (Urban Minor Arterial)	*PCES* - ci_geo_std	1 if GS 6, 0 otherwise
g_{6} - GS 7 (Urban Collector Street)	*PCES* - ci_geo_std	1 if GS 7, 0 otherwise
g_{7} - GS 8 (Urban Local Street)	*PCES* - ci_geo_std	1 if GS 8, 0 otherwise
g_{8} - Interstate	*PCES* - ci_geo_std	1 if Interstate, 0 otherwise
g_3 x ADT	NA	If GS 4: ADT * 1, ADT * 0 otherwise

Table 5.7 – Full-Depth Section, Condensed Model Data Source

Term	Source (System - Field Name)	Transform
h_{4} - Interstate	*DWMIP* - State Highway System	1 if Interstate, 0 otherwise
ln Construction Estimate	*PCES* - cn_estimate	ln[(cn_estimate adjusted to Jan. 1, 2008) / $1,000,000]
h_4 x ln Construction Estimate	NA	If Interstate: 1 * ln[(cn_estimate adjusted to Jan. 1, 2008) / $1,000,000]; 0 * ln[(cn_estimate adjusted to Jan. 1, 2008) / $1,000,000] otherwise

Table 5.8 – Highway Improvement, Full Model Data Source

Term	Source (System - Field Name)	Transform
ADT	*PCES* - ci_current_adt	None
M_{1} - Has Median	*PCES* - ci_median_type	1 if has Median, 0 otherwise
New Signal Count	*PCES* - ci_new_signal_count	None
g_{5} - GS 6 (Urban Minor Arterial)	*PCES* - ci_geo_std	1 if GS 6, 0 otherwise
g_{6} - GS 7 (Urban Collector Street)	*PCES* - ci_geo_std	1 if GS 7, 0 otherwise

Table 5.9 – Highway Improvement, Condensed Model Data Source

Term	Source (System - Field Name)	Transform
a_{2} - Medium Urban (50,000 - 199,999)	*DWMIP* - Area Location Name	1 if Urbanized (50,000 - 199,999), 0 otherwise
Construction Estimate	*PCES* - cn_estimate	(cn_estimate adjusted to Jan. 1, 2008) / $1,000,000

As seen, most data fields are taken from *PCES*. However, *DWMIP* data also has a positive effect on the explanatory powers of the models, and are included in three of the four models developed.

5.4.1.3 Construction Duration Prediction

This section offers examples of how the construction duration can be predicted by VDOT using the regression models developed. This example includes retrieving and inputting data into the models, calculating the construction duration, and a description of using the predicted construction duration.

As the condensed models proved better candidates for prediction of construction duration, the examples will make use of the condensed models for Full-Depth Section and Highway Improvement project types. Future work should continually reevaluate the models and the relationships identified as underlying construction means and methods and technologies evolve. However, at present, the full models developed are included to identify and describe the duration-influential project factors while the condensed models are recommended for construction duration prediction.

5.4.1.3.1 Full-Depth Section Construction Duration Prediction

The condensed models are exemplified here to demonstrate the prediction of construction duration and the output of the models developed. The project data used is fictitious and intended for representation purposes only. However, note that the factor values used are within the limits of the data analyzed.

The condensed Full-Depth Section model includes three variables: ln(*PCES* Construction Cost Estimate (2008 \$M)), $h_{4-Interstate}$, and an interaction term between the two variables. The presence of the indicator variable ($h_{4-Interstate}$) evidences a second regression line. The interaction effect shows that the two lines are not of equal slope. The regression lines created by the condensed model are shown in Figure 5.4.

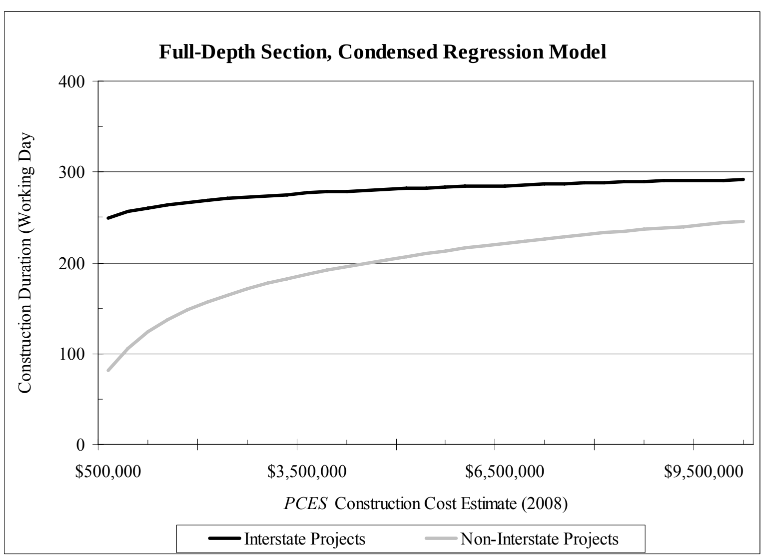

Figure 5.4 – Full-Depth Section, Condensed Model Regression Lines

As seen, the regression lines converge at higher values of the *PCES* construction cost estimate. This indicates that as the overall cost of constructing the project increases, the effect of the roadway being an interstate on construction duration decreases. These lines ultimately converge, but well beyond the limits of the data analyzed.

Predicting construction duration can be done by inserting the appropriate values into the regression equation or using the plot above. An example of using the regression equation to predict construction duration for Full-Depth Section projects for both interstate and non-interstate projects is presented below.

Example: Imagine the need for a construction duration estimate for an interstate project, with a *PCES* construction cost estimate of $5,750,000. Use Equation 5.2 (presented below).

$$DURATION = y = \beta_0 + \beta_1 \bullet CostEst + \beta_2 \bullet HWSystem + \beta_3 HWSystem \times CostEst$$

where: DURATION = Construction Duration (working-days) = y

β_o = Intercept = 118.2

β_1 = 54.8

β_2 = 140.1

β_3 = -40.6

HWSystem = 1 (since interstate)

CostEst = ln(*PCES* Construction Cost Estimate (2008 $M)) = ln($5,750,000 / $1,000,000) = 1.75

Inserting the appropriate values into Equation 5.2:

$$DURATION = 118.2 + (54.8 \times 1.75) + (140.1 \times 1) + (-40.6 \times 1 \times 1.75)$$
$$DURATION = 283 \ working\text{-}days$$

Note, had the project not been an interstate, the equation above would reduce to:

$$DURATION = 118.2 + (54.8 \times 1.75) + (140.1 \times 0) + (-40.6 \times 0 \times 1.75)$$
$$DURATION = 214 \ working\text{-}days$$

Again, these values can be taken from the plot in Figure 5.4 as well. It is important to note that the construction duration is in working-days. The scheduler would therefore need to apply a calendar to the duration predicted in order to determine the number of calendar days needed to complete the project. As previously stated, the calendar duration will depend on non-working days and the time of year in which the project is begun.

5.4.1.3.2 Highway Improvement Construction Duration Prediction

Because of its relatively similar explanatory power and ease of use, the condensed model is suggested for predicting construction duration for Highway Improvement projects. As before, the plot of the regression lines is presented as well as an example of model use.

The Highway Improvement condensed model has two variables: *PCES* Construction Cost Estimate (2008 \$M) and $a_{2\text{ – Medium Urban (50,000 – 199,999)}}$. The indicator ($a_{2\text{ – Medium Urban (50,000 – 199,999)}}$) shows that there are two regression lines of equal slope, separated vertically by the coefficient of the indicator variable. Figure 5.5 below shows the two regression lines for the Highway Improvement condensed model.

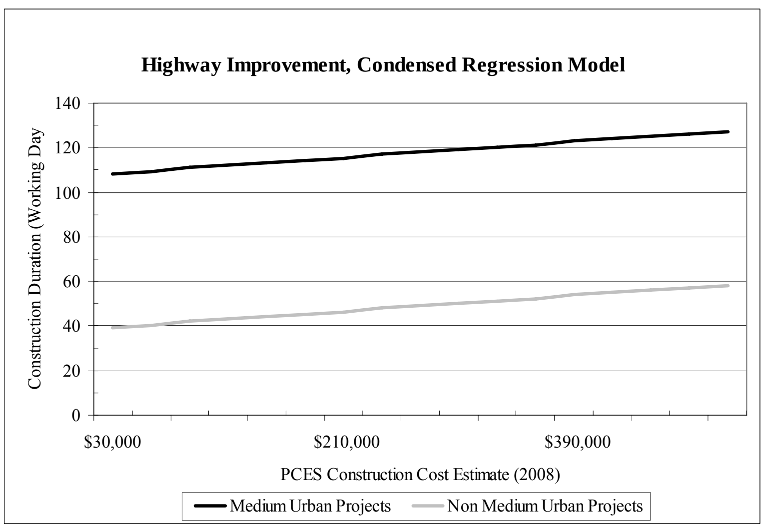

Figure 5.5 – Highway Improvement, Condensed Model Regression Lines

Unlike the Full-Depth Section lines, the Highway Improvement regression lines do not converge (since there is no interaction term). These regression lines indicate that the project being constructed in the Medium Urban setting (population 50,000 – 199,999) has the effect of increasing the construction duration by 69 working days.

Predicting construction duration can be done by inserting the appropriate values into the regression equation or using the plot above. An example of using the regression equation to predict construction duration for Highway Improvement projects for projects located in the medium urban and non medium urban settings is presented below.

Example: Imagine the need for a construction duration estimate for a medium urban Highway Improvement project, with a *PCES* construction cost estimate of \$230,000. Use Equation 5.4 (presented below).

$$DURATION = y = \beta_0 + \beta_1 \bullet CostEst + \beta_2 \bullet Area$$

where: DURATION = Construction Duration (working-days) = y

β_o = Intercept = 37.4

β_1 = 40.2

β_2 = 69.0

CostEst = *PCES* Construction Cost Estimate (2008 \$M) = \$230,000 / \$1,000,000 = 0.23

Area = 1 (since Medium Urban)

Inserting the appropriate values into Equation 5.4:

$$DURATION = 37.4 + 40.2 \times 0.23 + 69.0 \times 1$$
$$DURATION = 116 \textit{ working - days}$$

Note, had the project not been in the medium urban setting, the equation above would reduce to:

$$DURATION = 37.4 + 40.2 \times 0.23 + 69.0 \times 0$$
$$DURATION = 47 \textit{ working - days}$$

Again, these values can be taken from the plot in Figure 5.5 as well and that the construction duration is in working-days. The scheduler would therefore need to apply a calendar to the duration predicted in order to determine the number of calendar days needed to complete the project.

5.4.1.3.3 Historical Project Data Examples

Beyond hypothetical examples, the results of this work are intended to aid the highway construction industry in understanding construction duration-influential factors and their relationship to construction duration. Further, VDOT may immediately apply the results of the work to their current initial design duration estimating practices and procedures. To further exemplify the use of the models, three case study projects are presented here. The projects are from two separate VDOT districts, are from both Full-Depth Section and Highway Improvement project types, and demonstrate the use of both the full and condensed prediction models.

The case studies below also present a significant assumption of the regression analysis performed. Williams [2008a] used data from two sources to identify duration-influential factors and model construction duration. The contract level data used was obtained from a historical, as-built database of project information which included the dependent duration variable. The project level data used was obtained from an as-designed cost estimating system. Therefore, the regression analysis assumes that the dependent construction duration is reflective of the independent design factor values.

This assumption is justified in a number of ways. First, the *PCES* factors analyzed were derived from the most current estimate generated by the system. As such, early-design changes (those that tend to be major scope changes) are accounted for in the data collected. Second, of the early-known factors which influence construction duration, most have fixed values that will not change in the latter stages of design or construction. For instance, the project location, traffic volume, and design standard are fixed project aspects. Finally, as an early-design prediction tool, the models developed cannot encompass or account for late design or construction changes or nuances of the work. One of the limitations of a retrospective analysis is the industry 'photograph' from which analysis is based. The work intends to aid in early-design prediction of construction duration and it is felt that discrepancies between as-planned and as-built project factor values are not likely to alter the predicted construction duration beyond the range expected of a conceptual design estimate.

The case studies below use actual project data included in the regression analysis. For those factors whose as-built value may differ from as-planned, both the design and actual values are listed. The estimated and actual contract durations extracted from the *DWMIP* system are shown to identify design changes that may have occurred during construction. The predicted construction durations are based on the design factor values. Both full and condensed models are demonstrated for each project. A discussion of the results follows.

Route 638 Reconstruction

Project Information

Project Type:	Full-Depth Section
Work Type:	Reconstruction
Location:	Lynchburg District
Setting:	Rural
Functional Classification:	Major Collector
Start Date:	January 1, 2006
End Date:	October 31, 2006
PCES Estimate (2006):	$847,000
PCES Estimate (2008):	$945,882

Condensed Model

Factor	Coeff.	Project Value	
		Design	*Actual*
Intercept (β_0)	118.20	-	
$h_{4\text{ - Interstate}}$	140.13	0	
ln Con. Est. (2008 $M)	54.78	-0.06	
h_4 x ln Con. Est. (2008 $M)	-40.62	0	
Estimated Contract Duration (Calendar Days):			307
Actual Contract Duration (Calendar Days):			299
Actual Construction Duration (Working Days):			147
Predicted Construction Duration (Working Days):			115

Full Model

Factor	Coeff.	Project Value	
		Design	*Actual*
Intercept (β_0)	271.35	-	
ADT	0.003	645	
New Signal Count	141.47	0	0
$CG_{1\text{ - Has Curb \& Gutter}}$	84.03	No C&G	No C&G
$M_{1\text{ - Has Median}}$	-130.59	No Median	No Median
$d_{8\text{ - Staunton District}}$	-29.10	0	
$g_{1\text{ - GS 2 (Rural Minor Arterial)}}$	-250.38	0	
$g_{2\text{ - GS 3 (Rural Collector Road)}}$	-133.10	1	
$g_{3\text{ - GS 4 (Rural Local Road)}}$	-201.06	0	
$g_{4\text{ - GS 5 (Urban Principal Arterial)}}$	-172.25	0	
$g_{5\text{ - GS 6 (Urban Minor Arterial)}}$	-347.86	0	
$g_{6\text{ - GS 7 (Urban Collector Street)}}$	-97.91	0	
$g_{7\text{ - GS 8 (Urban Local Street)}}$	-228.90	0	
$g_{8\text{ - Interstate}}$	-298.47	0	
g_3 x ADT	0.12	0	
Estimated Contract Duration (Calendar Days):			307
Actual Contract Duration (Calendar Days):			299
Actual Construction Duration (Working Days):			147
Predicted Construction Duration (Working Days):			141

Figure 5.6 – Historical Data Case Study #1

Route 604 Reconstruction

Project Information

Project Type:	Full-Depth Section
Work Type:	Reconstruction
Location:	Lynchburg District
Setting:	Rural
Functional Classification:	Local Road
Start Date:	September 20, 2004
End Date:	August 9, 2005
PCES Estimate (2005):	$1,421,000
PCES Estimate (2008):	$1,833,299

Condensed Model

Factor	Coeff.	Project Value *Design*	Project Value *Actual*
Intercept (β_0)	118.20	-	
$h_{4\text{ - Interstate}}$	140.13	0	
ln Con. Est. (2008 $M)	54.78	0.61	
h_4 x ln Con. Est. (2008 $M)	-40.62	0	
Estimated Contract Duration (Calendar Days):			324
Actual Contract Duration (Calendar Days):			323
Actual Construction Duration (Working Days):			152
Predicted Construction Duration (Working Days):			152

Full Model

Factor	Coeff.	Project Value *Design*	Project Value *Actual*
Intercept (β_0)	271.35	-	
ADT	0.003	450	
New Signal Count	141.47	0	0
$CG_{1\text{ - Has Curb \& Gutter}}$	84.03	No C&G	No C&G
$M_{1\text{ - Has Median}}$	-130.59	No Median	No Median
$d_{8\text{ - Staunton District}}$	-29.10	0	
$g_{1\text{ - GS 2 (Rural Minor Arterial)}}$	-250.38	0	
$g_{2\text{ - GS 3 (Rural Collector Road)}}$	-133.10	1	
$g_{3\text{ - GS 4 (Rural Local Road)}}$	-201.06	0	
$g_{4\text{ - GS 5 (Urban Principal Arterial)}}$	-172.25	0	
$g_{5\text{ - GS 6 (Urban Minor Arterial)}}$	-347.86	0	
$g_{6\text{ - GS 7 (Urban Collector Street)}}$	-97.91	0	
$g_{7\text{ - GS 8 (Urban Local Street)}}$	-228.90	0	
$g_{8\text{ - Interstate}}$	-298.47	0	
g_3 x ADT	0.12	0	
Estimated Contract Duration (Calendar Days):			324
Actual Contract Duration (Calendar Days):			323
Actual Construction Duration (Working Days):			152
Predicted Construction Duration (Working Days):			140

Figure 5.7 – Historical Data Case Study #2

Route 1566 Reconstruct Right Turning Lane

Project Information

Project Type:	Highway Improvement
Work Type:	Safety/Traffic Operations
Location:	Northern Virginia District
Setting:	Urban (200,000 +)
Functional Classification:	Minor Collector
Start Date:	April 25, 2006
End Date:	November 7, 2006
PCES Estimate (2006):	$602,000
PCES Estimate (2008):	$664,930

Condensed Model

Factor	Coeff.	Project Value	
		Design	*Actual*
Intercept (β_0)	37.40	-	
$a_{2\text{ - Medium Urban (50,000 - 199,999)}}$	69.00	0	
Con. Est. (2008 $M)	40.18	0.665	
Estimated Contract Duration (Calendar Days):			198
Actual Contract Duration (Calendar Days):			196
Actual Construction Duration (Working Days):			89
Predicted Construction Duration (Working Days):			65

Full Model

Factor	Coeff.	Project Value	
		Design	*Actual*
Intercept (β_0)	47.45	-	
ADT	0.001	19,000	
New Signal Count	18.53	1	1
$M_{1\text{ - Has Median}}$	-55.19	Has Median	Has Median
$g_{5\text{ - GS 6 (Urban Minor Arterial)}}$	42.30	1	
$g_{6\text{ - GS 7 (Urban Collector Street)}}$	59.66	0	
Estimated Contract Duration (Calendar Days):			198
Actual Contract Duration (Calendar Days):			196
Actual Construction Duration (Working Days):			89
Predicted Construction Duration (Working Days):			73

Figure 5.8 – Historical Data Case Study #3

As seen in the case studies above, there are few project factors which may differ between design and construction. Also, the estimated and the actual contract durations are nearly the same, indicating that major design or constructability changes did not likely occur during construction. In the projects cited, these factor values remained constant throughout design and construction. Also, as they make use of different factors and project aspects, the full and condensed models provided different construction duration predictions for each project. By chance, the construction durations predicted were all less than the actual construction duration for the projects. As a tenet of regression analysis and the principle of randomly distributed errors about the mean, it is expected that for each prediction under the actual duration, there is a corresponding value above the actual duration. Finally, the difference between predicted and actual construction durations are generally small, and in one case, zero. Where differences exist, the magnitude of the difference is commensurate with that expected of a conceptual design estimate.

5.4.2 Highway Construction Application

The highway construction industry does not currently have an understanding of the early-known project factors which influence construction duration [Williams et al. 2008c]. Williams et al. [2008a] established and quantified those factors and their relationship with construction duration. In doing so, multiple linear regression (MLR) models were developed using VDOT construction data. While the MLR model coefficient values are specific to VDOT, the research concept, methodology, and early-known project factors which are influential on construction duration are pertinent to the larger highway construction industry in the United States and abroad (see Table 5.5). This section discusses the applicability of the research to U.S. state highway agencies (SHAs).

5.4.2.1 Concept and Methodology

The research concept and methodology employed is applicable to SHAs outside of VDOT. A review of SHAs across the U.S. found that, while 80% of SHA's recognize a need for early, systematic, historically-based construction duration estimates, few have available tools or procedures to complete such early estimates [Williams et al. 2008c].

Existing tools and procedures included CPM schedules with multiple assumptions, personnel experience, or comparison of similar projects [Williams et al. 2008c]. Despite the lack of existing procedures or tools, SHAs recognize that more accurate and reliable duration estimates completed early in project design have the potential to enhance planning efforts, improve funding knowledge and financial planning efforts, improve public convenience and safety, enhance resource planning, improve the cost estimating process, and allow them to consider multiple design alternatives.

The details of the research methodology are defined elsewhere [Williams et al. 2008a, Williams et al. 2008b], however, the basic steps include: data collection, data treatment, regression analysis, and model validation. Here, the significant findings from each of these basic categories are highlighted. These findings are intended to illustrate the similarities between the methodology employed using VDOT data and how similar methodology can be adapted and applied by other SHAs.

The data collection process presented a unique challenge as many SHAs do not currently collect and maintain the type of early-known project factor data used in this research. Two SHAs having both contract and project level data were identified and VDOT selected as the optimal data source. The data analyzed was extracted from *DWMIP* and *PCES* systems. *DWMIP* maintains contract level data of the type that most SHAs regularly maintain. These fields include the project start and end dates, location, roadway type, and setting. *PCES* is a construction cost estimating system used by VDOT to prepare a cost estimate during the initial design phases. Such a system appears unique to VDOT. To collect the data for this study, projects were matched via uniform project code (UPC) which was common in both systems.

The data treatment process included both topical data analysis for erroneous or incomplete entries and seasonal adjustment to account for non-working days. The seasonal adjustment process involved developing historically-typical adverse weather calendars for each month from National Oceanic and Atmospheric Administration (NOAA) weather data and impact factors developed through a review of existing literature and subject-matter knowledge. These calendars were applied to the contract duration and the historically-typical construction duration calculated. This adjustment removed nonworking days from weather, weekends, and holidays and has the effect of removing 'noise' from the data. The calculated historically-typical construction duration was used as the response variable for the regression analysis.

The regression analysis completed used a stepwise regression approach to identify statistically-significant project factors and quantify the coefficient values for each. This approach was used to create a full and condensed model for both Full-Depth Section and Highway Improvement project types. The full models were constructed through analysis of both *DWMIP* and individual *PCES* factors. The condensed models were developed through analysis of *DWMIP* factors and the *PCES* construction cost estimate, which is developed from the aforementioned individual *PCES* factors. The model validation process considered correlation, coefficient validity, and model stability.

5.4.2.2 Highway Construction Factors of Interest

The construction project factors known to influence construction duration are identified and described in section 5.3. The factors identified varied for the full and condensed models as well as the two project types (Full-Depth Section and Highway Improvement).

The factors identified as influential on highway construction duration include location (district), traffic count, new signal count, construction cost estimate, project setting (area location name), roadway primacy (state highway system), roadway type (geometric design standard), as well whether the project has a median and curb and gutter. The factors identified are not unique to Virginia projects. Instead, these factors, or a close resemblance to them, are common to highway construction projects across the U.S. and abroad.

These factors offer the SHA a basic understanding of the early-known construction project factors which influence duration and their significance to construction duration. This provides other SHAs a starting point for future analysis to quantify the coefficient values in their own states. Further, the factors identified in Virginia were limited to the available data from *DWMIP* and *PCES*. As such, there are undoubtedly other project factors, known early in project design, which influence construction duration. SHAs can build on the list of known factors to identify additional influential factors in their own states. As shown in Table 5.5, the highway construction duration-influential factors are applicable, in principle but not detail, to all SHAs.

5.4.2.3 Highway Construction Duration Prediction Equations

The prediction models developed are applicable to other SHAs in principle, but not detail. Because the equations are built from VDOT data, the coefficient values, sign, and relationship are not directly applicable to other SHAs. However, knowledge regarding the sign, magnitude, and function of the relationships between highway construction duration and these factors is principally relevant to other SHAs.

This information is the first entry in the SHA body of knowledge regarding the form of the relationships between highway construction duration and early-known factors. Also, because of the likelihood of common factors across highway construction, other SHAs may build on this knowledge as a basis for understanding their own duration-influential project factors.

5.4.3 Construction Industry Application

The concept behind this work and the methodologies developed through the work are applicable to the general construction industry. As construction duration prediction is a project planning function, the construction owner sees the most immediate benefits to the work. This section discusses how the concept behind the research and the methodologies developed are applicable to the general construction industry. Both public and private owner needs and potential for application are considered herein.

5.4.3.1 Research Concept

There are various reasons for having an accurate, reliable construction duration estimate early in the design process. Chief among these reasons is to improve knowledge of and decisions about the project finances and funding [Williams et al. 2008c]. Public owners are required to establish long-term budgets for submission to legislators. Private owners must undertake the process of securing funding from a bank or through the sales of company stocks/bonds. Such funding must be established early in project design, soon after the concept and initial design is established, and must be maintained throughout the life of the construction project. As such, the duration of the construction process is critical to the owner. The completion of construction and transfer of facility to the owner, coincides with the beginning of the return on investment for most private owners, when capital expended for construction may begin to be recouped.

Further, accurate and reliable construction duration estimates completed early in the project design lifecycle is one of the key components to the front-end planning process. Adequate front-end planning is crucial to successful project execution. The Construction Industry Institute (CII) states that:

> *"effective front end planning can improve performance of overall capital costs by 10 percent or greater, reduce schedule by seven percent or better, and reduce change orders by five percent"* [CII 2006].

Historically, however, initial construction duration estimates have been based on personnel experience and comparison to similar projects [Williams et al. 2008c]. Work that has been performed to establish construction duration early in the project design lifecycle has typically been based on high-level contractual details. Furthermore, the initial duration estimate is often unrefined until completion of the project design (see Figure 5.9).

This work has established the methodologies by which the construction project duration may be predicted early in the project design lifecycle, based on historical construction data, using known duration-influential project factors, and owner-specific relationships between the project factors and construction duration. The research described here makes use of both contract and project level design information to predict construction duration throughout initial design (see Figure 5.10).

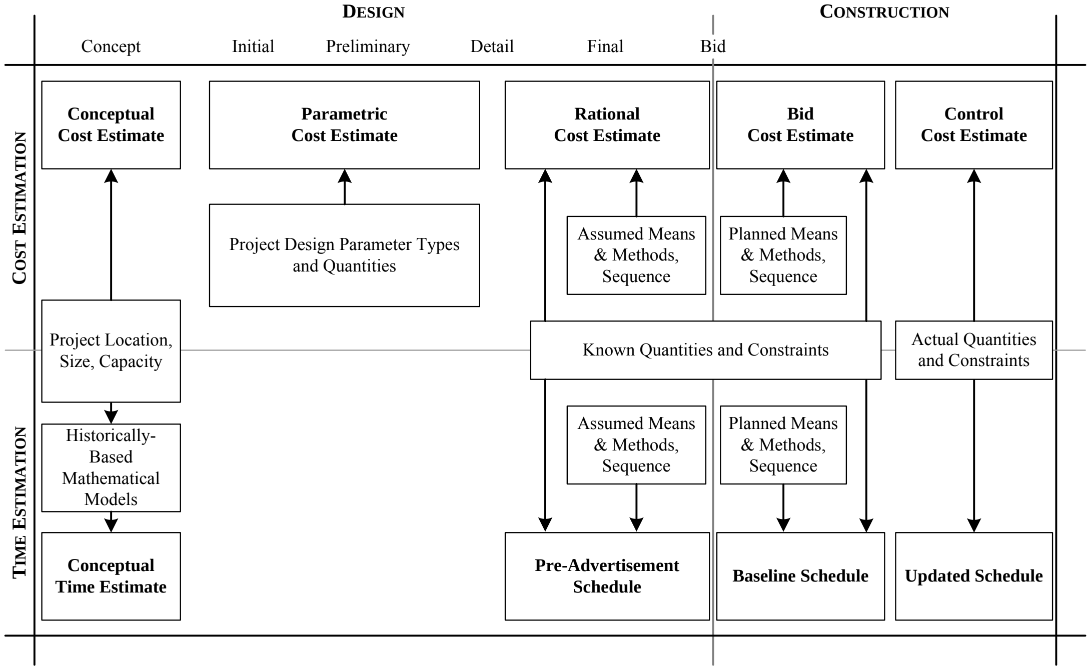

Figure 5.9 – Existing Construction Industry Project Development Structure

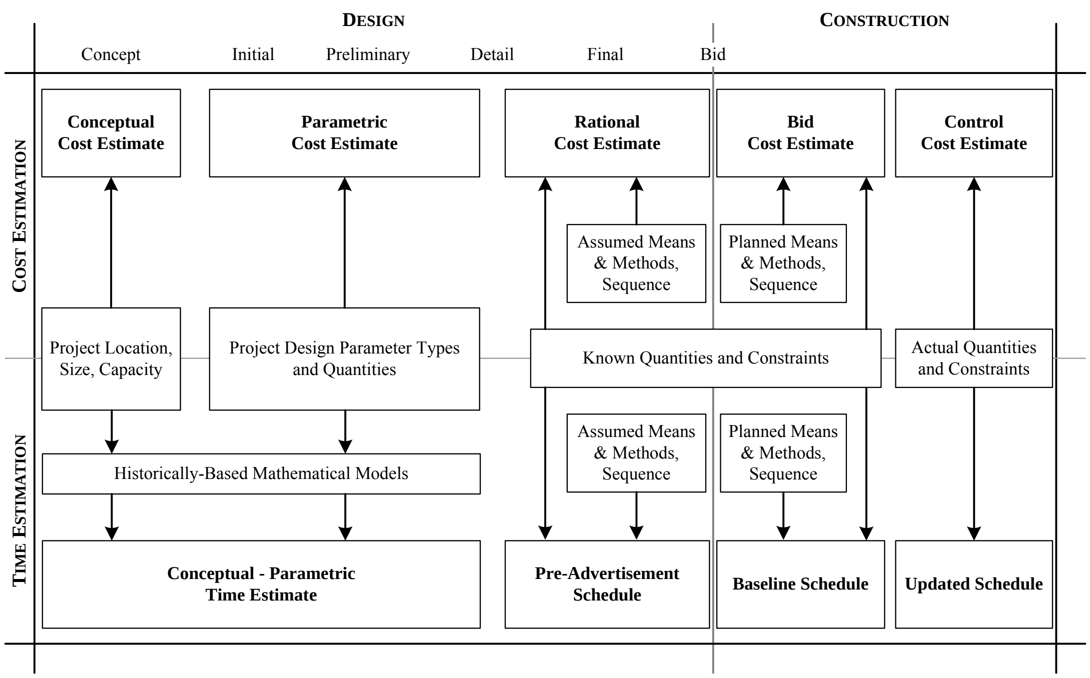

Figure 5.10 – Proposed Construction Industry Project Development Structure

While the project owner or design professional is in the best position to apply the research for design-bid-build delivered projects, the applicability of the research concept extends to contractors for design-build projects. Design-build projects often begin construction work early in design, once initial design is complete. In such delivery methods, the contractor works directly with the owner to complete the design and construct the project. As such, the contractor is able to recommend design and constructability changes to enhance project quality and timeliness. Therefore, design-build contractors recognize many of the same needs for early, accurate, historically-based duration estimates as design-bid-build project owners.

To further describe the applicability of the research concept and the principles of the methodology, the following sections present a review of the current state of practice found among public and private construction owners outside of highway construction. These sections describe how the organizations currently estimate construction duration early in the design lifecycle, the need for such estimates, and benefits derived from establishing an accurate and reliable construction duration during initial design.

5.4.3.1.1 Public Construction Owners

Public construction owners exist in every construction industry sector. Public construction owners are entities which oversee construction for the general public, constituency, or geographical locality. Public construction owners make use of taxes, fees, or other funds provided by constituents to complete projects intended for public use. Generally, these entities have large capital budgets and are habitual owners – completing multiple construction projects each fiscal year. As such, public construction owners are stewards of the public funds and have increasingly demonstrated an interest in improving the project planning process. This work seeks to enhance the body of knowledge regarding prediction of construction duration early in the project development lifecycle. While this work made use of SHA data (specifically VDOT data), the concept behind the research is relevant to all public construction owners.

Interest in the work was discussed in an interview with a member of the Virginia Tech (VT) Planning, Design, and Construction (PDC) office. The PDC office is the programming and project management level entity which oversees construction of all projects on the VT campus, with a budget over $1,000,000 in value. The VT PDC office currently develops an early-design duration estimate using templates (created from past projects) to estimate the design duration and ‘rule of thumb’ construction duration estimates garnered from experienced personnel.

The link between construction cost and duration was emphasized by the VT PDC office. The continuing escalation of construction costs is recognized by the VT PDC office as well. As such, multi-year projects are often broken down into a number of phases so construction costs can be more accurately estimated. Further, the duration estimate established is used to generate the initial project cost estimates, as the VT PDC office generally inflates current/historical construction costs to the estimated midpoint of construction.

The VT PDC office sees the research concept and methodology (in principle) as applicable to their current needs and work. As a university, the VT PDC office must coordinate its programming and construction calendars to the academic calendar. The office must often commit to facility turnover dates well before construction breaks ground. An earlier, historically-based method of predicting construction duration is expected to enhance the overall campus construction programming efforts. Also, the VT campus has undergone significant expansion in recent years. As such, the VT PDC office sees a significant need for earlier, more reliable planning tools and procedures to cope with the issues surrounding infrastructure planning and integration. Finally, at the resource level, the VT PDC office recognizes that earlier construction duration planning has the potential to improve their long-range personnel planning, particularly in the area of Project Manager level personnel which oversee construction progress.

5.4.3.1.2 Private Construction Owners

Private construction owners may also apply the general concept behind this research. The chief difference between public and private owners is the source of funding used to secure construction services. Public construction owners operate using constituent-supplied funding. As such, public construction owners, can create a deficit, but generally, are not allowed to go bankrupt. Meanwhile, construction financing is a business decision for private owners which must secure funds using company held assets. Just as with any other business decision, poor construction financing decisions have the potential to bankrupt the private construction owner and potentially drive them out of business.

Private construction owners exist in every business sector. Regardless of the sector, construction duration is crucial to financial viability as the return on the construction investment is directly tied to the completion of construction and the turnover of the facility for operation.

To help demonstrate the applicability of this work to private construction owners, a representative from a privately held commercial property owner was interviewed. The commercial property owner interviewed is in the business of constructing leasable commercial facilities. As such, the transition between monetary outflow (in the form of construction costs) and monetary inflow occurs upon commencement of the lease period. Therefore, the construction owner has a need for a fixed date of construction completion. The date on which the lease begins not only signals the end of construction, but is also the basis for establishing the monthly lease price of the facility. This price is established based on the forecasted date of completion and turnover of the facility, the cost of facility construction, and the desire to recoup these costs over a period of time.

To develop the date of occupancy, and subsequently begin planning for return on investment, the commercial owner first estimates the project duration based on personal knowledge and experience in the construction business (as a habitual construction owner). The owner also considers significant factors about the project such as the location and type of facility being constructed. Next, a number of general contractors are contacted and asked to provide a conceptual cost and duration estimate based on these conceptual details. The duration estimate provided is usually provided as a number of months at this stage. A front-end contingency for site inspections and securing permits is also included in this initial duration estimate.

Once potential constructors have offered these conceptual estimates, the commercial owner begins working to secure a lessee to occupy the facility once constructed. Once a lessee is secured, the construction financing and date of completion are fixed. As the construction contract is drawn, this fixed-date of construction completion is agreed upon and established. The commercial owner interviewed cited two scenarios which arise should this fixed-date be missed. In the best case, the commercial owner may be able to maintain the lease, allow the lessee to occupy the facility later, and subsequently forfeit the portion of the lease which was not available to the lessee. The worst case envisioned by the commercial owner, is that construction completion forces the owner to forfeit the entire lease. This would require the owner to secure a new lessee and would undoubtedly cost the owner revenue and additional expenses.

Though private owners have more flexibility in establishing the construction duration and often work with the general contractor to establish a reasonable construction duration, there is still a need for an early, accurate, reliable prediction of construction duration upon initial project conception. Once the need for a new facility is established, the construction owner must evaluate the feasibility of the project in terms of the cost of construction, duration of construction, date of occupancy, and return on investment. In the example cited above, the construction owner begins seeking potential lessees at this conceptual design stage. As such, the owner must have a reliable estimate of the lease price, which is tied directly to the project costs and construction duration, soon after the facility is conceived. This initial estimate may then be used to evaluate general contractor estimates for completion and as a baseline from which to gauge project performance.

The historically-based construction duration prediction concept put forward by this research can be applied by all habitual, or recurring, construction owners. All construction owners, whether public or private, have a need for accurate, reliable construction duration estimates based on initial-design details. For the public owner, the construction duration is primarily tied to operational calendars which they are responsible to the general public for. For the private owner, the construction duration correlates with the beginning of a return on investment, and therefore, the successfulness of the business operation.

5.4.3.2 Methodology Application

Though the models, coefficients, and duration-influential factors are specific to the data analyzed and, hence, VDOT, the methodology employed is pertinent to all construction owners, in principle. This work develops a methodology for identifying, quantifying, and modeling early-known project factors which influence construction duration. Replicating and applying the methodology has numerous benefits to the construction owner [Williams et al. 2008].

In order to apply the methodology developed herein, the construction owner must possess accurate, reliable, and copious data. Habitual owners, whether public or private, have the adequate resources and populations from which to draw the necessary project data. The data used to develop the models and identify the project factors above includes highway construction contract and project factor level data. Contract data includes the project type, location, roadway system, and setting. These factors are important in the specification of mathematical models and data segregation. By refining the data analyzed at this level, relationships within the project factors can be more easily described where they exist. For other industry sectors, this high-level project data may include the construction type (for owners who fund various project types), location, and estimated cost.

Project level data serves to quantify the project size and complexity. The factors identified for highway construction include the parametric cost estimate, geometric design standard, traffic volume, new signal count, presence of median, and presence of curb and gutter. The value of these factors is indicative of the project size and complexity in a quantifiable measure. Further, the value of these factors are determined and known during the initial project design. These qualities make them ideal for statistical regression analysis in the highway construction sector.

Habitual owners in other industry sectors maintain similar project factors on which to base regression models. Previous works cited referenced the use of square footage, building height, and number of stories as explanatory variables in the building construction sector [Chan and Kumaraswamy 1995, Kumaraswamy and Chan 1995, Khosrowshahi and Kaka 1996]. Process industries such as petroleum or chemical companies may use anticipated output volume, average pipe diameter or total length, number or volume of storage tanks as project parameters. Manufacturing companies may opt to use daily output, assembly line length, or factory volume as explanatory variables.

5.5 Summary

Williams et al. [2008a] presents the highway construction duration factors known early in design development, having an influence on construction duration. Williams et al. [2008a] also developed four MLR models which may be used to predict construction duration for two project types: Full-Depth Section and Highway Improvement projects. For each project type, a full model (which incorporates contract and project level explanatory variables) and a condensed model (which incorporates contract level explanatory variables and the estimated project construction cost) was developed. The full models serve to identify and quantify the duration-influential project factors. Meanwhile, the condensed models proved better suited to predicting construction duration due to their simplicity and stability.

This work describes the project factors found to be statistically-significant to construction duration. This discussion includes the meaning of each of the nine duration-influential factors, where the data may be obtained, the limits of the models for each factor, and the confidence interval for each of the factor coefficient estimates. The models developed from these factors are also described.

This work also discusses the applicability and use of the factors, models, underlying methodologies, and research concept to VDOT, highway construction, and the construction industry at large. As data from VDOT construction projects was used for the analysis, the factors determined, models developed, and coefficient values estimated are specific to application within Virginia. However, the highway construction industry can benefit through knowledge of the duration-influential project factors as these factors are common to construction projects across the United States. Further, the research concept and methodology is relevant to the highway construction industry and benefits are expected through adaptation of similar methodologies and analysis. Finally, the work is applicable to the general construction industry as well. Multiple linear regression analysis is verified as a suitable means for identifying and quantifying duration-influential project factors. The concept behind the research is verified citing sources from public and private construction owners outside of highway construction. It is expected that the research concept and methodologies, in principle, can be easily adopted to residential, building, retail, industrial, and process industries in the U.S. and abroad.

5.6 References

Chan, D. W. M., and Kumaraswamy, M., M. (1995). "A study of the factors affecting construction durations in Hong Kong." *Construction Management and Economics*, 13, 319-333.

Construction Industry Institute (CII). (2006). "Front End Planning: Break the Rules, Pay the Price." 213-1.

Hildreth, J. C. (2005). "Dura-Lator: A Project Duration Calculator" Technical Report TR-0503. VDOT-VT Partnership for Project Scheduling, Blacksburg, VA.

Khosrowshahi, F., and Kaka, A. P. (1996). "Estimation of Project Total Cost and Duration for Housing Projects in the U.K." *Building and Environment*, 31(4), 373-383.

Kumaraswamy, M., M., and Chan, D. W. M. (1995). "Determinants of construction duration." *Construction Management and Economics*, 13, 209-217.

Kyte, C. A., Perfater, M. A., Haynes, S., and Lee, H. W. (2004). "Developing and Validating a Highway Construction Project Cost Estimation Tool." Virginia Transportation Research Council, Charlottesville, VA.

Montgomery, D. C., Peck, E. A., and Vining, G. G. (2006). *Introduction to Linear Regression Analysis*, 4th Ed., John Wiley & Sons, Inc., Hoboken, NJ.

Ott, R. L., and Longnecker, M. (2001). *An Introduction to Statistical Methods and Data Analysis*, Duxbury, Pacific Grove, CA.

Williams, R. C., Anderson-Cook, C. M., Hildreth, J. C., and Vorster, M. C. (2008a). "Statistical Regression Analysis to Identify and Model Relationships between Conceptual Design Parameters and Highway Construction Data." *ASCE Journal of Construction Engineering and Management*, Submitted for review.

Williams, R. C., Hildreth, J. C., and Vorster, M. C. (2008b). "Highway Construction Data Collection and Treatment for Statistical Regression Analysis." *ASCE Journal of Construction Engineering and Management*, Submitted for review.

Williams, R. C., Vorster, M. C., and Hildreth, J. C. (2008c). "The Need for Highway Construction Duration Estimating Using Early-known Design Details." *Transportation Research Record: Journal of the Transportation Research Board*, Submitted for review.

Chapter 6 – Conclusions and Recommendations

The research work performed is presented as four distinct, self-standing manuscripts. These manuscripts have been submitted to various research journals and review boards and are in various stages of the submission and review processes. This chapter describes how the manuscripts, taken as a whole, present the research. The contributions of the work to the construction industry body of knowledge are described. This work presents contributions to various industry entities, each of which are detailed in this chapter. Finally, future research is recommended to further this work and to build on the foundational methodology presented here.

6.1 Summary of the Work

This work sought to achieve three principle objectives:

1. Identify the early-known project factors which influence construction duration,
2. Quantify the relationship between these project factors and construction duration, and
3. Incorporate the early-known construction duration-influential project factors and their relationship with construction duration into mathematical models to facilitate prediction of construction duration.

These objectives were achieved through this work. The methodologies employed and outcomes of various aspects of the research are discussed in four manuscripts submitted to industry research journals. Here, each manuscript is summarized, discussing the work performed and the results achieved.

6.1.1 Need for the Research

Chapter 2 presents the current state of practice for predicting highway construction duration across the United States. To complete this review, state highway agencies (SHAs) and project design and management professionals from across the U.S. were interviewed. Also, experienced planning personnel from the Virginia Department of Transportation (VDOT) were interviewed to present a thorough look at the project development process and the role of the construction duration estimate within that process.

The industry review discovered that while most (80%) of SHAs contacted recognize the need for a systematic method for predicting construction duration during initial design, few had existing tools or procedures for preparing such an early estimate. Of those who had existing tools and procedures for predicting construction duration during preliminary design, the commonly used methods included *1)* the development of a critical path method (CPM) schedule using assumed quantities, means and methods, and sequence of construction, *2)* the comparison of the current project to previously completed projects, and *3)* estimates from experienced personnel. Finally, the SHAs contacted cited several benefits to having an accurate and reliable construction duration estimate during preliminary design. Of the benefits cited, the most popular included enhanced project and program planning efforts, improved funding knowledge, improved public safety and convenience, and improved resource planning.

Design and management professionals offered a variety of preliminary design construction duration prediction methods including using historical schedule data, outline schedules to build up the project schedule from components of the project, and personnel experience with similar work. The design professionals contacted all recognized a need for and benefits of a systematic methodology for predicting construction duration during preliminary project design. The benefits cited varied from improved equipment management and planning to supplementing the experience of the schedulers with historically-based and statistically-significant information.

Finally, the in-depth analysis of the VDOT project development process revealed that procedures, even within an individual SHA, varied. While the organization operates on the same project development lifecycle, individual VDOT districts prepared preliminary construction duration estimates differently. However, contacts with the location and design (programming) division recognized many benefits to a systematic method for predicting construction duration, based on historical project information. The benefits cited relate to the organization-level construction program planning.

Chapter 2 also presents a review of the literature concerning early duration prediction, and statistical regression analysis as a potential method for modeling construction duration. The review shows that other construction industry sectors have successfully used statistical regression analysis to identify, quantify, and model construction duration using contract level information. However, such work has not been done in the U.S., the highway construction industry, or to incorporate project level information to predict construction duration.

6.1.2 Data Collection and Treatment

Chapter 3 presents the results of identifying and selecting a construction duration data source, collecting the necessary contract and project data, topical treatment of the data, and seasonal-adjustment of the contract duration to prepare a seasonally-adjusted construction duration.

Habitual or recurring owners are in the best position to supply the quantity and quality of data necessary for this study. SHAs are the most prevalent highway construction habitual owners and, therefore, the targeted data supplier. Two SHAs are identified as having such information: Texas Department of Transportation (TxDOT) and VDOT. VDOT was selected as the optimal choice for supplying the data required by this study.

Data was gathered from two sources within VDOT: *Data Warehouse Management Information Portal* (*DWMIP*) and *Project Cost Estimating System* (*PCES*). *DWMIP* contains contract-level data used to obtain historical project durations, as well as factors which describe the location, type, and setting of the projects studied. *PCES* stores project data from factors known to influence construction cost. These factors describe the project size, complexity, and quantities of work. It was hypothesized that many of these same factors can be used to describe construction duration.

Data was collected for projects which began and were completed between January 1, 2000 and January 1, 2007. Operating within this timeframe offered a large amount of data. Once collected, potential explanatory variables were identified and selected. Finally, the data was analyzed for completeness and erroneous entries that clearly did not belong in the data.

To remove the effects of adverse weather, non-working days, and the time of year in which work is completed, the contract duration collected from *DWMIP* was seasonally-adjusted to a historically-typical construction duration. This was done using National Oceanic and Atmospheric Administration (NOAA) weather data from 1971 –

2000 and event impact factors. The weather impact factors were developed using engineering judgment and a review of the literature regarding weather impacts on construction. These impact factors account for potential delays and obstructions due to various thresholds of rain, snow, and temperature events in Virginia. The historically-typical construction duration is a working-day duration, indicative of the expected duration required to complete construction activities on site. This duration would serve as the response (or dependent) variable for the ensuing statistical regression analysis.

6.1.3 Statistical Regression Analysis

Chapter 4 presents the results of the statistical regression analysis aimed at identifying duration-influential project factors, quantifying the relationships between individual factors and construction duration, and modeling construction duration for purposes of prediction. Two major types of construction projects are studied separately: Full-Depth Section and Highway Improvement projects.

Full-Depth Section projects include new road construction, reconstruction, realignment, and major widening projects. Highway Improvement projects include resurfacing and rehabilitation. These project types are different in the nature of the work and effects of adverse weather on construction activities. Therefore, the weather impact factors vary for the two project types and the construction duration is calculated separately. Further, the project factors that influence construction duration vary by project type. Therefore, statistical regression analysis is performed separately for each project type.

The nature of the data collected from *PCES* presented two options for regression analysis of each of the project types. While *PCES* stores data from project factors input by users, the construction cost estimate is also stored within the system. The construction cost estimate is generated using the project factors input and historically-derived construction costs. As such, a full model could be analyzed which incorporated both *DWMIP* data and individual *PCES* project factors. Conversely, a condensed model was analyzed which used *DWMIP* data and the *PCES* construction cost estimate, derived from the individual project factors.

The full models developed, for both Full-Depth Section and Highway Improvement projects, identified the early-known duration-influential construction project factors. These factors include District, Highway System, Area Location Name, Geometric Design Standard, Traffic Volume, New Signal Count, presence of Curb & Gutter, and presence of Median. As expected the factors indentified and their relationship to construction duration varied by project type.

The condensed models were initially prepared to test the hypothesis that while a relationship between the construction cost and construction duration likely exists, the ability to explain variability in construction duration would be improved by using the individual project factors. This hypothesis proved correct as the full models developed had an R^2_{adj} that was 5% and 9% higher for Full-Depth Section and Highway Improvement projects, respectively. In validating the models, however, the condensed models proved to be more stable predictors of construction duration. This is attributed to the higher number of independent variables in the full models versus the condensed models. Given the small reduction in R^2_{adj}, their stability, and simplicity, the condensed models are deemed the better practical predictors of construction duration.

Chapter 4 also offers recommendations for future research which builds on the statistical regression analysis portion of this research. Additional recommendations for future work are presented later in this chapter.

6.1.4 Methodology Application

Chapter 5 presents the practical application of the models developed and the methodologies employed through this research. Because they each serve a specific purpose, both the full and condensed models are represented here and described on their merits. The discussion of the models includes description of project factors identified as statistically-significant, the coefficient estimates generated, and the sensitivity of these estimates (in terms of 95% confidence intervals).

Application of the models developed and the research methodologies are discussed by the three entities to which they are applicable: VDOT, highway construction, and the general construction industries. Because they are generated using VDOT data, the factors identified and coefficient estimates are specific to VDOT.

Therefore, the application discussion includes information regarding the location and range of the duration-influential project factor data. Location pertains to the system from which the data is taken, field name, and any transforms that must be done to the data before entering it into the models. The range of the data pertains to the limits within which the models developed should be used. Extrapolation beyond the limits of the statistical analysis data limits is dangerous as the models do not represent the relationships beyond these limits. The limits of each factor studied are included in the discussion. Examples of using the condensed models to predict construction duration within VDOT are presented. Finally, information regarding the construction duration (working-day) output of the models is presented.

The findings of the research also apply to the highway construction industry. Particularly, knowledge of the duration-influential project factors, the magnitude of the coefficient estimates, and the sign of the coefficients is relevant and applicable to SHAs across the U.S. Finally, the research methodology is applicable to both highway construction and the general construction industries. Chapter 5 describes the type of owner who could apply the information and the methodologies, as well as the parallels between the highway construction duration-influential factors and those from other habitual owners in other construction industry sectors.

6.2 Contributions to the Body of Knowledge

The methodology developed and results of this research contribute to entities across the construction industry spectrum. This section describes the chief contributions of the work, the benefitting entity, and how this information enhances the construction industry body of knowledge.

Figure 6.1 summarizes these contributions and the construction industry entity which benefits or receives the contribution. As seen, the pyramid has three levels which speak to the three main contributions to the body of knowledge and the entities which benefit most readily from them. These entities include VDOT who can directly use the results of the research, the highway construction industry sector who can apply the methodologies developed, and the general construction industry who may adapt the methodologies developed.

The pyramid in Figure 6.1 does not come to an apex as the research presents future areas of study and understanding. The apex of the pyramid is discussed in *6.3 Recommended Future Work*. Here, the lower three levels of contributions are discussed.

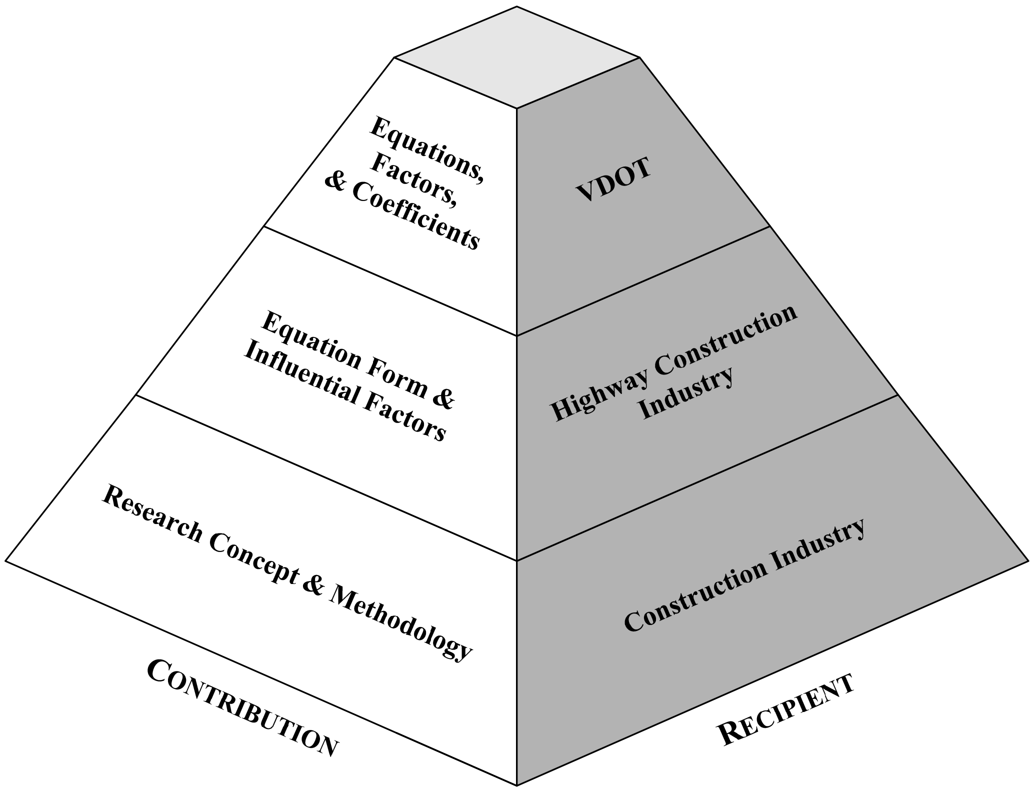

Figure 6.1 – Contributions to the Body of Knowledge

6.2.1 Research Concept and Methodology

This work builds on international studies from various construction industry sectors. Other authors have demonstrated the validity of using statistical regression analysis to model the relationship between duration and various project factors. However, the duration sought has been either the conceptual contract duration or the detailed activity duration [Bromilow 1980, Nkado 1992, Chan and Kumaraswamy 1995, Chan 1999, Skitmore and Ng 2001, Burrows et al. 2005, Jiang and Wu 2007]. To model such durations, the project factors studied have been of commensurate levels, either very conceptual contract details or very detailed task level details [Bromilow 1980, Nkado 1992, Chan and Kumaraswamy 1995, Chan 1999, Skitmore and Ng 2001, Burrows et al. 2005, Jiang and Wu 2007]. Further, these studies have taken place internationally, with domestic duration estimating practices relying less on historic project data and statistically-significant project factors, and more on personnel experience and comparison to similar projects [Williams et al. 2008]. Finally, highway construction has seen very little representation in the literature.

This work develops a methodology for modeling highway construction duration using early-known contract and project level factors. The construction duration modeled is a working-day duration which reduces the noise and variability induced by the start date of construction and the adverse weather and non-working days included in the longer contract duration. The contract level factors classify the project by type, location, system, and area or setting. Meanwhile, project level factors serve to quantify the project size, components, and quantities of work.

It is felt that performance of this work completes a significant gap in the construction industry body of knowledge regarding preliminary duration prediction. The concept of using both contract and project level factors to model construction duration is unique. Incorporation of project size, complexity, and quantity information found in project level data refines the magnitude duration estimate which was based on broad contract level details such as the location and setting of the project.

The performance of such work and attention to the highway construction industry fills a void in the existing literature. While the duration-influential factors vary by industry sector, understanding the highway construction duration-influential project factors can be used to understand the level and type of factors which influence construction duration in other industry sectors.

Within highway construction, two project types are analyzed and the construction duration of each determined to be influenced by different project factors. These project types (Full-Depth Section and Highway Improvement) are developed by grouping work types with similar means and methods of construction. Full-Depth Section projects are influenced by traffic volume, location, number of new signals, presence of curb and gutter, presence of median, and geometric design standard. Meanwhile, Highway Improvement projects are influenced by traffic volume, presence of median, number of new signals, and geometric design standard.

The methodology underlying the performance of the research is unique in its collection and combination of various levels of construction data, seasonal adjustment of contract duration to calculate a historically-typical construction duration, identification of duration-influential project factors, and modeling of construction duration using these early-known factors. As such, the methodology may be adapted by other construction industry sectors for their own purposes.

As with any research, the concept and methodology is the base on which all else is built. Within construction then, contributions arising from the development of new concepts and methodologies are most readily applicable and pertinent to the general construction industry. This is depicted by the bottom level of the pyramid in Figure 6.1.

6.2.2 Equation Form and Influential Factors

This work provides a number of contributions specifically to the highway construction body of knowledge. Using highway construction data, duration-influential project factors were determined. These factors are those having a statistically-significant relationship with construction duration. Multiple linear regression analysis was chosen as the optimal statistical method for this work, due to the ability to describe data and its "rich and flexible framework that suits the needs of many analysts" [Montgomery et al. 2006].

In performing this analysis, projects were divided into two types (Full-Depth Section and Highway Improvement). Due to the nature of the data obtained, two models were able to be developed for each project type: a full model using all duration-influential contract and project factors and a condensed model which used contract level factors and a surrogate cost estimate developed using the project level factors. Validation of these models revealed that the full models were useful in identifying the duration-influential factors that effect construction duration, while the condensed models were best for predicting highway construction duration.

The highway construction industry benefits in three ways from this work. First, the methodology developed and employed to perform the analysis contributes to the highway construction industry body of knowledge. These contributions are derived from the identification and selection of a data source, data collection and compilation, initial data treatment, and seasonal-adjustment to calculate construction duration. This methodology, and the principles guiding it, is applicable to any similar study for SHAs across the U.S. and abroad.

Second, this work determined several early-known, duration-influential project factors. As highway construction work is generally similar across the U.S., knowledge of these factors can be applied by highway construction owners across the nation. Such knowledge adds to the understanding of the way in which duration is impacted by various project factors. The factors are known to influence duration and, therefore, warrant additional design and planning attention. With this information, the planner may begin to make more informed decisions regarding the overall construction process design.

Finally, multiple linear regression analysis proved a viable method for quantifying the relationship between these duration-influential factors and construction duration. The sign and magnitude of these relationships (or coefficient estimates) contribute to the understanding of the relationship (positive or negative) between duration and factor quantity, as well as the impact of the factor and its quantity on construction duration. While the coefficient estimates developed using this VDOT data are specific to Virginia, estimates in other states are expected to have similar magnitudes and signs, adding to the understanding of the factors which influence highway construction duration.

6.2.3 Equations, Factors, and Coefficients

VDOT benefits directly from this work as the supplier of the data analyzed. As stated, the regression analysis *1)* identified early-known, duration influential project factors, *2)* quantified the relationship between these factors and construction duration, and *3)* developed models for prediction of construction duration using these factors and relationships to construction duration. Prior to regression analysis, the methodology developed for identifying, collecting, and treating the data was also specified to VDOT. Finally, the overall research methodology and concept contributes to the VDOT body of knowledge as it is the base of the theoretical pyramid of contributions.

The VDOT body of knowledge benefits most readily as the factors identified, the relationships quantified, and prediction models developed are based on Virginia construction data and, hence, readily applicable for the purposes of predicting construction duration. Through the statistical regression analysis, four models were developed: two full models and two condensed models. The full models developed serve to identify the early-known factors that influence construction duration, the sign of the relationships, and the magnitude of the relationships with construction duration. Meanwhile, the condensed models proved more stable and easier to use while explaining variability in construction duration nearly as well as the full models. Therefore, the information garnered from these condensed models (statistically-significant factors, relationship to construction duration, sign and magnitude of the relationships, and prediction model) contribute to the VDOT body of knowledge.

VDOT can directly use these results to predict construction duration, systematically, early in the project design lifecycle, using duration-influential project factors. The results of this research are a 'snapshot' of statistical trends from historical project data. As such, the results of this research provide VDOT a benchmark on which future research and improvements in the construction process may be based. This basic understanding and datum opens the door to numerous future research studies which will further enhance the construction industry body of knowledge.

6.3 Final Conclusions

The performance of this work generated numerous conclusions regarding the preliminary prediction of highway construction duration. The individual manuscripts document these conclusions as the methodology is developed and documented. This section summarizes the main conclusions reached through the completion of this work.

The processes and procedures used by SHAs for predicting highway construction duration during initial design are inadequate. Those states with consistent methods for prediction construction duration early in the design lifecycle often use either overly-complex detailed scheduling techniques or rely personnel experiences, which is a dynamic knowledge-base. The highway construction industry needs an understanding of the duration-influential factors, known early in project design, which influence construction duration. Further, the industry needs improved methods and processes by which construction duration can be calculated during initial design.

Details known during initial design include information which explains the conditions under which work is performed, as well as the size and complexity of the project work. A compilation of data was used here to provide an adequate field of descriptors to obtain duration-influential project factors. This data is maintained by SHAs, often for cost estimating or monitoring purposes, and can be used to explain the variability in construction duration. Because highway construction contracts are often fixed-date, the contract duration is reflective of the calendar days between contractor selection and project acceptance. This duration includes allowances for adverse weather and non-working days which are dependent on the time of year in which work begins and is performed. As such, seasonal adjustment is necessary to reduce this induced bias. Such adjustment converts the historical contract duration to construction duration. Here, generalized weather impact factors and monthly weather event averages were used to perform the required seasonal adjustment. This level of detail is easily automated, less labor-intensive, and commensurate with the conceptual design level of detail under which the duration prediction will be performed.

Statistical regression analysis was used to develop multiple linear regression (MLR) models for identifying duration-influential project factors and modeling construction duration. Based on related work by other authors, and the results of this study, MLR is verified as an excellent method for identifying individual duration-influential project factors, known early in project design. The relationships between duration-influential factors and construction duration can be developed using MLR. Finally, these duration-influential project factors and their relationships with construction duration can be modeled using MLR. The nature of MLR enhances the understanding of the influence of each factor and provides a robust, yet straightforward, tool for modeling construction duration.

Several early-design factors have been shown to influence construction duration. These factors are both contract and project level design details. The majority of the factors are project level details, which serve to explain the project size and complexity. The early-known project factors which influence construction duration are traffic volume, location, geometric design standard, presence of a median, presence of curb and gutter, project setting, and new signal count. Further, a collection of individual project factors are better able to explain the variability in construction duration. However, a condensed model, made up of contract level details and a conglomerate of the project level details, in the form a construction cost estimate, is an adequate explainer of construction duration variability, using fewer variables or factors.

Finally, the methodologies developed and the results of this work are applicable to numerous construction industry entities. From the concept and methodology, to the equations and factors developed, the research is applicable to VDOT, highway construction owners and contractors, and the general construction industry at large.

6.4 Recommended Future Research

Several areas for future research presented themselves through the completion of this work. Recommended future research forms the theoretical apex of the pyramid in Figure 6.1. The research topics proposed here build on the work performed and generate new lines of questioning, research, and understanding.

These areas for future research pertain to VDOT, highway construction, and general construction industries and are expected to be beneficial to the overall construction industry body of knowledge.

6.4.1 *DWMIP* Data Collection

DWMIP is used primarily for the purposes of maintaining historical contract and project data and as a data source for high-level project monitoring. As such, the project and contract factors maintained are not tailored to the needs of this study in investigating duration-influential project factors. A review of the literature and discussions with experienced industry representatives suggest that additional project factors could be gathered to aid in the understanding of construction duration. Such data includes project funding information, project delivery methods, and information regarding the conditions under which work was performed. Therefore, future research should be aimed at identifying and collecting additional contract and project factors in *DWMIP* which could be used to further analyze the development of construction duration.

6.4.2 *PCES* Data Collection

PCES is tailored to estimating construction cost. However, the factors obtained from *PCES* proved to be influential on construction duration as well. Because the information is similar, *PCES* is a suitable tool for collecting the data needed for similar research. However, the *PCES* factors collected could be expanded to incorporate more early-known influential factors. Such expansion would likely improve the accuracy of the *PCES* cost estimate and lend additional potential explanatory variables, or factors, to the construction duration estimate as well.

Also, as a multiple-user tool, *PCES* data can suffer from inaccurate or invalid data entries. A significant portion of this work was dedicated to identifying erroneous entries both in the data treatment and residual analysis processes. Further, many *PCES* entries were created by the user opening the system, but not completing the estimate. This generated another database entry with little useful data. Improvement of the data collection practices would provide quality data for future research.

6.4.3 Continued Refinement of Factors, Coefficients, and Models

This research performed is analogous to a photograph of the Virginia highway construction industry state of practice. The retrospective analysis performed could not account for specific project circumstances beyond that demonstrated in the data. Further, the factors identified, their relationships with construction duration, and the models developed are specific to the data analyzed. Therefore, several areas of future work present themselves through additional data collection.

First, the factors identified should be studied and refined as necessary. Numerous circumstances have the ability to alter the significance of the duration-influential factors discovered in this study. These circumstances include, but are not limited to, improvements in construction equipment, labor productivity, materials, and means and methods. Population, congestion, and land development would also have the ability to alter the statistically-significant factors as many of the factors identified relate to, not only the construction project itself, but the area, location, and environment in which it is performed.

As new data is collected, the coefficients or relationships between duration-influential factors and construction duration will undoubtedly change. As such, these coefficients should be periodically refined to ensure an adequate representation of the historical trends.

Finally, as the significant factors and their relationship with construction duration changes, the models developed from them will also change. Therefore, as factors and coefficients are updated, the models developed, particularly the condensed models suggested for prediction purposes, should be updated.

6.4.4 Analysis of Bridges

A limitation of this work is that bridge projects were not included in the analysis. Unfortunately, two few bridge projects were present among the data to adequately analyze those factors which influence bridge construction duration. As data collection practices are improved and data collection continues, it is felt that bridge projects should also be analyzed using a similar methodology to that developed herein.

6.4.5 Work Type Specific Analysis

In this study, projects were classified as either “Full-Depth Section” or “Highway Improvement” because of the expected differences in the construction methods and project scopes. These project types are developed by grouping work types. Full-Depth Section projects include new construction, reconstruction, major widening, etc. Highway Improvement projects include resurfacing and rehabilitation works. This grouping is the result of the distinct types of construction operations, weather impacts, and means and methods associated with each project type.

While these work types are generically grouped for this study, construction duration-influential factors and prediction models for individual work types should be examined in the future. Examining individual work types and understanding the duration-influential factors of each will aid in better understanding the differences in various highway construction work types and the impacts of contract and project level factors on construction duration.

6.4.6 Project Financing Implications on Duration

As mentioned, a number of project factors have the potential to impact construction duration, beyond those studied in this analysis. Further, the establishment of construction and contract duration by the owner is influenced by much more than project geometry, complexity, and size. Often, the establishment of contract, and hence construction, duration is often influenced by strategic financing and cash flow decisions. Contractor work and project progression can slow or stop if the funding available to the owner for payment of the contractor is not forthcoming. In such instances, the overall project duration is affected and the impact immeasurable by the analysis of project size, complexity, and quantity factors used here. Discussions with industry experts suggest that there is potential for future research in the area of understanding the business decisions which influence construction and contract duration. It is felt that such research would compliment the work presented here and is, therefore, warranted.

6.4.7 Develop a Consensus of Duration-Influential Factors

While the statistically-significant factors discovered here are generally applicable to the U.S. highway construction industry, it is expected that these factors will vary by state. For instance, states that are more homogenous may not see an influence on duration due to location. Further, states will likely assess project setting and highway systems differently. As such, the early-known factors that influence construction duration should be determined by SHAs for their specific purposes. An industry review showed that most SHAs do not have a standardized method for predicting construction duration early in project design. Identification of these factors will not only lend itself to improving the prediction of construction duration among individual states but begin to develop a consensus of project factors across the country. Such information will inform other states as they assess their own early-known duration-influential factors and also provide a more thorough understanding of the heterogeneity of highway construction across the country.

6.4.8 Adoption of the Methodology

This research is unique in the methodology developed and the research undertaken to use both contract and project level duration data. Other industry sectors and researchers in other countries have used either contract level data or task-specific data to model duration. Therefore, it is felt that opportunities exist for adopting the methodology developed in many other industry sectors. Such adoption will undoubtedly contribute to the construction body of knowledge and further understanding of duration estimation based on early-known construction details.

6.5 References

Bromilow, F. J., Hinds, M. F., and Moody, N. F. (1980). "AIQS Survey of Building Contract Time Performance." *The Building Economist*, 19(2), 79-82.

Burrows, T. K., Pegg, I., and Martin, J. (2005). "Predicting Building Construction Duration." *AACE International Transactions*, PS.14.11-PS.14-19.

Chan, D. W. M., and Kumaraswamy, M., M. (1995). "A study of the factors affecting construction durations in Hong Kong." *Construction Management and Economics*, 13, 319-333.

Chan, D. W. M., and Kumaraswamy, M., M. (1999). "Modelling and predicting construction durations in Hong Kong public housing." *Construction Management and Economics*, 17, 351-362

Jiang, Y., and Wu, H. (2007). "A Method for Highway Agency to Estimate Construction Durations and Set Contract Times." *International Journal of Construction Education and Research*, 3, 199-216.

Nkado, R. N. (1992). "Construction time information system for the building industry." *Construction Management and Economics*, 10, 489-509.

Skitmore, R. M., and Ng, S. T. (2001). "Australian project time-cost analysis: statistical analysis of intertemporal trends." *Construction Management and Economics*, 19, 455-458.

Williams, R. C., Vorster, M. C., and Hildreth, J. C. (2008). "The Need for Highway Construction Duration Estimating Using Early-known Design Details." *Transportation Research Record: Journal of the Transportation Research Board*, Submitted for review.

Chapter 7 - Bibliography

(2005). In: *23 Code of Federal Regulations*, United States of America.

Akintoye, A. (2000). "Analysis of factors influencing project cost estimating practice." *Construction Management and Economics*, 18, 77-89.

Anani, S. M. (2003). "Scheduling Results: Automation of Remaining Durations." *Transactions of the American Association of Cost Engineers*, PS.13.11-PS.13.17

Arioli, D. L., and Masi, J. G. (2002). "Development and Maintenance of a Parametric Building Estimating System." *AACE International Transactions*, ES.10.11-ES.10.16.

Attalla, M., and Hegazy, T. (2003). "Predicting Cost Deviation in Reconstruction Projects: Artificial Neural Networks versus Regression." *ASCE Journal of Construction Engineering and Management*, 129(4), 405-411.

Attoh-Okine, N. O. (2002). "Probabilistic analysis of factors affecting highway construction costs: a belief network approach." *Canadian Journal of Civil Engineering*, 29, 369-374.

Baloi, D., and Price, A. D. F. (2003). "Modelling Global Risk Factors Affecting Construction Cost Performance." *International Journal of Project Management*, 21(4), 261-269.

Berrey, L. G. (1992). "Re-engineering the Proposal Process Using Parametric Cost Models," Virginia Polytechnic Institute and State University, Blacksburg, VA.

Black, J. H. (1984). "Application of Parametric Estimating to Cost Engineering." *AACE International Transactions*, B.10.11 - B.10.15.

Boothroyd, G. (1988). "Estimate costs at an early stage." In: *American Machinist*, 54-57.

Brandon, P. S., and Newton, S. (1986). "Improving the forecast." *Chartered Quantity Surveyor*, 24-27.

Bromilow, F. J. (1969). "Contract time performance expectations and the reality." *Building Forum*, 1(3).

Bromilow, F. J., Hinds, M. F., and Moody, N. F. (1980). "AIQS Survey of Building Contract Time Performance." *The Building Economist*, 19(2), 79-82.

Bureau, U. S. C. (2006). "Strategic Plan FY 2006-2010." D. o. Commerce, ed., 19.

Burrows, T. K., Pegg, I., and Martin, J. (2005). "Predicting Building Construction Duration." *AACE International Transactions*, PS.14.11-PS.14-19.

Chan, A. P. C. (1999). "Modelling building durations in Hong Kong." *Construction Management and Economics*, 17, 189-196.

Chan, A. P. C., and Chan, D. W. M. (2004). "Developing a benchmark model for project construction time performance in Hong Kong." *Building and Environment*, 39, 339-349.

Chan, A. P. C., Ho, D. C. K., and Tam, C. M. (2001). "Design and Build Project Success Factors: Multivariate Analysis." *ASCE Journal of Construction Engineering and Management*, 127(2), 93-100

Chan, D. W. M., and Kumaraswamy, M., M. (1995). "A study of the factors affecting construction durations in Hong Kong." *Construction Management and Economics*, 13, 319-333.

Chan, D. W. M., and Kumaraswamy, M., M. (1996). "An Evaluation of Construction Time Performance in the Building Industry." *Building and Environment*, 31(6), 569-578.

Chan, D. W. M., and Kumaraswamy, M., M. (1999). "Forecasting construction durations for public housing projects: a Hong Kong perpective." *Building and Environment*, 34, 633-646.

Chan, D. W. M., and Kumaraswamy, M., M. (1999). "Modeling and predicting construction durations in Hong Kong public housing." *Construction Management and Economics*, 17, 351-362

Chong, W. K., O'Conner, J. T., Chou, J.-S., and Lee, S.-H. (2005). "Predicting the Production Rates of Foundation Construction Using Factor and Regression Analysis." *ASCE Construction Research Congress*, 1-10.

Chou, J.-S. (2005). "Item-Level Quantity-Based Preliminary Cost Estimating System for Highway Earthwork, Landscape, Subgrade Treatments, Base, Surface Courses, Pavement and Traffic Control," University of Texas at Austin, Austin, TX.

Clough, R. H., Sears, G. A., and Sears, S. K. (2000). *Construction Project Management*, 4th Ed., John Wiley & Sons, Inc., New York.

Connolly, T., and Porter, A. L. (1990). "Discretionary Databases in Forecasting." *Journal of Forecasting*, 9, 1-12.

Construction Industry Institute (CII). (2006). "Front End Planning: Break the Rules, Pay the Price." 213-1.

de la Garza, J. M., and Rouhana, K. G. (1995). "Neural Networks Versus Parameter-Based Applications in Cost Estimating." *Cost Engineering*, 37(2), 14-18.

Dean, J. (1936). *Statistical Determination of Costs, with Special Reference to Marginal Costs*, I Ed., University of Chicago, Chicago, IL.

Dreyfus, G. (2005). *Neural Networks: Methodology and Applications*, Springer, New York, Berlin.

Dysert, L. R. (2001). "Developing a Parametric Model for Estimating Process Control Costs." *Cost Engineering*, 43(2), 31-34.

East, E. W., Hassanein, A., and Melin, J. (1992). "Evaluation Factors that Affect Construction Project Duration." US Army Corps of Engineers.

El-Rayes, K., and Moselhi, O. (2001). "Impact of Rainfall on the Productivity of Highway Construction." *Journal of Construction Engineering and Management*, 127(2), 125-131

Farqahl, S. H., and Everett, J. G. (1997). "Learning Curves: Accuracy in Predicting Future Performance." *ASCE Journal of Construction Engineering and Management*, 123(1), 41-45.

Finke, M. R. (1990). "Weather-Related Delays on Government Contracts." *Transactions of the American Association of Cost Engineers*, F.3.1-F.3.8.

Finney, D. J. (1972). *An Introduction to Statistical Science in Agriculture*, 4th Ed., John WIley & Sons, New York, NY.

Fourie, F. (2003). "Time-for Versus Time-of Performance." *AACE International Transactions*, CDR211-CDR215.

Gibson, J. G. E., Irons, K. T., and Ray, M. P. (Year). "Front End Planning for Buildings." *AEI 2006*, ASCE, Omaha, Nebraska, USA, 41-41.

Gosselin, R., and McMullan, L. E. (1989). "Parametric Estimating - In Search of Expert Systems." *Transactions of the American Association of Cost Engineers*, C.1.1-C.1.9.

Hamburger, D. H. (1989). "Scheduling for On Time Project." *Transactions of the American Association of Cost Engineers*, D.1.1-D.1.8.

Hancher, D. E., and Rowings, J. E. (1981). "Setting Highway Construction Contract Duration." *ASCE Journal of the Construction Division*, 107(CO2), 169-179.

Harbuck, R. H. (2002). "Using Models in Parametric Estimating for Transportation Projects." *AACE International Transactions*, EST.05.01-EST.05.09.

Hegazy, T., and Ayed, A. (1998). "Neural Network Model for Parametric Cost Estimation of Highway Projects." *ASCE Journal of Construction Engineering and Management*, 124(3), 210-218.

Hendrickson, C. (2000). *Project Management for Construction*, 2nd Edition. Available at: http://pmbook.ce.cmu.edu.

Huang, R.-y., and Sun, K.-S. (2006). "Non-Unit-Based Planning and Scheduling of Repetitive Construction Projects." *ASCE Journal of Construction Engineering and Management*, 132(6), 585-597.

Isidore, L. J., and Back, W. E. (2001). "Probabilistic Optimal-Cost Scheduling." *ASCE Journal of Construction Engineering and Management*, 127(6), 431-437.

Isidore, L. J., and Back, W. E. (2002). "Multiple Simulation Analysis for Probabilistic Cost and Schedule Integration." *ASCE Journal of Construction Engineering and Management*, 128(3), 211-219.

Jervis, B. M., and Levin, P. (1987). *Construction Law - Principles and Practice*, McGraw-Hill, New York.

Jiang, Y., and Wu, H. (2007). "A Method for Highway Agency to Estimate Construction Durations and Set Contract Times." *International Journal of Construction Education and Research*, 3, 199-216.

Johnson, J. A., and Kwong, A. K. H. (2001). "Cost Modeling-Accurate Preconstruction Estimates Quicker." *Transactions of the American Association of Cost Engineers*, EST.01.01-EST.01.03

Kahane, L. H. (2001). *Regression Basics*, Sage Publications, Inc., Thousand Oaks, CA.

Kaka, A. P., and Price, A. D. F. (1991). "Relationship between value and duration of construction projects." *Construction Management and Economics*, 9, 383-400

Kang, L. S., and Paulson, B. C. (1998). "Information Management to Integrate Cost and Schedule for Civil Engineering Projects." *ASCE Journal of Construction Engineering and Management*, 124(5), 381-389.

Khosrowshahi, F., and Kaka, A. P. (1996). "Estimation of Project Total Cost and Duration for Housing Projects in the U.K." *Building and Environment*, 31(4), 373-383.

Kim, G. H., Seo, D. S., and Kang, K. I. (2005). "Hybrid Models of Neural Networks and Genetic Algorithms for Predicting Preliminary Cost Estimates" *ASCE Journal of Computing in Civil Engineering*. 19(2), 208-211

Knoke, J. R., Liu, L. Y., and Spittler, J. R. (1993). "Square Foot Range Estimating." *Transactions of the American Association of Cost Engineers*, L.6.1-L.6.8.

Koehn, E., and Brown, G. (1985). "Climatic Effects on Construction." *Journal of Construction Engineering and Management*, 111(2), 129-137.

Konchar, M., and Sanvido, V. (1998). "Comparison of U.S. Project Delivery Systems." *ASCE Journal of Construction Engineering and Management*, 124(6), 435-444.

Krokowski, A. W. (1992). "Estimating from Process Flow Diagrams: A Conceptual Approach." *AACE International Transactions*, 1, A.3.1-A.3.6.

Kumaraswamy, M., M., and Chan, D. W. M. (1995). "Determinants of construction duration." *Construction Management and Economics*, 13, 209-217.

Kwak, Y. H., and Watson, R. J. (2005). "Conceptual estimating tool for technology-driven projects: exploring parametric estimating technique." *Technovation*, 25(12), 1430-1436.

Kyte, C. A., Perfater, M. A., Haynes, S., and Lee, H. W. (2004). "Developing and Validating a Highway Construction Project Cost Estimation Tool." Virginia Transportation Research Council, Charlottesville, VA.

Lee, H.-S., and Yi, K. J. (1999). "Application of Mathematical Matrix to Integrate Project Schedule and Cost." *ASCE Journal of Construction Engineering and Management*, 125(5), 339-346.

Lofton, W. D., and Monteith, M. D. (2004). "Collaborative Estimating." *Transactions of the American Association of Cost Engineers*, EST.01.01-EST.01.03

Lowe, D. J., Emsley, M. W., and Harding, A. (2006). "Predicting Construction Costs Using Multiple Regression Techniques." *ASCE Journal of Construction Engineering and Management*, 132(7), 750-758.

Lucko, G. (2003). "A Statistical Analysis and Model of the Residual Value of Different Types of Heavy Construction Equipment," Virginia Polytechnic Institute and State University, Blacksburg, VA.

McCaffer, R. (1975). "Some examples of the use of Regression Analysis as an Estimating Tool." *The Quantity Surveyor*, 81-86.

McDonald, D. F. (2000). "Weather Delays and Impacts." *Construction Engineering*, 42(5), 34-39.

Middleton, J. P. (2004). "Field Operations Analysis Laboratory Construction Process Simulator." Virginia Polytechnic Institute and State University, Blacksburg, VA.

Mitchell Jr., Z. W. (1998). "A Statistical Analysis of Construction Equipment Repair Costs Using Field Data and the Cumulative Cost Model," Virginia Polytechnic Institute and State University, Blacksburg, VA.

Molenaar, K. R. (2005). "Programmatic Cost Risk Analysis for Highway Megaprojects." *ASCE Journal of Construction Engineering and Management*, 131(3), 343-353.

Montgomery, D. C., Peck, E. A., and Vining, G. G. (2006). *Introduction to Linear Regression Analysis*, 4th Ed., John Wiley & Sons, Inc., Hoboken, NJ.

Moselhi, O., Gong, D., and El-Rayes, K. (1997). "Estimating Weather Impact on the Duration of Construction Activities." *Canadian Journal of Civil Engineering*, 24, 359-366.

Myers, R. H. (1986). *Classical and Modern Regression with Applications*, Duxbury Press, Boston, MA.

Myers, R. H., Montgomery, D. C., and Vining, G. G. (2002). *Generalized linear models: with applications in engineering and the sciences*, Wiley-Interscience, New York.

Nassar, H. M., Hegab, M. Y., and Jack, N. W. (2005). "Design Cost Analysis of Transportation Projects." *ASCE Construction Research Congress*, 1-8.

National Institute of Standards and Technology (NIST) / SEMATECH, I. (2004). "NIST/SEMATECH e-handbook of statistical methods." <http://www.itl.nist.gov/div898/handbook>.

Neter, J., Kutner, M. H., Nachtsheim, C. J., and Wasserman, W. (1996). *Applied Linear Statistical Models*, 4th Ed., Irwin, Chicago.

Ng, S. T., Mak, M. M. Y., Skitmore, R. M., Lam, K. C., and Varnam, M. (2001). "The predictive ability of Bromilow's time-cost model." *Construction Management and Economics*, 19, 165-173.

Nkado, R. N. (1992). "Construction time information system for the building industry." *Construction Management and Economics*, 10, 489-509.

O'Conner, J. T., and Huh, Y. (2005). "Crew Production Rates for Contract Time Estimation: Bent Footing, Column, and Cap of Highway Bridges." *ASCE Journal of Construction Engineering and Management*, 131(9), 1013-1020.

Orczyk, J. J. (1989). "Parametric Construction Scheduling," Dissertation, Purdue University, West Lafayette, Indiana.

Organization, F. a. A. (1982). *Estimation of Crop Areas and Yields in Agricultural Statistics*, United Nations, Rome, Italy.

Ott, R. L., and Longnecker, M. (2001). *An Introduction to Statistical Methods and Data Analysis*, Duxbury, Pacific Grove, CA.

Parkinson, C. N. (1980). *Parkinson: The Law*, Houghton Mifflin.

Phaobunjong, K., and Popescu, C. M. (2003). "Parametric Cost Estimating Model for Buildings." *AACE International Transactions*, EST.13.11-EST.13.12.

Poulton, C. G. (1989). "The Dangers of Early Estimates." *Transactions of the American Association of Cost Engineers*, A.3.1-A.3.3.

Program, N. C. H. R. (1978). *Effect of Weather on Highway Construction*, Washington, D.C.

Proverbs, D. G., and Holt, G. D. (2000). "A theoretical model for optimum project (time) performance based on European best practice." *Construction Management and Economics*, 17, 657-665.

Proverbs, D. G., Holt, G. D., and Olomolaiye, P. O. (1999). "Factors impacting construction project duration: a comparison between France, Germany and the U.K." *Building and Environment*, 34, 197-204.

Roebber, P. J., Bruening, S. L., Schultz, D. M., and Jr., J. V. C. (2003). "Improving Snowfall Forecasting by Diagnosing Snow Density." *Weather and Forecasting*, 18(2), 264-287.

Sanders, S. R., Maxwell, R. R., and Glagola, C. R. (1992). "Preliminary Estimating Models for Infrastructure Projects." *Cost Engineering*, 34(8), 7-13.

Sanders, S. R., and Thomas, H. R. (1993). "Masonry Productivity Forecasting Model." *ASCE Journal of Construction Engineering and Management*, 119(1), 163-179.

Schexnayder, C. J., and Mayo, R. E. (2004). *Construction Management Fundamentals*, McGraw Hill Higher Education, Boston, MA.

Sidwell, A. C. (1984). "The Time Performance of Construction Projects." *Architectural Science Review*, 27(4), 85-91.

Skitmore, R. M., and Ng, S. T. (2001). "Australian project time-cost analysis: statistical analysis of intertemporal trends." *Construction Management and Economics*, 19, 455-458.

Skitmore, R. M., and Ng, S. T. (2003). "Forecast models for actual construction time and cost." *Building and Environment*, 38, 1075-1083.

Smith, G. R., and Hancher, D. E. (1989). "Estimating Precipitation Impacts for Scheduling." *Journal of Construction Engineering and Management*, 115(4), 552-566.

Smith, S. D. (1999). "Earthmoving Productivity Estimation Using Linear Regression Techniques." *ASCE Journal of Construction Engineering and Management*, 125(3), 133-141.

Snee, R. (1977). "Validation of Regression Models: Methods and Examples." *Technometrics*, 19(4), 415-428.

Sonmez, R., and Rowings, J. E. (1998). "Construction Labor Productivity Modeling with Neural Networks." *ASCE Journal of Construction Engineering and Management*, 124(6), 498-504.

Stevens, J. P. (2001). *Applied Multivariate Statistics for the Social Sciences*, 4th Ed., Lawrence Erlbaum Associates, Mahwah, New Jersey.

Tam, C. M., Leung, A. W. T., and Liu, D. K. (2002). "Nonlinear Models for Predicting Hoisting Times of Tower Cranes." *ASCE Journal of Computing in Civil Engineering.* 16(1), 76-81.

Transportation, V. D. o. (2008). "Project Development Concurrent Engineering Process." (January 15, 2008).

Trost, S. M., and Oberlender, G. D. (2003). "Predicting Accuracy of Early Cost Estimates Using Factor Analysis and Multivariate Regression." *ASCE Journal of Construction Engineering and Management*, 129(2), 198-204.

Walker, D. H. T. (1995). "An investigation into construction time performance." *Construction Management and Economics*, 13, 263-274.

Williams, R. C. (2006). "The Framework of a Multi-Level Database of Highway Construction Performance Times," Virginia Polytechnic Institute and State University, Blacksburg, VA.

Williams, R. C., Anderson-Cook, C. M., Hildreth, J. C., and Vorster, M. C. (2008). "Statistical Regression Analysis to Identify and Model Relationships between Conceptual Design Parameters and Highway Construction Data." *ASCE Journal of Construction Engineering and Management*, Submitted for review.

Williams, R. C., and Hildreth, J. C. (2006). "Framework of a Multi-Level Database of Highway Construction Performance Times." VDOT-VT Partnership for Project Scheduling, Blacksburg, VA.

Williams, R. C., Hildreth, J. C., and Vorster, M. C. (2006). "The Framework of a Multi-Level Database of Highway Construction Performance Times," Virginia Polytechnic Institute and State University, Blacksburg, VA.

Williams, R. C., Hildreth, J. C., and Vorster, M. C. (2007). "BIDDS: A Bid Item Level Performance Time Database Management System." *ASCE Workshop on Computing in Civil Engineering*, Carnegie Mellon University, Pittsburgh, PA, 8.

Williams, R. C., Hildreth, J. C., and Vorster, M. C. (2008). "Highway Construction Data Collection and Treatment for Statistical Regression Analysis." *ASCE Journal of Construction Engineering and Management*, Submitted for review.

Williams, R. C., Vorster, M. C., and Hildreth, J. C. (2008). "The Need for Highway Construction Duration Estimating Using Early-known Design Details." *Transportation Research Record: Journal of the Transportation Research Board*, Submitted for review.

Williams, T. P., Lakshminarayanan, S., and Sackrowitz, H. (2005). "Analyzing Bidding Statistics to Predict Completed Project Cost." *ASCE Journal of Computing in Civil Engineering*.

Williams, W. H. (1966). "A Linear Model Approach to Time and Cost Analysis." *Management Science*, 12(6), 8 pgs.

Wilmot, C. G., and Mei, B. (2005). "Neural Network Modeling of Highway Construction Costs." *ASCE Journal of Management in Engineering*, 131(7), 765-771.

Zayed, T. M., and Halpin, D. W. (2005). "Productivity and Cost Regression Models for Pile Construction." *ASCE Journal of Construction Engineering and Management*, 131(7), 779-789.

Zentner, R. S. (1996). "Accurate Estimates Start With Clear Pictures." *Transactions of the American Association of Cost Engineers*, C.10.11-C.10.16