

Addis Ababa
University

(Since 1950)



ADDIS ABABA UNIVERSITY

ADDIS ABABA INSTITUTE OF TECHNOLOGY

SCHOOL OF CIVIL AND ENVIRONMENTAL ENGINEERING

***Model for Estimating Construction Duration of Public Building
Projects in Addis Ababa***

A thesis submitted to the School of Graduate Studies of the Addis Ababa
University in partial fulfillment of the requirements for the degree of Master
of Science in Civil Engineering

(Stream: Construction Technology and Management)

(Sponsor: Ethiopian Roads Authority)

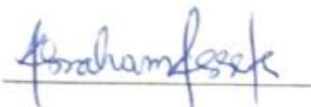
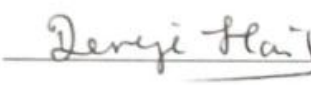

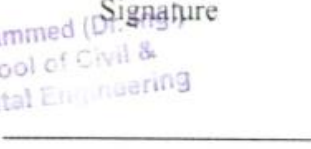
By: Dawit Birhanu

ID No.: GSE/6633/08

Advisor: Abraham Assefa Tsehayae, (PhD)

June 2020

The undersigned have examined the thesis entitled “**Model for Estimating Construction Duration of Public Building Projects in Addis Ababa**” presented by Dawit Birhanu, a candidate for the degree of Master of Science and hereby certify that it is worthy of acceptance.

<u>Abraham Assefa Tsehavae (PhD)</u>	<u></u>	<u>09/09/2020</u>
Advisor	Signature	Date
<u>Dr. Dereje Hailu</u>	<u></u>	<u>Aug 31, 2020</u>
Internal Examiner	Signature	Date
<u>Girmay Kahsay</u>	<u></u>	<u>23/09/2020</u>
External Examiner	Signature	Date
<u>Mebruk Mohammed (Dr. Eng)</u>	<u></u>	<u></u>
Chairman of the school	Signature	Date



Declaration

I, the undersigned, declare that this thesis entitled, “**Model for Estimating Construction Duration of Public Building Projects in Addis Ababa**”, is my original work and it has not been presented for a degree in any other university and that all sources of materials used for the thesis have been duly acknowledged.

Name: - Dawit Birhanu

Signature: - _____

Place: - Addis Ababa University School of Graduate studies, Addis Ababa Institute of Technology, School of Civil and Environmental Engineering.

Date of submission: - June 2020

Acknowledgement

First and foremost, I would like to praise the almighty God for giving me the patience and stamina to accomplish this study. Then I would like to express my heartfelt gratitude to my advisor, Abraham Assefa Tsehayae (PhD), whose priceless reviews and remarks guided me through the study. I would also like to extend my appreciation to the Engineers and Co. at Addis Ababa City Administration Construction Bureau (AACACB) and Federal Government Buildings Construction Project Office (FGBCPO), for availing the data set required for the study from their respective offices. Last but not least, I am grateful to my family and friends for their moral support.

Abstract

Predicting project duration with reasonable accuracy is a formidable challenge to industry practitioners and alike. Consequently, several research studies have focused on this subject, and thus multiple models have been developed. Notwithstanding the conflicting findings of its predictive ability, the 50-year-old “time-cost” model of Bromilow (BTC) remains the baseline, to which the outcomes of most developed models are compared. The Building and Transport Construction Design Authority (BaTCoDA) had developed a model based on BTC in 1987 for building projects but it is hardly used in the industry. The objective of this study, therefore, is to examine the validity, and hence the applicability, of the BTC model to estimate contract durations for public building projects in Addis Ababa, and to propose an alternative model, taking into consideration other potentially influential project scope variables on time performance. In this regard, relevant data for 34 public buildings in Addis Ababa, which were completed during the period 2010 and 2019, were collected by an archival analysis.

The results of the analysis show that while the BTC model holds true for the projects under the study, a model with a better statistical significance was developed with gross floor area as a sole predictor variable. The p -value and F -value for the BTC was 0.00019 and 18.403 respectively, while that of the developed model was 0.000076 and 21.734 respectively. Furthermore, the adjusted R^2 for the BTC was 0.397 while that of the developed model was 0.425 which implies the developed model was a better fit for the data set. Furthermore, the models were compared with the model developed by BaTCoDA and the result shows that the model developed based on gross floor area is superior on account of predictive

Key Words: Construction duration, Bromilow’s Time-Cost model, Regression model

Table of Contents

Declaration	i
Acknowledgement	ii
Abstract	iii
Table of Contents	iv
List of Tables and Figures.....	vii
1. Introduction.....	1
1.1. Study Background	1
1.2. Statement of the Problem	2
1.3. Objectives of the Study	3
1.3.1. General Objective	3
1.3.2. Specific Objectives	3
1.4. Research Questions	3
1.5. Research Hypotheses.....	3
1.6. Scope of the Study.....	4
1.7. Significance of the Study	4
1.8. Organization of the Research	5
2. Literature Review.....	6
2.1. Project and Project Management.....	6
2.1.1. Project: Definition.....	6
2.1.2. Project Management	6
2.1.3. Project Management Process Groups	7
2.1.4. Project Life Cycle	8
2.2. Construction Duration	11

2.3.	Factors Affecting Construction Duration.....	12
2.4.	Construction Duration Estimation Modelling.....	15
2.4.1.	Regression Analysis.....	16
2.4.2.	Steps in Regression Analysis.....	17
2.4.3.	Bromilow’s Time-Cost Model (BTC).....	19
2.4.4.	Previous Studies.....	21
2.4.5.	The Ethiopian Context.....	30
3.	Research Design and Methodology.....	33
3.1.	Introduction.....	33
3.2.	Study Design.....	33
3.2.1.	Study Approach.....	33
3.2.2.	Study procedure.....	34
3.2.3.	Data collection.....	34
3.2.4.	Data analysis.....	38
3.2.5.	Test of Hypothesis.....	39
3.2.6.	Model Validation.....	39
4.	Results and Discussion.....	41
4.1.	Data Characteristics.....	41
4.1.1.	Data Collection.....	41
4.1.2.	Data Processing.....	43
4.1.3.	Data Correlation.....	44
4.1.4.	Data Transformation.....	46
4.2.	BTC Model.....	48
4.2.1.	Test of Hypothesis 1.....	50
4.3.	Multiple Linear Regression Model.....	51

4.3.1. Model Specification	51
4.3.2. Choice of Fitting Method.....	52
4.3.3. Model Fitting	52
4.3.4. Test of Hypothesis 2	55
4.4. Regression Diagnostics and Model Validation	55
4.4.1. Assumptions of Regression Analysis.....	56
4.4.2. Model Validation	60
5. Conclusion and Recommendation	62
5.1. Conclusion.....	62
5.2. Recommendations	63
References.....	65
Appendices.....	70

List of Tables and Figures

List of Tables

Table 2-1 Summary of inputs used to model construction duration.....	27
Table 4-1 Description of Variables.....	41
Table 4-2 Description of Projects for the Study.....	42
Table 4-3 Correlation Matrix for Unadjusted Data.....	44
Table 4-4 Correlation Matrix for Transformed Data.....	46
Table 4-5 Model Summary - BTC Actual Data.....	48
Table 4-6 ANOVA – BTC Actual Data.....	49
Table 4-7 Model Coefficients – BTC Actual Data.....	49
Table 4-8 Model Summary - BTC Contract Data.....	49
Table 4-9 ANOVA - BTC Contract Data.....	49
Table 4-10 Model Coefficients - BTC Contract Data.....	49
Table 4-11 Variables Entered/Removed ^a – Developed Model.....	53
Table 4-12 Model Summary ^b – Developed Model.....	53
Table 4-13 ANOVA ^a – Developed Model.....	53
Table 4-14 Coefficients ^a – Developed Model.....	53
Table 4-15 Excluded Variables ^a – Developed Model.....	54
Table 4-16 Comparison of Models.....	55
Table 4-17 Actual and Calculated Duration of Projects used for calibration.....	61
Table 4-18 Predictive ability of models.....	61
Table A-0-1 Study Data Set.....	70

List of Figures

Figure 2-1 Project Life Cycle (PMI, 2017).....	9
Figure 2-2 Typical Project Life Cycle –Traditional Design/Bid/Build (Dawson & Zeevaart, 2009)	10
Figure 2-3 Typical sequence of Principal Events during Contracts for Construction (FIDIC, 1999)	11
Figure 2-4 Some factors affecting construction project duration (Kumaraswamy & Chan, 1995)	13
Figure 4-1 Scatter Plot Matrix for Unadjusted Data.....	45
Figure 4-2 Scatter Plot Matrix for Transformed Data	47
Figure 4-3 Residual Scatter Plot - BTC	57
Figure 4-4 Residual Scatter Plot - Developed Model.....	57
Figure 4-5 Normal P-P Plot of Regression Standard Residual - BTC.....	58
Figure 4-6 Normal P-P Plot of Standardized Residual - Developed Model.....	58
Figure 4-7 Residual Histogram - BTC.....	59
Figure 4-8 Residual Histogram - Developed Model.....	59

1. Introduction

1.1. Study Background

Construction duration is the period between possession of site by the contractor and completion of the project. Construction time, along with quality and cost, is one of the three major success criteria of any construction project, be it building or any other civil structure. However, Ethiopia's construction industry is often characterized by considerable project time overrun. A recent study made by Anshebo (2017) in Addis Ababa indicates that 100% of the public building projects under the study were delayed. Studies have shown that several factors can be attributed to the delay of projects and inaccurate or unrealistic estimation of construction duration at the time of contract completion is among them.

Besides a success criterion, accurate estimation of project duration is important for both; contractor and the client. The client can create a financial, cash and material flow plan in a pre-set time and can make optimum funds available to the Project (ODABAŞI, 2009). Moreover, general contractor's ability to predict the construction time accurately and performing the works on time will gain power in the construction market and will take good decisions and take precautions against delays (ODABAŞI, 2009). Precise estimate of construction duration is essential in negotiating contract conditions and there are various methods of estimation but according to (Lin, et al., 2011) in practice, most of them depend on the subjective skills of the estimators and planners, rather than on objective assessment. Consequently, project time estimation has been the subject of interest for many researchers across the globe. In this regard, parametric models of estimation that make use of project information as basis to estimate contract time are prevalent.

The first study that tried to establish mathematical formula for construction duration prediction was Bromilow (1969), who used the cost of the building as basis to predict time, half a century ago (Bromilow's Time-Cost Model-BTC). Since then, researchers have tested the applicability of the BTC model in their respective countries with contrasting results as to the predictive ability of the model and also adopted other models by incorporating other parameters that would best predict construction duration.

The fact that predictive modelling of any kind is applicable only in the context it is developed, necessitates a model to be developed for different parts of the world reflecting the context of the locality. In this regard, there has been some effort done to develop a model for the Ethiopian context. The Building and Transport Construction Design Authority (BaTCoDA) had developed a model based on BTC in 1987 for building projects. Furthermore, (Assefa, 2008) had tried to establish a model for predicting the construction duration of public road projects. This study intends to build on these works and hopes to add to the existing body of knowledge by developing a predictive model that could be used to predict/estimate the duration of construction for public building projects in Addis Ababa.

1.2. Statement of the Problem

In practice, estimators commonly use two methods to forecast construction time and cost; (1) pursuant to the client's accessible budget and time, i.e. occupancy need, or (2) through a detailed analysis of work to be done and resources available, using estimates of the time and cost requirements for each specific activity (Skitmore & Thomas, 2003). The problem with the above methods is that, the former fails to consider the actual extent of the works while the latter is often time consuming and reliant on the subjective skills of the estimator.

In Ethiopia, the Building and Transport Construction Design Authority (BaTCoDA), the former government building regulatory body, in 1987 formulated completion time based on cost by using the data on completed projects (Assefa, 2008). However, it is hardly used by estimators because it has not been updated to reflect for inflation and contemporary practices of construction. In this regard, the Addis Ababa City government Construction Bureau, who is the Engineer/consultant for most of the public buildings in the city, has put in place a regulation to be followed whilst estimating construction duration based on the number of stories i.e. for G + 0 buildings, up to 3-months; up to G + 4, 6-12 months; and for buildings above G + 4, 12-24 months. The exact duration of construction would then be fixed by the subjective skill of the estimator based on previous projects. The downside to this method is that it only considers the number of stories to predict the construction duration irrespective of the cost or the built-up area of the building or the type of the building. The problem lies right at the center of this fact, as the estimated duration is some way off

the actual duration. This is manifested by the 100% delay the projects faced as per the findings of (Anshebo, 2017).

1.3. Objectives of the Study

1.3.1. General Objective

The principal objective of this study is to develop a best-fit model that can be used to estimate duration of public building projects in Addis Ababa based on parameters that are known to the estimator at the early stages of a project.

1.3.2. Specific Objectives

- To assess the applicability of Bromilow's Time-Cost model for estimating contract durations of public building projects in Addis Ababa;
- To propose an alternative model that incorporates other influential variables on construction duration, and compare the results with the BTC model, so that building contract durations can be estimated with reasonable accuracy.
- To determine the major factors that are determinants of construction duration of building projects;

1.4. Research Questions

To realize the objectives of the study, the following questions are raised.

- Is Bromilow's model applicable for modelling construction duration of public building projects in Addis Ababa?
- What are the major factors that affect construction duration of a building project?
- Can construction parameters other than cost be used to develop a model that can reliably predict construction duration?

1.5. Research Hypotheses

The hypotheses stemming out of the main research objectives are:

- I. Verify the applicability of Bromilow's Time-Cost model for public building projects in Addis Ababa. And

- II. Verify whether incorporating influential factors other than cost could result in an enhanced best-fit model that can predict duration

The corresponding null hypotheses that would be the subject of acceptance and rejection based on the study findings are:

H_{01} = “There is no significant relationship between project duration and project cost” And

H_{02} = “Including influential project parameters other than cost doesn’t improve the statistical significance.”

1.6. Scope of the Study

The scope of the study is bounded by three main aspects: geographical, sectoral, and project stage.

Geographical scope: the area selected for the study is Addis Ababa, the capital city of Ethiopia, owing to time and cost feasibility and relatively higher volume of construction.

Sectoral Scope: there are various sectors in the construction industry such as roads, railways, water works, residential housing, commercial buildings, public buildings and so on. In this respect the study is limited to public buildings that are managed by the government. The rationale behind this is that the projects carried out the government are assumed to have a similarity in many aspects which in turn allows the development of a single model.

Project Stage Scope: the study is limited to the pre-construction stage of the project (planning and design) where fast, practical and reliable construction duration estimate is needed to be used as the contract duration.

1.7. Significance of the Study

Estimation of construction duration at the early stages of the project is irrefutably one of the key success essentials since it forms the basis for budgeting, planning, executing, monitoring and even for litigation aims (Skitmore & Thomas, 2003). At the end of the day, the estimated duration is the basis against which the project’s success with respect to time is judged. This study, tailored towards developing a model for estimating construction duration, is believed to have paramount importance for both client and contractor. Increasing the accuracy of contract time estimates would

help the client in securing optimum funds for the project and can make practical arrangements for the occupancy of the building. The building contractor executing the works under his contract in due time will enhance its reputation in the construction industry and will make necessary provisions so as to avoid delays.

In addition, adversarial relationships among project stakeholders would be avoided as far as completion time is concerned, considering all other obligations are carried out in due course. Furthermore, the study results may also be used to form the basis for establishing, or refuting, construction claims for inadequate or unrealistic construction durations imposed by clients, a practice that is occasionally encountered, but most often lacks supporting evidence, particularly the lack of local reliable project duration forecasting models, against which disputed durations can be evaluated.

The findings of this study not only fill an important gap in predicting building construction project durations in Addis Ababa, but can be also used during the preconstruction phase to assist in feasibility studies, i.e., cash-flow management, financial planning, and budgetary control for contract and project management scopes, in addition to activity scheduling and efficient utilization of available resources. Moreover, determining the factors that can significantly impact or that have the biggest say in construction duration could be used in further studies.

1.8. Organization of the Research

The thesis is organized in five major chapters namely, introduction, literature review, research methodology, results and discussions, and conclusions and recommendations. The first part of the thesis puts forth the background information along with the study objectives and hypotheses. The second chapter presents the thorough review of literature made to establish the theoretical and conceptual frameworks for the study. The third chapter deals with the methodological framework used to realize the study objectives. The fourth chapter entails the presentation of the major results of the study in a step by step discussion. Finally, the fifth chapter presents the conclusions drawn from the study along with the recommendations forwarded to stakeholders.

2. Literature Review

2.1. Project and Project Management

2.1.1. Project: Definition

Project Management Institute (PMI) defines a project as, a temporary endeavor undertaken to create a unique product, service, or result. The terms unique product, service, or result infer that, projects are undertaken to fulfill objectives by producing deliverables. An objective is defined as an outcome toward which work is to be directed, a strategic position to be attained, a purpose to be achieved, a result to be obtained, a product to be produced, or a service to be performed. A deliverable is defined as any unique and verifiable product, result, or capability to perform a service that is required to be produced to complete a process, phase, or project (PMI, 2017).

The temporary nature of a project indicates that a project has a definite beginning and end (PMI, 2017). Projects are temporary, but their deliverables may exist beyond the end of the project. Projects may produce deliverables of a social, economic, material, or environmental nature. For example, a project to build a national monument will create a deliverable expected to last for centuries. This means that a project is done only one time. If it is repetitive, it's not a project. A project should have definite starting and ending points (time), a budget (cost), a clearly defined scope—or magnitude—of work to be done, and specific performance requirements that must be met (Heagney, 2012).

As an alternative definition, a project is a series of activities aimed at achieving specific goals within a defined budget and schedule. Every project has a certain objective (scope), has defined start and end dates (schedule) and has funding limits (budget). For the professionals of construction project management, three words, “scope, schedule and budget”, seem to be enough to define the borders of a construction project. Since budget and schedule are the main project constraints to be worked with, any estimation concerned with cost and duration are very helpful in the early stages of construction project management process (Karanci, 2010).

2.1.2. Project Management

Project management is the application of knowledge, skills, tools, and techniques to project activities to meet the project requirements (PMI, 2017). Project management is accomplished

through the application and integration of the project management processes of initiating, planning, executing, monitoring and controlling, and closing (Heagney, 2012).

One of the common causes of project failures is that the project sponsor demands that the project manager must finish the job by a certain time, within budget, and at a given magnitude or scope, while achieving specific performance levels. In other words, the sponsor dictates all four of the project constraints i.e. performance, cost, time and scope (PCTS). And according to Heagney (2012), this doesn't work.

The relationship among the PCTS constraints can be written as follows:

$$C = f(P, T, S)$$

In words, this says, "Cost is a function of Performance, Time, and Scope (Heagney, 2012).

2.1.3. Project Management Process Groups

Project management is achieved through the integration of the five project management processes of initiating, planning, executing, monitoring and controlling, and closing. The project life cycle is managed by executing a series of project management activities known as project management processes. Every project management process produces one or more outputs from one or more inputs by using appropriate project management tools and techniques. The output can be a deliverable or an outcome. Outcome is an end result of a process. Project management processes apply globally across industries (PMI, 2017).

A Project Management Process Group is a logical grouping of project management processes to achieve specific project objectives. Process Groups are independent of project phases. PMI (2017) divides project management processes into the following five Project Management Process Groups:

- **Initiating Process Group.** Those processes performed to define a new project or a new phase of an existing project by obtaining authorization to start the project or phase.
- **Planning Process Group.** Those processes required to establish the scope of the project, refine the objectives, and define the course of action required to attain the objectives that the project was undertaken to achieve.

- **Executing Process Group.** Those processes performed to complete the work defined in the project management plan to satisfy the project requirements.
- **Monitoring and Controlling Process Group.** Those processes required to track, review, and regulate the progress and performance of the project; identify any areas in which changes to the plan are required; and initiate the corresponding changes.
- **Closing Process Group.** Those processes performed to formally complete or close the project, phase, or contract.

2.1.4. Project Life Cycle

A project life cycle is the series of phases that a project passes through from its start to its completion. It provides the basic framework for managing the project. This basic framework applies regardless of the specific project work involved. The phases may be sequential, iterative, or overlapping (PMI, 2017). All projects can be mapped to the generic life cycle shown in figure 2.1.

Karanci (2010) stipulated that, when the project delivery stages of a construction project are considered, the process can be summarized in 6 different stages:

1. Feasibility Stage
2. Conceptual Stage
3. Engineering
4. Procurement
5. Construction
6. Turnover

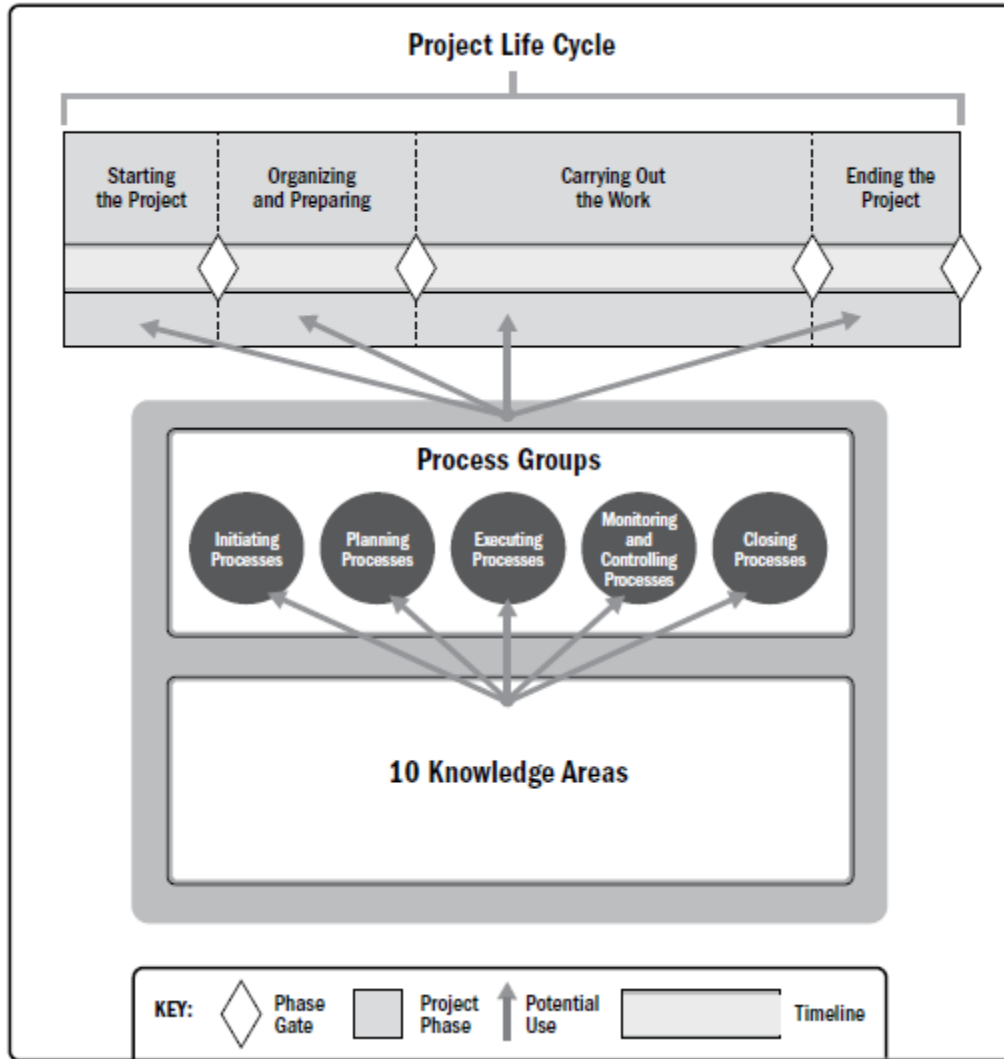


Figure 2-1 Project Life Cycle (PMI, 2017)

In a traditional design/bid/build (D/B/B) project, the project life cycle begins with the initiation of planning (including environmental and funding), and design (including conceptual design). These phases overlap to some extent. During these phases the project evolves through consideration of various alternatives and the concept for the preferred alternative is formed. The design phase continues through the preliminary engineering effort to further analyze, validate, and define the preferred alternative and arrive at the baseline scope, budget, and schedule. Then the design phase concludes with the final design, which further details the design features to provide the permitting agencies and the contractor a set of construction drawings and specifications to permit and build the project. The construction phase proceeds with the bid and award process. At the end of the

construction phase, the work of the contractor needs to be integrated with operations and Agency furnished activities, technology, and equipment, and evaluated for acceptance through the commissioning phase to bring the project to a successful completion (Dawson & Zeevaart, 2009). Figure 2-2 shows a typical project life cycle for traditional D/B/B delivery.

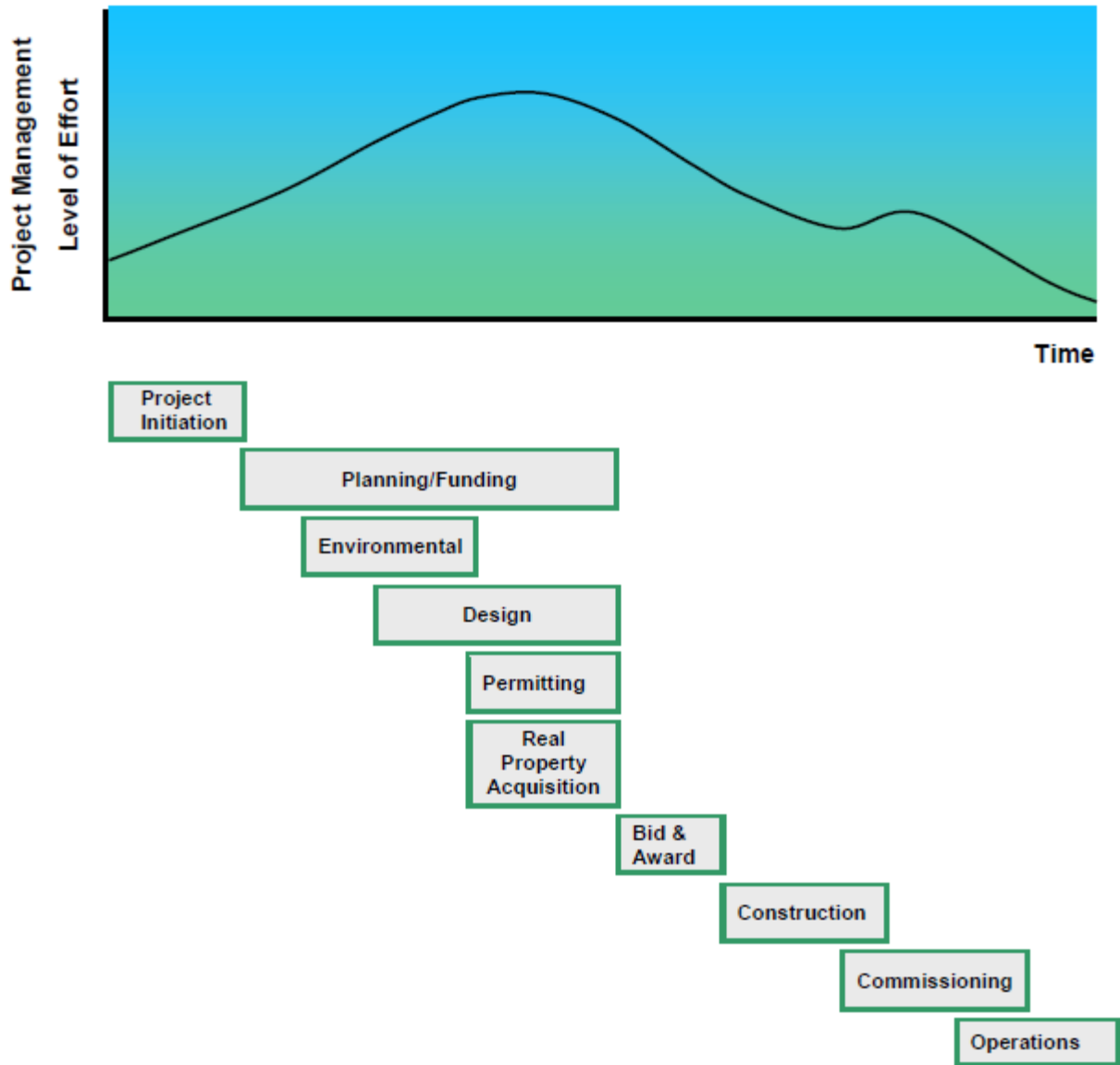


Figure 2-2 Typical Project Life Cycle –Traditional Design/Bid/Build (Dawson & Zeevaart, 2009)

2.2. Construction Duration

Although there are many definitions for construction duration, for the purposes of this study, construction duration is defined based on the work of Nkado (1995), whereby building construction duration is defined as the elapsed period from the commencement of site works to the completion time of building to the client. It is usually specified prior to the commencement of construction. The FIDIC general conditions of contract (Fédération Internationale des Ingénieurs-Conseils or International Federation of Consulting Engineers, 1999) specifies the time for completion as the period elapsed between the commencement date up to the issue of the Taking-Over Certificate including any added extension of time, as shown in figure 2-3.

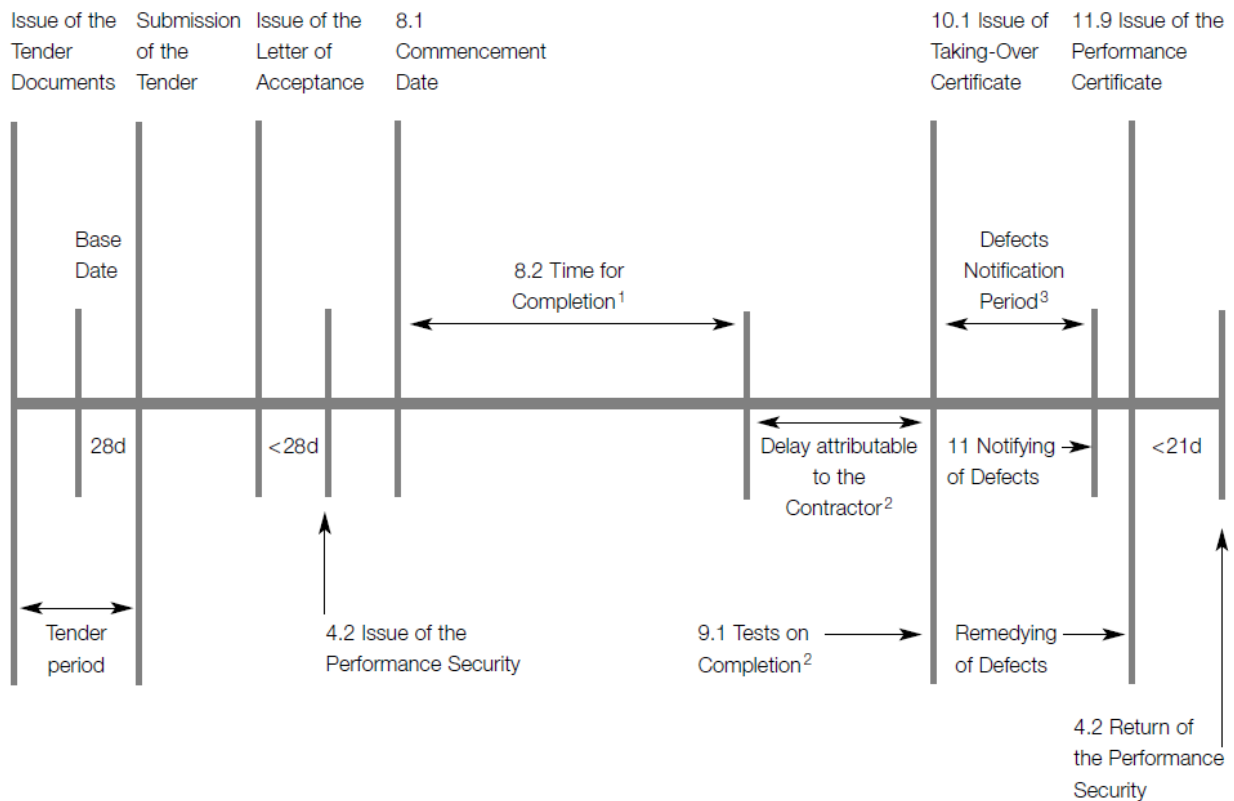


Figure 2-3 Typical sequence of Principal Events during Contracts for Construction (FIDIC, 1999)

As can be shown from Figure 2-3 and inferred from the definition, construction duration starts when the contractor receives the instruction to proceed and ends at the completion of construction works on site attested by the taking-over certificate. It also includes delays caused by unanticipated

circumstances, e.g. alteration of works (changed conditions and change orders), extra works, and supply of materials, location, weather, and site work conditions. Major changes that after the scope of work significantly are not included.

2.3. Factors Affecting Construction Duration

Odabasi (2009) indicated that, starting from the early 1970s, there were many researches into the factors influencing construction durations across various categories of projects for many reasons (ODABAŞI, 2009). The overriding reason for this is that, the first step in developing a model is selection of factors or predictor variables that are used to estimate the parameter to be modelled i.e. the response variable. As such, the determinants of construction duration have been the subject of research throughout the world. It is worth pointing out that the factors that affect duration comprise of both quantitative and qualitative variables. More often than not, the quantitative variables are used in the model whilst qualitative variables are used to categorize the projects under consideration.

The pioneer study conducted to determine the factors that affect construction duration was made by Kumaraswamy & Chan in Hong Kong (Kumaraswamy & Chan, 1995). They established a hierarchy of factors that could decide the duration of construction, as derived firstly from the experience of the authors and the literature in general and secondly from the survey in Hong Kong in particular. Construction time can be considered to be a function of all such primary, secondary and tertiary factors in the hierarchy of determinants of construction project duration. The primary factors in the hierarchy include: cost, type of construction, location, client's and other imperatives/priorities, total factor productivity, type of contract, post-contractual developments, and other factors as illustrated in Figure 2-4. The secondary and tertiary factors, as indicated in the bottom hierarchy of Figure 2-4 include, product type, technical meters, quality, complexity, risk allocation, tenderer selection method, management structure, contract modalities, variation orders, conflicts, managerial productivity, organizational productivity, labor productivity and technology.

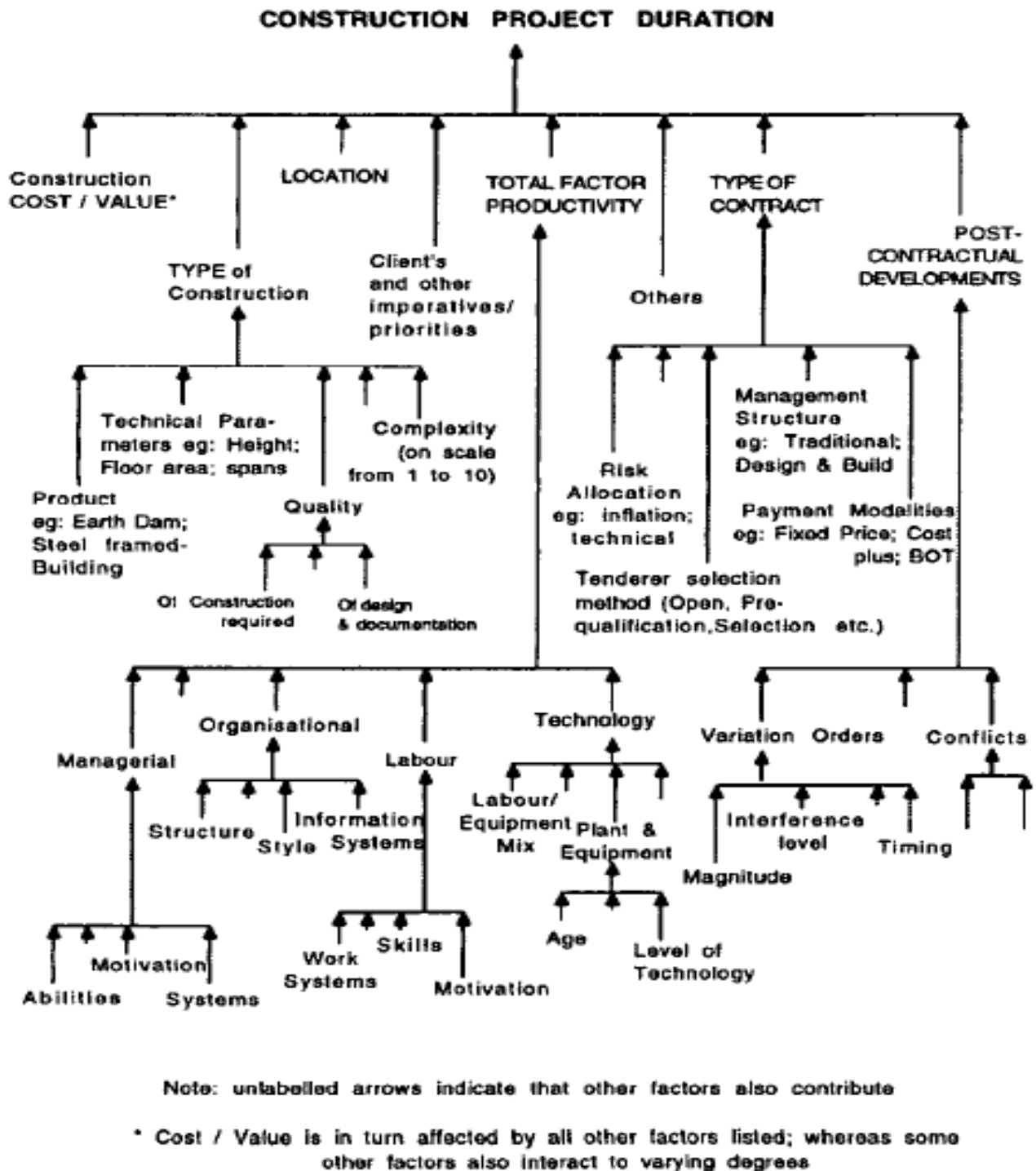


Figure 2-4 Some factors affecting construction project duration (*Kumaraswamy & Chan, 1995*)

A study made on building projects in Germany by Dursun & Stoy (2012), stipulated that gross external floor area and cost of construction works as the major variables for describing the construction duration (Dursun & Stoy, 2012), out of 1,676 observations with 30 quantitative and four qualitative variables as shown on the table below. Moreover, they stressed on the categorization with respect to variables such as type of facility, project location, availability of construction area, and market conditions, as they were found to be statistically significant (Dursun & Stoy, 2012).

Moreover, Odabasi (2009) selected the following eleven factors as the major factors affecting construction duration, from the extensive literature review conducted (ODABAŞI, 2009).

1. Cost:
2. Cash flow:
3. Productivity of on-site:
4. Procurement:
5. Project Related Factors:
6. Technology and Methodology of Construction:
7. Experience:
8. Coordination:
9. Weather:
10. Construction site:
11. The degree of completeness of design project:

A study made in Nigeria by Oyedele (2017) concluded that, construction projects duration is predictable but ‘poor cash flow, corruption, contractor type (whether professional or non-professional), year of project (election year or non-election year), getting approval from government, type of work either new build or refurbishment, insecurity, disagreement between contractors and clients or clients’ representatives, lack of teamwork among participants, type of design (whether simple or complex, low-rise or high-rise), unstable weather and client type (whether private or public)’ are the major factors adduced by respondents (construction participants) as affecting duration of construction projects in Nigeria.

In the context of Addis Ababa, Anshebo (2017) evaluated the performance of 35 public building projects carried out in the city. He surveyed the opinion of practitioners from contractor, consultant, client and regulatory body in a quest to determine the main factors affecting construction duration. According to the study, the 10 most influential factors that affect construction duration of building projects in order of relevance are: (Anshebo, 2017)

1. Contractors' organizational structure;
2. Project team work experience;
3. Project team turnover;
4. Test sample of material;
5. Labor deployment of contractor;
6. Project complexity;
7. Speed of decision making;
8. Rate of feedback;
9. Productivity of labour; and
10. Experience of contractor.

A comparison between the work done by Anshebo (2017) and the findings of the works cited above, with respect to the influential factors of construction duration, indicates that some of the factors stipulated by Anshebo (2017) are unique for the Ethiopian context. Influential factors such as complexity, productivity of labour, experience of contractor, project team work experience and contractors' organizational structure are similar to the ones mentioned in the works cited above. However, the findings of Anshebo (2017) designate that influential factors such as project team turnover, labor deployment of contractor, speed of decision making, and rate of feedback are unique to the Ethiopian construction industry. This goes to show that factors affecting construction duration vary based on the context and location.

2.4. Construction Duration Estimation Modelling

Duration of construction can be estimated in any stage of the project. As discussed above, on the basis of project delivery, a construction project has six different stages: feasibility, conceptual, engineering, procurement, construction and turnover stage. The techniques to be used for prediction are different across all the stages as the availability of information varies from one stage

to another. As the project progresses from one stage to another, more detailed information would be obtained about the project and this in turn would improve the accuracy of the duration estimate. For instance, detailed project parameters that can be used to estimate duration are available at the engineering stage compared to the feasibility stage.

In practice, there are two common methods of estimating project completion time: (1) according to the client's time constraints e.g. occupancy need, or (2) through a detailed analysis of work to be done and resources available, using estimates of the time requirements for each specific activity. Method (2) is known to be very tedious and is often impractical in view of the time limitations imposed on contractors at the tendering stage. Detailed estimating of construction activities also relies on the estimators' experience and judgment to correctly interpret project and site information and make the best possible decisions. In most cases, however, time expectation is formed based on previous experience, rather than in the context of best practice (Ng, et al., 2001).

Accordingly, studies have tried to come up with an estimation mechanism that can be used to objectively forecast the construction duration of projects. To this end regression analysis has proven to be essential in developing a simple and robust model that could predict duration based on other project parameters.

2.4.1. Regression Analysis

According to most of the works reviewed under this study, regression analysis is the most widely used technique to predict duration of construction. Regression analysis uses a statistical approach that defines a relationship between a response (dependent) variable and a predictor (independent) variable(s). In construction time estimation, the response variable is construction time, and a predictor variable is a project parameter or parameters that have dominant impact on construction time. Regression analysis is generally considered cost effective and simple compared with other estimation techniques because it is easier to modify and update the model than neural networks or other artificial intelligence– based models (Son, et al., 2018).

Regression analysis is a conceptually simple method for investigating functional relationships among variables (Chatterjee & Hadi, 2006). And according to Odabasi (2009) regression analysis (multiple) is used to express a dependent variable (y) in terms of the independent variables X_1, X_2

... X_n for investigating the relationship between a dependent variable and one or more independent variables. The equation representation is as follows:

$$y = \alpha_0 + \alpha_1 X_1 + \alpha_2 X_2 + \dots + \alpha_n X_n \quad (\text{Eq.2.1})$$

Where,

Y = dependent variable, response variable

α_0 = regression constant

$\alpha_{1, 2, 3, \dots, n}$ = partial regression coefficient of $X_{1, 2, 3, \dots, n}$

$X_{1, 2, 3, \dots, n}$ = independent variables, predictor variables (ODABAŞI, 2009)

According to some researchers such as Kim, et al., (2004), the major disadvantage of regression analysis is its accuracy as compared to other predictive estimation models such as the Artificial Neural Network (ANN). On the other hand, Nani, et al., (2017) established that there were no significant differences in accuracy between the two approaches estimating duration. Nevertheless, researchers argue that regression analysis is a preferred option for planning-phase time estimation considering its simplicity, cost effectiveness, and practical applicability compared to neural networks or other estimation methods such as simulation or fuzzy theory (Son, et al., 2018).

2.4.2. Steps in Regression Analysis

Chatterjee and Hadi (2006) had put in place eight distinct steps to be followed to develop a regression model.

1. Statement of the problem
2. Selection of potentially relevant variables
3. Data collection
4. Model specification
5. Choice of fitting method
6. Model fitting
7. Model validation and criticism
8. Using the chosen model(s) for the solution of the posed problem.

2.4.2.1 Statement of the Problem

Regression analysis usually starts with a formulation of the problem. This includes the determination of the question(s) to be addressed by the analysis. The problem statement is the first and perhaps the most important step in regression analysis (Chatterjee & Hadi, 2006). The problem statement defines the variables that must be analyzed to solve the problem. For instance, in this particular study, the problem is the inaccurate estimation of construction duration. Accordingly, the variables to be analyzed revolve around those having a significant impact on construction duration.

2.4.2.2. Selection of Potentially Relevant Variables

The next step after the statement of the problem is to select a set of variables that are thought by the experts in the area of study to explain or predict the response variable (Chatterjee & Hadi, 2006). The factors affecting construction duration, as described above, are the potentially relevant variables in that regard.

2.4.2.3. Data Collection

Having defined the problem and selecting the relevant variables that would help propose a solution to the problem, the next step would be to collect the data from the environment under study to be used in the analysis.

2.4.2.4. Model Specification

The form of the model that is thought to relate the response variable to the set of predictor variables can be specified initially by the experts in the area of study based on their knowledge or their objective and/or subjective judgments. The hypothesized model can then be either confirmed or refuted by the analysis of the collected data (Chatterjee & Hadi, 2006). Based on the pattern of the relationship between the dependent and independent variables, the regression function can be defined as linear or non-linear as shown in Eq.2.2 and Eq.2.3. respectively.

$$Y = \beta_0 + \beta_1 X_1 + \varepsilon \text{ ----- } \mathbf{Linear} \quad (\text{Eq.2.2})$$

$$Y = \beta_0 + e^{\beta_1 X_1} + \varepsilon \text{ ----- } \mathbf{Non - linear} \quad (\text{Eq.2.3})$$

Furthermore, a regression function is specified by the number independent and dependent variables it has. A regression equation containing only one predictor variable is called a simple regression equation. An equation containing more than one predictor variable is called a multiple regression

equation (Chatterjee & Hadi, 2006). When I comes to the response variable, a regression analysis containing only one response variable is called univariate regression and in cases where there are more than one response variable, the regression is called multivariate regression.

2.4.2.5. Choice of Fitting Method

After the model has been defined and the data have been collected, the next task is to estimate the parameters of the model based on the collected data. This is also referred to as parameter estimation or model fitting. The most commonly used method of estimation is called the least squares method. Under certain assumptions, least squares method produce estimators with desirable properties (Chatterjee & Hadi, 2006). Computer software packages such as Statistical Package for Social Sciences (SPSS) use the least squares method to do regression analysis.

2.4.2.6. Model fitting

The next step in the analysis is to estimate the regression parameters or to fit the model to the collected data using the chosen estimation method.

2.4.2.7. Model Validation and Criticism

The validity of a statistical method, such as regression analysis, depends on certain assumptions. Assumptions are usually made about the data and the model. The accuracy of the analysis and the conclusions derived from an analysis depends crucially on the validity of these assumptions (Chatterjee & Hadi, 2006).

2.4.2.8. Using the Model

The explicit determination of the regression equation is the most important product of the analysis. It is a summary of the relationship between Y (the response variable) and the set of predictor variables X_1, X_2, \dots, X_n . The equation may be used for several purposes. It may be used to evaluate the importance of individual predictors, to analyze the effects of policy that involves changing values of the predictor variables, or to forecast values of the response variable for a given set of predictors (Chatterjee & Hadi, 2006).

2.4.3. Bromilow's Time-Cost Model (BTC)

The pioneer in developing a model based on regression analysis to predict construction duration is Bromilow (1969). He used cost as a determinant to predict time taken by construction, by using

data from 329 completed building projects in Australia (Bromilow, 1969). His model which is based on the power regression formula is:

$$T = KC^B \quad (\text{Eq.2.4})$$

Where,

T = construction time (in days),

C = estimated cost of project (in million dollars),

K = a constant describing the general level of time performance,

B = a constant describing how the time performance was affected by project size as measured by cost.

The rationale behind the BTC is cost and time are highly correlated and a model developed with such relationship would reduce the subjectivity in duration estimation. The model was intended to be used by clients and contractors for estimating and benchmarking the contract period of construction projects, thereby reducing the subjective duration estimation approach.

Bromilow's model is supported by the logic that the time required to complete a given project is highly correlated with the project's size which in turn is highly correlated with the cost of the building. However, a major shortcoming of this model stems from the fact that besides cost, other factors may significantly impact the duration of a project, e.g., contract procurement method, project type and sector, construction method, construction area, number of stories above and below ground, buildability level, and productivity, in addition to the quality and quantity of site management (Jarkas, 2015).

The assertion that the BTC fails to consider other parameters other than cost is confirmed by different studies such as Lin, et al., (2011), Love, et al., (2005), and Hoffman, et al., (2007). Further to the above, Hoffman, et al., (2007) indicated the importance of partitioning of the Data to reflect different project characteristics while developing a model based on BTC.

Moreover, Odabasi (2009) argued that cost could not be used as a sole predictor of duration. For instance, two building projects having the same design but different quality would have different construction cost while their duration could be the same. On the other hand, two different construction projects having different cost values may take the same time to construct (ODABAŞI, 2009). Another possibility is that two different project having same cost values may take same

time to construct because of the different working productivity or experience of different construction teams.

Consequently, selection of the inputs to be used to develop the regression model to predict construction duration has continued to be a formidable challenge. As such it is worthwhile to go through previous studies to look into the prominent parameters that have significant impact on duration and hence can be used to model duration prediction.

2.4.4. Previous Studies

Over the years researchers across the world, have tried to develop a model to predict construction duration based on Bromilow's principle and came up with contrasting results. Some have found the time-cost relationship as a sound principle while others found time-cost relationship not having a statistical significance and used other parameters that have a significant impact on duration.

In a quest to determine whether the time-cost relationship held, Bromilow carried out another study in 1980. He studied 408 building projects done between 1970 and 1976 in Australia, of which 290 were government projects and 118 were private ones. The results showed that the relationship between construction duration and project cost in the 1960s still holds.

Kaka and Price (1991) studied data collected from 661 building projects and 140 road projects in the UK, built in the period of 1984 to 1989, to test Bromilow's time-cost model. They classified the data into groups based on the project type, form of contract and form of procurement. Under this study, the BTC was found to be valid for public buildings and civil engineering projects, while private buildings varied significantly (Kaka & Price, 1991).

Chan and Kumaraswamy (1999) took data from 56 public housing projects completed between 1990 and 1994 in Hong Kong to develop a model for predicting duration (Chan & Kumaraswamy, 1999). This work was a continuation on their previous study that established the determinants of construction duration (Kumaraswamy & Chan, 1995). They considered 94 (84 pre-identified and 10 derived) variables, both qualitative and quantitative, affecting construction duration. They developed a multiple linear regression model by using the stepwise regression analysis and the resulting model was based on construction cost; presence/absence of precast facades; height of building; nature of site; and type of scheme (rental/ purchase).

$$\text{Loge EST_TIME} = 2.6031 + 0.0834 \log \text{ESTCOST} + \text{FACADE (0 for with facades; 0.0497 for without facades)} + 0.0024 \text{HEIGHT} + \text{NATSITE (0.2352 for level; 0.2221 for built-up; 0 for sloping)} + \text{TYPESCH } (\pm 0.0453 \text{ for purchase; 0 for rental}) \quad (\text{Eq.2.5})$$

Ng, et al., (2001) used records of 93 completed building projects constructed between 1991 and 1998 in Australia to check whether the BTC model still fit. The data also allowed for partition according to the method of contractor selection (selective tendering, open tendering, negotiated price, etc.), type of project (recreational buildings, industrial buildings, educational buildings, residential buildings, etc.) and type of contractual arrangements (lump sum, design and construct, construction management, etc.) enabling further regression models to be fitted (Ng, et al., 2001). Their study stipulated that the BTC model was still applicable and the classification of the data didn't show considerable difference except for the type of the project, industrial and non-industrial. This study, unlike its predecessors, showed that, while project cost continued to be the major parameter to predict duration and there is a need to develop different models based on project type.

Subsequently, additional studies to further calibrate the time-cost model for use across a variety of project types and locations have been performed. These efforts confirmed the continuing validity of Bromilow's initial model, in which the time-cost relationship is highly correlated. However, the studies also highlighted the value of classifying data to develop different parameter values (K and B) for varying project characteristics. Classification thereby results in a number of separate models that, when taken as a whole, are able to account for a variety of factors influencing construction duration. Instead of identifying relevant factors and partitioning the data accordingly, several studies have attempted to improve the accuracy of the BTC model by incorporating predictor variables i.e., factors having a significant influence on the dependent variable i.e., duration. These additional variables have included both quantitative scope-related as well as qualitative management-related factors.

The major study that defied the applicability of the BTC model was carried out by Love, et al., in 2005. The project time–cost relationship was examined by using project scope factors as predictors for 161 building construction projects that were completed in various Australian States in 2000. Project scope factors included project size, project type, facility type, contract value, project

duration, procurement method, the tender type, Gross floor area (GFA) and number of floors (i.e., project height), and their influence on the time and cost performance was examined. The analysis was performed between project duration (i.e., time), project type, procurement method, tender type, gross floor area (GFA), and number of stories. They developed an alternative model to the BTC based on GFA and the number of stories. Furthermore, the results indicate that cost is a poor predictor of time performance (Love, et al., 2005). The resulting model with root-mean-square error of 221.15, mean absolute error of 130.3, and mean absolute percentage error of 50.4 is:

$$\text{Log (T)} = 3.178 + 0.274 \log(\text{GFA}) + 0.142 \log(\text{Floor}) \quad (\text{Eq.2.6})$$

The study developed a multiple linear regression models and found that GFA and number of stories are key determinants of time performance in projects. The other important result is that cost was found to be poor indicator of time performance because it is not possible to know the cost at the start of the project where the cost by itself is an estimate. Although Love et al. (2005) found the cost parameter as an insignificant factor; they found the BTC model to be applicable for estimation during the inception of the project..

A study made by Hoffman, et al. (2007), aspired to gain insight into the significant factors impacting construction duration by developing a regression model. Data were collected for 856 facility projects completed between 1988 and 2004 in the United States of America. These data were analyzed using Bromilow's time-cost BTC model as well as multiple linear regression. The multiple linear regression model was found to provide the most acceptable prediction. A forward stepwise procedure (with a p-value of 0.05) was used to identify statistically significant variables for inclusion in the model. Using a natural logarithmic transformation, the final model contained cost as the lone quantitative variable and was in the form:

$$y = 3.44 + 0.198x_1 - 0.059x_2 - 0.070x_3 - 0.222x_4 - 0.193x_5 - 0.0146x_6 \quad (\text{Eq.2.7})$$

Where y = project construction duration (ln days); and x_1 = project total cost (ln \$). The remaining x_i values represent dummy variables for Air Combat Command (ACC), Air Education and Training Command (AETC), Air Force Special Operations Command (AFSOC), Northwestern COE region, and in-house design/construction agent, respectively (Hoffman, et al., 2007).

A study made in Turkey by (ODABAŞI, 2009) examined seven case study buildings that are located at the Middle East Technical University (METU) campus in Ankara. Based on the remarks

of previous works, the study used 10 project parameters to develop a multiple linear regression model. The parameters were: number of block, number of floor, total height, average height of floors, total area, average floor area, total volume, average volume of floors, façade area and adjusted detailed cost values. The developed model was based on the area of the building and the area of its façade (ODABAŞI, 2009), as shown below.

$$T = 53.73 + 0.03xTot.Area + 0.07xFaçade Area. \quad (Eq.2.8)$$

Lin, et al., (2011) used data from 56 buildings in Taiwan to develop a duration prediction model. A logical approach was employed to select the “good” regression model when the contract cost, the gross floor area and stories are known, while the number of change orders and the number rainy days are rationally estimated (Lin, et al., 2011). The final regression model based on the 56 cases, with an adjusted R2 value of 0.920, is:

$$Y_i = 1.238 + 0.218X_1 - 0.00092X_2 + 0.02177X_3 + 0.663X_4 + 0.01169X_5 \quad (Eq.2.9)$$

Where: Y_i is log (predictive construction duration); X_1 is log (contract initial cost); X_2 is GFA / expected contract duration; X_3 is number of stories; X_4 is modified contract duration (estimated rainy days + expected contract duration)/ expected contract duration; X_5 is change order; (Lin, et al., 2011)

Ameyaw et al. (2012) assessed the applicability of the BTC for building projects in Ghana. Historical cost data on 62 completed building projects carried out within the period 2000 - 2007 were therefore assessed for achievement of the aim of the study. The analyses revealed that the original model developed by Bromilow is not applicable to building projects in Ghana. Later on, a study made by Moses (2014) stipulated that the BTC was applicable for feeder road projects in Ghana. This goes to show that not only the BTC model’s applicability varies from country to country but also between different project types.

A study made in Vietnam by (Le-Hoai, et al., 2012) tried to develop a model based on multiple regression. Based on a data set of 70 projects (outliers discarded), the multiple regression technique is applied to integrate significant variables into a model used to predict or evaluate the project time index. The result suggests that six variables, namely underground site condition, project management works, estimating works, competency of subcontractor(s), accuracy and

completeness of design, and owner’s project financing emerged as important factors in regression model. The final model is:

$$\begin{aligned} T.I. = & 1.622 - 0.072 \times \text{Underground site condition} - 0.019 \times \text{Project} \\ & \text{management works} - 0.039 \times \text{Estimating works} - 0.033 \times \text{Accuracy and} \\ & \text{completeness of design} - 0.019 \times \text{Competency of subcontractors} - 0.013 \times \\ & \text{Owner's project financing} \end{aligned} \quad (\text{Eq.2.10})$$

In other studies, (Mizanur, et al., 2014), (Waziri & Yusuf, 2014), and (Sriana & Hayati, 2015) tested the applicability of the Bromilow Time-Cost model and verified its predictive ability in Bangladesh, Nigeria, and Indonesia respectively.

Jarkas (2015) collected data on 113 residential and 74 office buildings, which were completed during the period 2004–2010 in the state of Kuwait to analyze the applicability of the BTC model. He developed another multiple linear regression model to compare the predictive ability with the BTC model. The alternative model was based on project construction area as well as the number of floors above and below ground. The multiple regression model provided is as shown below in Eq.2.9 and Eq.2.10 for residential and office building types respectively.

$$T = 338.70 + 0.005574CA + 13.95NFAG + 27.58NFBG \quad (\text{Eq.2.11})$$

$$T = 453.97 + 0.0103CA + 1.92NFAG + 58.54NFBG \quad (\text{Eq.2.12})$$

While the results obtained further affirm the validity of the BTC model, a comparison between the outcomes of both models demonstrates that, on average, the multifactor model proposed is associated with an enhanced prediction accuracy of 16% and 14% for residential and office buildings, respectively, and therefore can be deemed more suitable for forecasting local construction duration for the building types investigated (Jarkas, 2015).

A study was carried out in Turkey by Bayram (2016) in a quest to investigate and to validate the ‘Bromilow’s Time-cost (BTC) model’ and the ‘Love et al.s Time-Floor (LTF) model’ for Turkish public building projects, to check their appropriateness and to compare their prediction performances on project duration. The study collected data from 530 public building projects in Turkey through archive search work. The study used 80% of the total data for the model development while the remaining 20% of the data were used for the validation of the model. The parameters collected in the data set are; total construction area, building height, gross floor area,

floor numbers, contract sum, actual cost, contract duration and the actual duration. According to the findings of the study the BTC model was found to be a better fit than the LTF model, which meant that cost is a more significant predictor of duration than gross floor area & floor numbers. The study goes on to develop a best-fit model that incorporated total construction area (TCA), building height (BH) and contract cost (CS).

$$T = 0.72 \times CS^{0.3352} \times TCA^{0.0926} \times BH^{0.1167} \quad (\text{Eq.2.13})$$

Mackova et al., (2017) performed a study to develop a model to predict duration based on computer experiment. The study used project parameters such as gross floor area, number of stories and floor area of one storey as variable inputs in prediction while taking into account the intensity of the deployment of labour resources (Mackova, et al., 2017). A data set was developed based on a computer experiment and multiple linear regression analysis was used to develop duration prediction model. The resulting model, with R^2 value of 0.808, is:

$$T = 151 \times NOF^{0.64} \times FA^{0.29} \quad (\text{Eq.2.14})$$

Where T = construction duration (in days),

NOF = number of floors,

FA = floor area of one story (in thousand m^2) (Mackova, et al., 2017)

A different approach with respect to the selection of model inputs was put forth by a study in Ghana by Nani, et al., in 2017. They argued that the major work items in the bill of quantities (BOQ) were hardly tried as model inputs in previous studies and hence tried to develop a model to predict duration of bridge construction based on items in the BOQ. The identified predictors, namely; in-situ concrete, weight of prefabricated steel components, gravel sub-base and haulage of aggregates used as independent variables resulted in the development of a regression model with a mean absolute percentage error (MAPE) of 25% (Nani, et al., 2017). They also developed an Artificial Neural Network (ANN) model and compared it with the regression model, with the regression model having a slightly better result than the ANN. The developed model is:

$$T = 799.513 + 0.373*IC - 0.006*BW + 0.076*SB + 0.001*H \quad (\text{Eq.2.15})$$

Where T is duration in days,

IC is the total volume of in –situ concrete in m³ in a BOQ.

BW is the total weight of the steel bridge component in kilograms

SB is the total quantity of gravel sub- base in m³ in a BOQ and

H is the total haulage of aggregates in m³ –km.

Yeom, et al., (2018) developed a model to predict construction duration for general office buildings in Korea. They determined the independent variables that have an impact on duration of the building projects under the study by thorough correlation analysis and multiple linear regression. Subsequently, they were able to develop a statistically significant regression model, based on the independent variables, which can be used to forecast construction for the general office buildings under the study, during the project planning phase, as shown in Equation 2.14.

$$T = 897.272 + 0.004 * \text{Total Floor Area} + 0.003 * \text{Building Area} + 9.896 * \text{Number of Above Ground Floors} + 40.646 * \text{Number of Basement Floors} - 437.319 * \text{SRC} - 578.234 * \text{RC} - 110.37 * \text{Winter Season Construction Commencement} \quad (\text{Eq.2.16})$$

Here, T represents the approximate construction duration, and the variables reflected in the prediction model comprise a total of seven types – total floor area, building area, number of floors above the ground, number of basement floors, steel framed reinforced concrete (SRC), reinforced concrete (RC), and winter season construction commencement (Yeom, et al., 2018). The units of the total floor area and construction area are m² and the units of the number of floors above the ground and basement are floors. The model was found to be excellent in applicability with high statistical significance with R² value of 0.905 and adjusted R² value of 0.882.

Table 2.2. below summarizes the studies in chronological order along with the various independent variables that were used in developing their respective models.

Table 2-1 Summary of inputs used to model construction duration

No	Author	Year	Country	Model Inputs/Independent Variables
1	Bromilow	1969	Australia	Cost
2	Bromilow	1980	Australia	Cost
3	Kaka and price	1991	UK	Cost

4	Chan & Kumaraswamy	1999	Hong Kong	Construction cost; Presence/absence of precast facades; Height of building; Nature of site; and Type of scheme (rental/ purchase)
5	Ng et al.	2001	Australia	Cost
6	love et al.	2004	Australia	Gross Floor Area; and Number of stories
7	Hoffman et al.	2007	USA	Cost; Building type; and Designer type (In-house/construction agent)
8	Stoy et al.	2007	Germany	Gross floor area; Number of winters; and Planning time
9	Abu Hammad et al.	2008	Jordan	Project type; Job type; Project area; Original bid price; and Original project duration
10	Odabasi	2009	Turkey	Area of the building; and Façade area
11	Ameyaw et al.	2012	Ghana	Cost
12	Lin et al.	2011	Taiwan	Contract initial cost; Gross floor area; Expected contract duration; Number of stories; Modified contract duration (estimated rainy days + expected contract duration; and change order
13	Le-Hoai et al.	2012	Vietnam	Underground site condition, Project management works, Estimating works; Competency of subcontractor(s); Accuracy and completeness of design; and Owner's project financing
14	Choudhury	2012	USA	Construction cost; Gross floor area; Number of floors
15	Mizanur et al.	2014	Bangladesh	Cost
16	Waziri and Yusuf	2014	Nigeria	Cost
17	Moses	2014	Ghana	Cost
18	Sriana and Hayati	2015	Indonesia	Cost

19	Jarkas	2015	Kuwait	Construction area; Number of floors above ground; and Number of floors below ground
20	Bayram	2016	Turkey	Total construction area; Height; and Cost
21	Mackova et al.	2017	Slovakia	Number of floor; and Floor area
22	Nani et al.	2017	Ghana	In-situ concrete; Weight of prefabricated steel components; Gravel sub-base; and Haulage
23	Yeom et al.	2018	Korea	Total floor area; Building area; Number of above ground floors; Number of basement floors; Structural form (steel framed reinforced concrete or reinforced concrete); and Winter Season Construction Commencement

As can be seen from the above studies, cost is the most used parameter in the model development, either alone or in combination with other determinants of construction duration. Other parameters that were mostly used include, gross floor area, floor height, number of floors above and below ground as well as other qualitative variables.

As can be inferred from the reviewed studies, although the predictive ability of Bromilow’s time-cost model to forecast construction project duration has been affirmed by several research findings, its applicability across different geographical settings, is questionable, as the contradictory findings reviewed above suggest. For instance, as discussed above, in Ghana the BTC was valid for feeder road projects while it didn’t hold true for building projects. Several factors related to a local industry may further impact construction duration, e.g., procurement type, construction methodology, inflation, availability of resources, prevalent weather conditions, cultural practices, technology and so on. The major shortcomings of the BTC as inferred from the above studies are:

1. Failing to incorporate other project parameters other than cost; and
2. Using cost to estimate time, while the cost by itself is an estimate.

Accordingly, validating the applicability of the time-cost model to a specific setting is of paramount importance, before it can be applied to either estimate or decide the most reasonable

construction duration. Furthermore, incorporating other project parameters other than cost would go a long way to address the shortfalls of the BTC.

2.4.5. The Ethiopian Context

In Ethiopia, there has been some effort made towards establishing a time-cost relationship for projects. A comprehensive study to test the applicability of the BTC model on road projects in Ethiopia was carried out by Assefa (2008). The study used data from 33 completed federal road projects in the country. He divided the projects as those carried out by International Contractors (IC) and Domestic Contractors (DC) as international contracts to allow for price escalation (21 projects done by ICs and 12 projects done by DCs). He checked the existence of a statistically significant relationship between contract time and contract cost as well as final time and final cost. The study findings show that the BTC model only holds true for road projects made by ICs. The study stipulated that the main reason for this distinction was because the international contracts had a provision for price escalation.

Moreover, the study tried to develop a multiple linear regression model based on cost and project length and concluded that the addition of project length didn't improve the model for projects carried out by ICs. On the other hand the inclusion of the length parameter did improve the model for projects made by DCs, although the developed model failed the t-test and hence was rejected.

In a quest to realize its objectives, Assefa (2008) continued his study to develop regression model by altering the transformation of the data in different forms. The forms were: linear equation (LIN), logarithmic equation (LOG), inverse equation (INV), quadric equation (QUA), cubic equation (CUB), composite equation (COM), power equation (POW), S-curve equation (S) and exponential equation (EXP). He found that the cubic transformation was the better option for both IC and DC projects and the resulting models are:

ALL Projects (ICs)

$$\text{Log}(T) = -32.759 + 6.1661 \text{Log}(C) - 0.0268 (\text{Log}(C))^3 \quad (\text{Eq.2.17})$$

AC (Asphalt concrete) Projects (ICs)

$$\text{Log}(T) = -47.058 + 8.7640 \text{Log}(C) - 0.0394 (\text{Log}(C))^3 \quad (\text{Eq.2.18})$$

Where T – Time in calendar days and C – cost in Ethiopian Birr (in million ETB)

ALL Projects (DCs)

$$\text{Log}(T) = 15.0615 + 20.2224(L) - 0.2631(L)^2 + 0.0013(L)^3 \quad (\text{Eq.2.19})$$

Where T – Time in calendar days and L – Project Length in Km

Assefa (2008) also reported that, “The Building and Transport Construction Design Authority (BaTCoDA), the former government building regulatory body, in 1987 formulated completion time based on cost by using the data on completed projects.” The formula is:

$$T = 7 \text{ to } 12 C^{1/3} \quad (\text{Eq.2.20})$$

Where:

7 - Stand for small sized buildings and 12- taken for most complex building projects.

T – Actual completion time in months

C – Actual project cost Birr in millions (Assefa, 2008)

Furthermore, (Assefa, 2008) goes on to mention another undergraduate research made by Andinet et al (2006) on 39 educational buildings in Ethiopia developed the following formula.

$$\text{Time} = 7.06 * \text{Cost}^{0.47} \quad (\text{Eq.2.21})$$

To summarize,

- A reliable construction duration estimation technique/model is important for the success of a construction project;
- Several factors could affect the construction duration of building projects and these factors can be used to develop a model that can predict duration before construction commences;
- Regression models are widely used to develop a model to predict construction duration in a reliable and practical fashion;
- Time-cost model (BTC) developed by (Bromilow, 1969) is the baseline to which the findings of most developed models are compared;
- Although many researches prove the existence of a profound correlation between time and cost some have proved otherwise and developed other models that incorporate other parameters that can best-predict construction duration. Notwithstanding the documented success of applying the time-cost model to predict project durations in some geographical settings, numerous studies reported poor prediction ability of the model in other parts of

the world. Furthermore, due to the differences in socioeconomic and political conditions, construction business environment, weather and geophysical conditions, and technological advancements between one country and another, it becomes necessary to validate the model before it can be successfully applied within a particular environment. and

- Little work has been done in Ethiopia to establish a model that can predict construction duration of building projects as the major work done in this subject was on road projects. Hence, practitioners rely on their subjective skills and client's demand as basis to estimate durations of construction at the early stage of a project.

Therefore, it is against this backdrop this study intends to base and fill the gap by developing a robust model that can be used to predict construction duration of public building projects in Addis Ababa, which is also believed to add to the global body of knowledge.

3. Research Design and Methodology

3.1. Introduction

The previous chapter described the concepts behind project duration and estimation models along with a thorough survey of previous studies. It also highlighted the gap in previous research in the context of the study and established the theoretical framework. This chapter outlines the methodological framework used to realize the research objectives.

3.2. Study Design

A research design is the arrangement of conditions for collection and analysis of data in a manner that aims to combine relevance to the research purpose with economy in procedure. In fact, the research design is the conceptual structure within which research is conducted; it constitutes the blueprint for the collection, measurement and analysis of data. As such, the design includes an outline of what the researcher will do from writing the hypothesis and its operational implications to the final analysis of data (Kothari, 2004).

The methodological design tailored to realize the research objectives is that of a combination of descriptive and analytical study under the umbrella of observational studies. As the major objective of the research is to develop a practical and reliable model that can be used to estimate construction duration, quantitative data on past building projects would be obtained and analyzed to develop an inferential model based on the variables that have a significant impact on duration. Furthermore, the descriptive nature emanates from the other objective of the study which is, to determine the major factors that are the determinants of construction duration of public building projects. Henceforth, methods associated with descriptive and inferential studies were deployed in the study as detailed in this chapter.

3.2.1. Study Approach

In research design, one has to decide the methodological approach in finding solutions/answers to the research problem or research questions. Khotari (2004) states that there are two common types of approaches in research design i.e. quantitative approach and qualitative approach. The former involves the generation of data in quantitative form which can be subjected to rigorous quantitative

analysis in a formal and rigid fashion, while the latter is concerned with subjective assessment of attitudes, opinions and behaviours. This study used quantitative approach, as quantifiable data on past project records related to construction duration was collected to develop the model.

3.2.2. Study procedure

To realize the research objectives the following sets of procedures were followed.

1. **Literature Review:** a thorough review of literature was carried out to explore the existing body of knowledge on project duration estimation modelling.
2. **Selection of Model Inputs:** based on the findings of the literature review the dominant determinants of construction duration were selected.
3. **Document Review:** past project records were then reviewed at relevant government offices to gather the data set needed to develop the model.
4. **Model Development:** the data set gathered in step-3 was then fitted into the BTC model and also was used to develop a regression model by incorporating other predictor variables, thus realizing the study objectives.
5. **Model Validation:** the model is then checked for statistical validity based on parameters that would be mentioned in detail later in this chapter.

3.2.3. Data collection

Ideally, the researcher and the existence of the research will have no influence on the data collected. The main goal of devising a mechanism to collect data is therefore to minimize the influence of the research on the data. Such minimization is sought by using objective methods designed to remove as much bias as possible and to conduct the research in the most unobtrusive way, whilst retaining goodwill of the collaborators and subjects of study – essential in studies of people and their behaviours (Fellows & Liu, 2008). The data source for this study is that of secondary data. A thorough literature review was carried out to assemble the theoretical framework for the study by reviewing similar works done and establish the parameters to be used as inputs in the development of the model. Furthermore, archival analysis of project records was carried out at relevant government offices to gather data on the parameters to be used in the development of the model.

3.2.3.1. Study area

As the title of the study indicates, the area selected for the study is Addis Ababa, the capital city of Ethiopia. Addis Ababa was selected on the basis of time and cost feasibility perspective. Moreover, it is assumed that the building construction activity in the capital is relatively greater than other parts of the country. Hence, an empirical formula developed to predict contract durations for building projects in Addis could be considered as a baseline for others. It is worthwhile to mention that a model developed for a particular context is only applicable for the study area. Accordingly, due attention shall be given to the variance in the location, to reflect the changes, before applying this model for other areas of the country.

3.2.3.2. Sampling Design

The objective of sampling is to provide a practical means of enabling the data collection and processing components of research to be carried out whilst ensuring that the sample provides a good representation of the population; i.e. the sample is representative. A sample design is a definite plan for obtaining a sample from a given population. It refers to the technique or the procedure the researcher would adopt in selecting items for the sample (Kothari, 2004). Some of the sampling techniques include deliberate sampling, simple random sampling, systematic sampling, stratified sampling and quota sampling.

According to Kothari (2004), the choice of the sampling technique to be used is based on type of universe, sampling unit, source list, size of sample, parameters of interest, budgetary constraint and sampling procedure. In this regard, the population of the study, as the title indicates, is completed public building projects in Addis Ababa. Two government offices overlook the construction of public buildings – not including housing projects - Addis Ababa City Administration Construction Bureau (AACACB) and Federal Government Buildings Construction Project Office (FGBCPO). But the Housing projects are administered independently by Addis Ababa Housing Development and Administration Bureau. Under this study, the housing projects were not included for the study as they have their own uniqueness when compared to other office buildings, with respect to budget, contract type and building types.

During the data collection phase of the study, the number of completed projects were found to be around 34 projects, 27 from the AACACB and 7 from the FGBCPO. Although there are other

projects done by these government agencies, they were not included in the population due to either their data being not available or provisional acceptance certificate not issued in light of unresolved contractual issues.

When it comes to the type of sampling technique, there are two types of sampling, probability (random) and non-probability sampling. This research uses probability sampling as it is the purest form of probability sampling in light of each member of the population having an equal chance of being selected.

3.2.3.3. Sample Size

There are many formulas to determine sample size to draw data from a given population. However, in this study, on account of the population being considerably small and on the merit of the increase in the accuracy of the developed model, this study has used all of the population in the development of the model. Accordingly, the sample size of this study was fixed at 34 completed projects.

3.2.3.4. Description of variables

The inferential nature of the study whereby, a model for predicting a parameter based on other project parameters, demands defining the research variables, i.e. dependent variables and independent variables. The hypotheses were tested by manipulating the independent variables and studying their impacts in predicting the dependent variable, which is the core objective of the research.

Dependent Variable

A dependent variable is the subject of the research, experiment or test (the output of the system); its attributes (quantity, quality) are affected by (depend upon) the attributes of the other variables with which it is associated, especially the independent variable. Dependent variables are also known as response variables, responding variables, explained variables, criterion variables or regressands (Fellows & Liu, 2008). The dependent variable that would be reliant on the independent variables is construction duration.

Independent Variables

An independent variable is a phenomenon, the attributes of which are controlled, selected and (possibly only) measured by the researcher so that any corresponding attributes of the dependent

variable can be examined; the relationship between the independent variable and the dependent variable are examined for nature, size and, most especially, for causality. Independent variables are also known as predictor variables, controlled variables, manipulated variables, explanatory variables or regressors (Fellows & Liu, 2008). The independent variables are those parameters related to construction of buildings that have an impact on duration. As discussed in the literature review, construction duration is a function of various parameters as can be seen in Eq. 3-1.

$$T = f(X_1, X_2, \dots, X_n) \quad (\text{Eq.3-1})$$

Where,

T = Construction duration (dependent variable)

X_1, X_2, \dots, X_n = Independent variables (Model inputs)

In choosing the types of independent variables to be used in the development of the model, one has to give due attention to the problem of singularity and multi-collinearity. Singularity occurs when a variable, retained in a model, is a combination of two or more variables retained in the same model whereas multi-collinearity is present in case high correlation exists between explanatory variables that are included in a model. When detected, unstable regression parameters due to large standard errors are observed. Also, model parameters present dramatic variations caused by small modifications in the sample (Dursun & Stoy, 2012). Accordingly, based on the review made on previous works, this study used the following independent variables to be used as model inputs.

- a) **Cost (C):** the total cost of the building expressed in Ethiopian Birr ETB. The study incorporated both the contract and actual project costs to make respective inference.
- b) **Average Floor Area (A):** the average external floor area of the building measured at each floor level, expressed in square meters, and measured as per Royal Institution of Chartered Surveyors (RICS) professional standards and guidance, code of measuring practice (RICS, 2015).
- c) **Gross Floor Area (A_g):** the total floor area of the building calculated by multiplying the average floor area by the total number of floors, expressed in square meters.
- d) **Average Floor Height (FH):** the average height of each floor measured from floor to floor, expressed in meters.

- e) **Number of Floors below Ground Level (NoF_{bg}):** the total number of floors below the ground level i.e. basement floors.
- f) **Number of Floors above Ground Level (NoF_{ag}):** the total number of floors above the ground level.
- g) **Total Number of Floors (NoF):** the total number of floors of the building.
- h) **Building Height (BH):** the total height of the building from basement to the roof, expressed in meters.
- i) **Building Volume (V):** the total volume of the building calculated by multiplying the total building height by the average floor area, expressed in cubic meters

Although some of the variable are linearly correlated with other, thereby posing the problem of multicollinearity, they are included to check whether different combinations of variables would result in a better model. For instance, Gross floor area is a product of average floor area and number of floors. However, the best-fit mode may be found by combining either gross floor area or average floor area with other variables.

3.2.4. Data analysis

Data analysis is a process of inspecting, cleansing, transforming, and modelling data with the goal of finding useful information, informing conclusions, and supporting decision-making. Data analysis forms the basis for accepting or rejecting the hypothesis. With respect to this study, the data on past project records pertaining to project duration is used to develop the model. The model was developed by using regression analysis.

As mentioned in the literature review, multiple regression is used to express a dependent variable (Y) in terms of the independent variables $X_1, X_2 \dots X_n$ for investigating the functional relationship between a dependent variable and two or more independent variables (ODABAŞI, 2009) as per Equation 2-1 shown in chapter 2.

Three popular iterative search algorithms for choosing a best-fir regression model are forward selection, backward elimination, and stepwise regression (Froby, 2008). Forward selection starts with no input variables and then add one input variable at a time. The input variable chosen to be added to the regression equation is the one which makes the largest contribution to R^2 on top of the input variables already in the equation. In backward elimination we start with all of the input

variables in the regression equation and then at each step we eliminate the input variable that reduces R^2 by the least amount when dropping the input variable from the regression.

Finally, there is the stepwise regression procedure which is like the forward selection and backward selection procedures except the once in, always in and once out, always out rules of forward and backward selection respectively, are no longer invoked and variables can come in and drop out, and even can come back in again until there is no remaining variable in the equation that is not statistically significant at the required levels (Froby, 2008).

Out of the three methods above, stepwise regression was used for the study to automatically develop the best fit model for predicting construction duration. Statistical Package for Social Sciences (SPSS) version 25 was used for the analysis.

3.2.5. Test of Hypothesis

The statistical significance of the independent variable's impact on the dependent variable was tested by the p -value quantified at the specified significance level; the smaller the p -value, the more significant the regression coefficient is. In general, if the p -value of the independent variable regression coefficient is less than the significance level, the variable is considered statistically significant in its effect on the dependent variable (Jarkas, 2015). The goodness of fit, i.e., how well the linear regression model relates to the data, and the overall adequacy of the model were assessed by the coefficient of determination (R^2) and the F-ratio, respectively. The closer the value of R^2 to unity, the better the goodness of fit, whereas an F-ratio greater than 1 indicates that at least one regression coefficient in the model is significantly different from zero, and hence the overall regression model is useful for predicting the dependent variable.

3.2.6. Model Validation

The models developed under this study were checked for their aptness by using regression diagnostics. Regression analysis is made under certain assumptions that must not be violated if the model is to be used for prediction. These assumptions were tested by assessing the appropriate graphs, both before and after model development.

Furthermore, the models were validated by testing the predictive ability of the developed models. To compare the results of the reference data of all models, two standard error measures, namely root mean square error (RMSE) and mean absolute percentage error (MAPE) (Bayram, 2016).

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (Y_{real,i} - Y_{model,i})^2}{n}} \quad (\text{Eq.3-2})$$

$$MAPE (\%) = \frac{1}{n} \sum_{i=1}^n \left(\frac{Y_{real,i} - Y_{model,i}}{Y_{real,i}} \right) * 100 \quad (\text{Eq.3-3})$$

Where, $y_{real,i}$ and $y_{model,i}$ represent the observed values and the model predictions respectively, and n represents the number of data.

4. Results and Discussion

4.1. Data Characteristics

4.1.1. Data Collection

Data pertaining to past project information was acquired from relevant government offices in the city. The population of the study, public buildings in the Addis Ababa, comprise buildings built for the city administration on one hand and buildings built for the federal government on the other hand, overseen by the Addis Ababa City Administration Construction Bureau (AACACB) and Federal Government Buildings Construction Project Office (FGBCPO), respectively. The lack of database on past project records makes it hard to acquire data on all of the projects overseen under these bureaus in the past. Accordingly, this study was able to acquire data on 34 public building projects completed between 2015 and 2019, after reviewing contract documents, handing over certificates, payment certificates and drawings. 27 of the projects were provided from AACACB, while 7 projects were availed from FGBCPO. The description of the variables comprised in the data set is shown in Table 4-1.

Table 4-1 Description of Variables

No	Variable	Code	Unit	Type	Scale
1	Actual Project Duration	D_a	Days	Dependent	Ratio
2	Contract Duration	D_c	Days	Independent	Ratio
3	Actual Project Cost	C_a	ETB	Independent	Ratio
4	Contract Cost	C_c	ETB	Independent	Ratio
5	Average Floor Area	A	m^2	Independent	Ratio
6	Gross Floor Area	A_g	m^2	Independent	Ratio
7	Number of Floors Below Ground Floor	NoF_{bg}	No	Independent	Ratio
8	Number of Floors Above Ground Floor	NoF_{ag}	No	Independent	Ratio
9	Total Number of Floors	NoF	No	Independent	Ratio
10	Average Floor Height	H	m	Independent	Ratio
11	Total Building Height	BH	m	Independent	Ratio
12	Building Volume	V	m^3	Independent	Ratio

The projects used for the study include administration offices for 9 of the 10 sub-cities of Addis Ababa, office buildings for hospitals, fire & rescue agencies, and government agencies. Therefore, the projects are of similar nature with respect to their types. As for their geographic location, they are fairly scattered throughout the city and in all the sub-cities.

Table 4-2 Description of Projects for the Study

No	Project Name	Dc (days)	Da (days)	Cc	Ca
				in M ETB	in M ETB
1	Bole Sub-city Office	558	1760	227.13	234.77
2	Kirkos Sub-city Office	730	1317	297.22	280.25
3	Gulele Sub-city Office	558	1035	202.76	197.62
4	Addis Ketema Sub-city Office	558	1245	228.16	237.36
5	Arada Sub-city Office	730	1192	290.08	220.51
6	Kolfe Sub-city Office	730	1223	250.48	231.92
7	Yeka Sub-city Office	558	1095	240.61	220.34
8	Akaki Kality Sub-city Office	558	1095	229.01	225.9
9	Yeka Sub-city Police Office	530	763	60.14	54.56
10	Addis Ketema Sub-city Police	530	560	63.97	53.16
11	Polytechnic (Tegbared) Building	365	1095	53.94	37.76
12	Akaki TVET Building	365	740	18.58	19.07
12	Entoto Technical College Building	180	410	11.08	8.13
13	Kotebe University College Building Project 1	180	977	8.99	9.4
14	Kotebe University College Building Project 2	180	702	8.2	8.27
16	Bole Sub-city Fire and Rescue Agency Building	365	365	28.46	23.87
17	Addis Ketema Sub-city Fire and Rescue Agency Building	180	768	10.96	11.19
18	Gulele Sub-city Fire and Rescue Agency Building	365	949	36.51	25.56
19	Minilik Hospital Building	540	1580	183.32	188.64
20	Yekatit 12 Hospital Office Building	540	1580	178.48	197.61
21	Gandi Hospital Building	365	450	20.39	23.23
22	Zewditu Hospital Office Building	365	480	8.19	11.56
23	Justice Bureau Building	540	803	66.64	61.42
24	Environment Protection Office Building	485	1760	43.51	46.75
25	Addis Ababa Construction Bureau Laboratory Building	180	730	13.48	14.12
26	Education Bureau Building	540	1085	60.34	56.47
27	Ethiopian Sport Commission Office Building	540	1420	45.16	51.63

28	Urban Integrated Land Information System Development Office Building	540	1,273	166.26	160.42
29	Ministry of Science and Technology Office Building	540	1,385	30.56	35.23
30	Industry Minister Office Building	540	1,547	30.48	33.76
31	Central Statistics Agency Office Building	555	1,273	23.65	27.35
32	Ethiopian Roads Authority Office Building	590	1,742	39.66	33.88
33	Ministry of Foreign Affairs Conference Hall Project	426	1,505	162.61	190.27
34	Mofa Car Parking Project	887	1,698	80.66	123.71

All of the above projects were carried out with the traditional Design-Bid-Build (DBB) form of contract delivery and hence the partitioning of data with this respect or the association of this variable in the analysis was disregarded. 97% of the projects were not completed within the contract duration while the rest of the projects were completed within their time frame. Whereas 50% of the projects faced cost overrun. This reaffirms the work of Anshebo 2017) where 100% of the public buildings under his study were not finished in the allotted time.

When reported, projects which have periods during which the construction progress had come to a complete halt for a considerable time, were removed from the data set. Such a measure is required to unravel the effects of the project scope factors on the effective construction duration, as the impact of interruption can mask or overshadow their influence, which may undermine the validity of the results obtained.

4.1.2. Data Processing

Most studies have outlined the importance of partitioning the data to reflect difference in context of the projects such as type of project, type of project delivery system, location, and so on. However, the projects under this study were fairly similar in most aspects and the partitioning of the gathered data was overlooked.

The projects were completed between 2010 and 2019 with construction mid-point dates ranging from 2012 to 2018. In this regard, it's imperative that the cost elements be adjusted for the effect of inflation. The final cost of the projects is therefore adjusted for inflation. A base year of 2018 was used and the rest of the cost data were adjusted for inflation based on the Consumer Price Index (CPI) provided by the Ethiopian Central Statistics Agency.

4.1.3. Data Correlation

The first step in developing regression model is analyzing the relationship among variables. This is done for three reasons – (1) to check how the independent variables are related with the dependent variable so as to show which variables can best predict the dependent variable, (2) to see the pattern of the relationship between each independent variable and the dependent variable and, (3) to check how the independent variables are correlated with each other i.e. the problem of multicollinearity. In this regard, Pearson’s correlation coefficient as well as graphical exploration methods such as scatter plots are widely used. Correlation coefficient is a measure of the degree of the relationship between two variables while the scatter plots graphically show the pattern of the relationship, most importantly whether the relationship is linear or not.

The correlation coefficient, r , is between -1.0 and $+1.0$, and a negative and positive correlation exists if $-1.0 < r < 0$ and $0 < r < 1.0$, respectively and the more the value is close to 1 the more related the variables are (Yeom, et al., 2018). The correlation matrix among the variables under this study shown in Table 4-2.

Table 4-3 Correlation Matrix for Unadjusted Data

	D_a	C_a	A	H	NoF_{ag}	NoF_{bg}	A_g	NoF	BH	V
D_a	1.000	0.471	0.513	0.154	0.304	0.453	0.429	0.347	0.344	0.429
C_a	0.471	1.000	0.762	0.539	0.581	0.638	0.912	0.628	0.669	0.925
A	0.513	0.762	1.000	0.457	0.176	0.524	0.695	0.238	0.276	0.728
H	0.154	0.539	0.457	1.000	0.236	0.352	0.425	0.269	0.373	0.481
NoF_{ag}	0.304	0.581	0.176	0.236	1.000	0.480	0.770	0.992	0.981	0.740
NoF_{bg}	0.453	0.638	0.524	0.352	0.480	1.000	0.591	0.589	0.606	0.591
A_g	0.429	0.912	0.695	0.425	0.770	0.591	1.000	0.796	0.817	0.997
Nof	0.347	0.628	0.238	0.269	0.992	0.589	0.796	1.000	0.993	0.768
BH	0.344	0.669	0.276	0.373	0.981	0.606	0.817	0.993	1.000	0.797
V	0.429	0.925	0.728	0.481	0.740	0.591	0.997	0.768	0.797	1.000

Evidently, average floor area had the highest coefficient of correlation with actual project duration, while actual project cost, gross floor area, number of floors below ground and building volume had considerable relationship with project duration.

The corresponding scatter plot matrix among the variables is shown in Figure 4-1. The pattern of the relationship is between the predictor variables and the response variable is different across the independent variables. The variables that have good correlation with duration seem to have a fairly linear relationship while some of the independent variables are linearly dependent with each other that could pose the problem of multicollinearity.



Figure 4-1 Scatter Plot Matrix for Unadjusted Data

The scatter plots among the variables have also shown that there are some outliers in the data that would skew the model. The removal of outliers from the data set is controversial. Some have

opined that the removal of outliers in regression analysis help linearize the relationship and result in a best-fit model while others argue the removal of outliers from the data set violates the purpose of the model and choose to use the outliers to give the full picture as it is presented to the researcher.

4.1.4. Data Transformation

Data do not always come in a form that is immediately suitable for analysis. Data transformation is often carried out on the variables before carrying out the analysis. Transformations are applied to accomplish certain objectives such as to ensure linearity, to achieve normality, or to stabilize the variance. It often becomes necessary to fit a linear regression model to the transformed rather than the original variables (Chatterjee & Hadi, 2006). In this regard, logarithmic transformations had proven to be productive in previous studies, as was in the case of the BTC. Accordingly, the data set was transformed into the natural logarithmic form. The correlation of the variables considerably increased from the unadjusted case as can be seen in Table 4-3.

Table 4-4 Correlation Matrix for Transformed Data

	ln(Da)	ln(Ca)	ln(A)	ln(H)	ln(NoF_{ag})	ln(A_g)	ln(NoF)	ln(BH)	ln(V)
ln(Da)	1.000	0.600	0.597	0.148	0.272	0.628	0.425	0.427	0.622
ln(Ca)	0.600	1.000	0.818	0.512	0.464	0.884	0.620	0.658	0.897
ln(A)	0.597	0.818	1.000	0.412	0.142	0.801	0.312	0.351	0.819
ln(H)	0.148	0.512	0.412	1.000	0.257	0.423	0.276	0.377	0.480
ln(NoF_{ag})	0.272	0.464	0.142	0.257	1.000	0.684	0.948	0.940	0.662
ln(A_g)	0.628	0.884	0.801	0.423	0.684	1.000	0.819	0.839	0.998
ln(NoF)	0.425	0.620	0.312	0.276	0.948	0.819	1.000	0.994	0.798
ln(BH)	0.427	0.658	0.351	0.377	0.940	0.839	0.994	1.000	0.825
ln(V)	0.622	0.897	0.819	0.480	0.662	0.998	0.798	0.825	1.000

The transformation of the data into natural logarithm form has resulted in a considerable increase in the coefficient of correlation. Gross floor area was found to be the independent variable with the highest correlation coefficient of 0.628 when compared with project cost. The corresponding scatter plot matrix for the transformed data is shown in figure 4-2. It is worth mentioning that the transformation into natural logarithmic form for the data corresponding to number of floors below

the ground has not been made in this study, as some of the buildings have no basement floors and this made the log-transformation for this data set inapplicable.

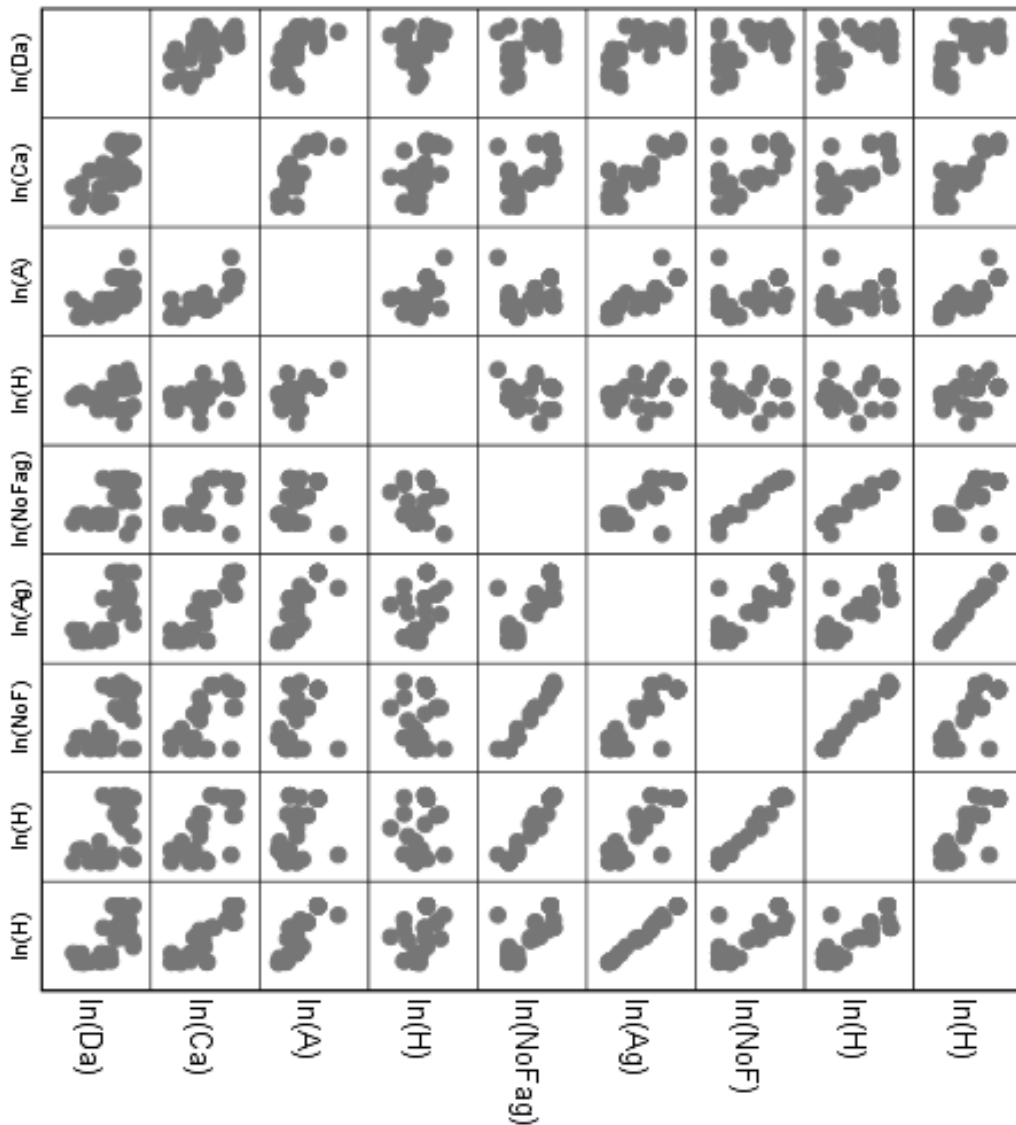


Figure 4-2 Scatter Plot Matrix for Transformed Data

The transformation of the data set into logarithmic form not only increased the value of the coefficient of correlation, but also enhanced the linearity of the relationship between each independent variable and the dependent variable. This is illustrated in the first column of the scatter plot matrix shown in Figure 4-2. The linear relationship shown the first column implies that the model to be developed based on these independent variables shall be linear.

4.2. BTC Model

The BTC model, based on the power function and presented in Eq 2.2., is transformed into a linear form by the natural logarithm form as shown below.

$$\ln(T) = \ln(k) + B\ln(C) \tag{Eq.4.1}$$

Where

$\ln(T)$ is the natural logarithm of project duration;

$\ln(C)$ is the natural logarithm of project cost

$\ln(K)$ is the natural logarithm of K

Equation 4.1 can be transformed to the standard simple linear regression form expressed as:

$$Y = a + bX \tag{Eq.4.2}$$

Where

$Y = \ln(T)$

$a = \ln(K)$

$X = \ln(C)$

In order to test the reliability of the developed model to predict project duration realistically, 30 of the projects were randomly selected to be used for model development while 4 projects were used for model validation. Accordingly, the data transformed into natural logarithmic form corresponding to duration and cost was then fixed to the above simple linear regression form. The model was developed by using SPSS version 25 and the output is shown in Table 4-4 to Table 4-9. For comparison purposes the contract data was incorporated in the model development.

Table 4-5 Model Summary - BTC Actual Data

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.630 ^a	.397	.375	.36588

a. Predictors: (Constant), $\ln(C)$

Table 4-6 ANOVA – BTC Actual Data

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	2.464	1	2.464	18.403	.00019 ^b
	Residual	3.748	28	.134		
	Total	6.212	29			

a. Dependent Variable: ln(Da)

b. Predictors: (Constant), ln(Ca)

Table 4-7 Model Coefficients – BTC Actual Data

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	2.350	1.071		2.194	.037
	ln(Ca)	.252	.059	.630	4.290	.00019

a. Dependent Variable: ln(Da)

Table 4-8 Model Summary - BTC Contract Data

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.744 ^a	.554	.538	.29212

a. Predictors: (Constant), ln(Cc)

Table 4-9 ANOVA - BTC Contract Data

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	2.969	1	2.969	34.787	.000002 ^b
	Residual	2.389	28	.085		
	Total	5.358	29			

a. Dependent Variable: ln(Dc)

b. Predictors: (Constant), (ln(Cc))

Table 4-10 Model Coefficients - BTC Contract Data

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	1.217	.831		1.465	.154
	ln(Cc)	.274	.046	.744	5.898	.000002

a. Dependent Variable: ln(Dc)

The above sets of results were selected on account of their relevance to describe the developed model. The first table shows the value of R and R² which show the goodness of fit, while the second and third result show the significance value and the regression coefficients respectively.

4.2.1. Test of Hypothesis 1

Regression models are derived under the assumption that the dependent variable is affected linearly by at least one independent variable. In order for the derived regression model to be meaningful, a null hypothesis, such as “None of the independent variable affects the dependent variable” has to be set, and the hypothesis test should return a statistically significant result (Yeom, et al., 2018). To achieve this objective, this study verified the suitability of the developed model, with a significance level of 0.05 or lower, based on the significance probability (p-value) and the corresponding F value in the ANOVA table derived from SPSS version 25.

The first hypothesis of the study is “Verify the applicability of Bromilow’s Time-Cost model for public building projects in Addis Ababa.” The corresponding null hypothesis that is subjected to acceptance or rejection is:

$$H_{01} = \text{“There is no significant relationship between project duration and project cost.”}$$

The overall regression model, based on the BTC, was significant with $F(1,28) = 18.403$, $\rho = 0.00019 \ll 0.05$ & Adjusted $R^2 = 0.397$ for the actual project data and $F(1,28) = 34.787$, $\rho = 0.000002 \ll 0.05$ & Adjusted $R^2 = 0.554$. Therefore, the null hypothesis is rejected based on both actual and contract cost and duration data.

The resulting regression equation based on the model coefficients shown in Table 4-6 and Equation 4.2 would be:

$$Y = 2.35 + 0.252X \tag{Eq.4.3}$$

K is calculated based on Eq. 4.2 as

$$K = e^{2.35} = 2.718^{2.35} = 10.48$$

Thus, the non-linear form of the BTC, as shown in Equation 2.1, is expressed in Eq. 4.4.

$$T = 10.48C^{0.252} \tag{Eq.4.4}$$

Where,

T – Actual Project Duration in days; and

C – Actual Project Cost in ETB.

The corresponding equation based on contract cost and duration is given in Equation 4.5.

$$T = 3.38C^{0.274} \quad (\text{Eq.4.5})$$

Where,

T – Contract Project Duration in days; and

C – Contract Project Cost in ETB.

4.3. Multiple Linear Regression Model

Chatterjee and Hadi's work (2006) reviewed in chapter 2 stipulates that a best-fit regression model can be developed by the steps shown below.

1. Statement of the problem
2. Selection of potentially relevant variables
3. Data collection
4. Model specification
5. Choice of fitting method
6. Model fitting
7. Model validation and criticism
8. Using the chosen model(s) for the solution of the posed problem.

At this stage, the study has already gone through the first 3 steps. The problem is defined in the first chapter, the relevant factors that have an impact on construction duration had been selected from the literature review and based on this data has been collected on the variables. Henceforth, subsequent steps of the model development would be dealt with hereafter.

4.3.1. Model Specification

In specifying a regression model, one must determine whether the form of the equation is linear or nonlinear. The typical equation of multiple linear regression is shown in Eq. 4.

$$Y = \beta_0 + \beta_1X_1 + \beta_2X_2 + \dots + \beta_nX_n + \varepsilon \quad (\text{Eq.4.6})$$

where $\beta_0, \beta_1, \beta_2, \dots, \beta_n$ are constants referred to as the regression coefficients and ε is error.

In the case of this study the model is linear i.e. the regression parameters enter the equation linearly. This is checked by observing the scatter plots between each independent variable and the dependent variable. As mentioned above, the data set was transformed into the natural logarithmic form to linearize the relationship and this is demonstrated in Figure 4-2.

Moreover, with respect to the number of dependent and independent variables, the model to be developed for this study would be a univariate multiple linear regression as there is only one dependent variable and there are possibly multiple independent variables.

4.3.2. Choice of Fitting Method

After the model has been defined and the data have been collected, the next task is to estimate the parameters of the model based on the collected data. The most commonly used method of estimation is called the least squares method (Chatterjee & Hadi, 2006). In some cases, especially when there are a large number of explanatory variables, we might use statistical criteria to include/exclude explanatory variables, especially if we are interested in the ‘best’ equation to predict the dependent variable. This is mainly done by using statistical software packages such as SPSS.

There are 5 options of model fitting in SPSS; Enter, Stepwise, Remove, Backward and Forward. Of these options, stepwise method was used in this study, whereby SPSS automatically develops the model by starting first with the independent variable with the most correlation coefficient with the dependent variable and progressively adding more variables to reach to the best fit model.

4.3.3. Model Fitting

The transformation of the data set into natural logarithmic form had resulted in a better correlation coefficient between each independent variable and the dependent variable, as shown in Table 4-3. Accordingly, the transformed data set was analyzed by SPSS and the output of the developed model is presented through Table 4-10 to 4-14. It is worthwhile to point out that number of floors below ground was omitted from the data set as its transformation into logarithmic form could not be achieved for some buildings which don't have basement floors.

Table 4-11 Variables Entered/Removed^a – Developed Model

Model	Variables Entered	Variables Removed	Method
1	ln(Ag)		Stepwise (Criteria: Probability-of-F-to-enter <= .050, Probability-of-F-to-remove >= .100).

a. Dependent Variable: ln(Da)

Table 4-12 Model Summary^b – Developed Model

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.668 ^a	.446	.425	.34942

a. Predictors: (Constant), ln(Ag)

b. Dependent Variable: ln(Da)

Table 4-13 ANOVA^a – Developed Model

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	2.654	1	2.654	21.734	.000076 ^b
	Residual	3.297	27	.122		
	Total	5.950	28			

a. Dependent Variable: ln(Da)

b. Predictors: (Constant), ln(Ag)

Table 4-14 Coefficients^a – Developed Model

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	3.682	.697		5.280	.000014
	ln(Ag)	.367	.079	.668	4.662	.000076

a. Dependent Variable: ln(Da)

Table 4-15 Excluded Variables^a – Developed Model

Model		Beta In	t	Sig.	Partial Correlation	Collinearity Statistics Tolerance
1	ln(Ca)	.105 ^b	.304	.764	.059	.178
	ln(Dc)	.270 ^b	1.310	.202	.249	.471
	ln(Cc)	-.068 ^b	-.203	.841	-.040	.190
	ln(A)	.041 ^b	.153	.879	.030	.303
	ln(H)	-.023 ^b	-.150	.882	-.029	.865
	ln(NoFag)	-.064 ^b	-.265	.793	-.052	.363
	ln(NoF)	-.035 ^b	-.129	.898	-.025	.296
	ln(BH)	-.075 ^b	-.257	.799	-.050	.252
	ln(V)	-.445 ^b	-.219	.828	-.043	.005

a. Dependent Variable: ln(Da)

b. Predictors in the Model: (Constant), ln(Ag)

The resulting model has only one predictor variable which is gross floor area as the inclusion of other variables in the model didn't improve the significance of the model.

The overall regression model was significant with $F(1,28) = 21.734$, $\rho = 0.000076 \ll 0.05$ & $R^2 = 0.425$ for the fitted data set.

The resulting regression equation based on the model coefficients shown in Table 4-13 and the typical linear regression equation expressed in Eq. 4.2 would be:

$$\ln(Da) = 3.682 + 0.367\ln(Ag) \quad (\text{Eq.4.7})$$

Where

ln(D_a) or Natural logarithm of project duration in days; and

ln(A_g) or Natural logarithm of gross floor area of the building in m².

The non-linear form of the developed model, based on Eq. 4.2., is expressed in Eq. 4.8.

$$T = 39.71Ag^{0.367} \quad (\text{Eq.4.8})$$

4.3.4. Test of Hypothesis 2

The second hypothesis of the study is “Verify whether incorporating influential factors other than cost could result in an enhanced best-fit model that can predict duration.” The corresponding null hypothesis is:

H_{02} = “Including influential project parameters other than cost doesn’t improve the statistical significance.”

The above hypothesis can be tested by comparing the statistical significance of the BTC model with the developed regression model as shown in Table 4-15. The developed model based on gross floor area was found to have a better statistical significance as compared with the BTC and hence the second hypothesis of the study was rejected.

Table 4-16 Comparison of Models

Criteria	BTC	Developed Model
Predictor	Cost (C_a)	Gross floor area (A_g)
Formula	$T = 10.48C^{0.252}$	$T = 39.71Ag^{0.367}$
Adjusted R^2	0.397	0.425
F	18.403	21.734
ρ -value	0.00019	0.000076

Although, the developed model based on gross floor area has a better statistical significance than the model developed based on BTC, it will have an inherent flaw on account of the gross floor area being the only variable. This implies that the model would predict the same amount of duration for two projects having the same gross floor area inspite of the number of floors which could have a practical impact on duration as a building having multiple floors would take longer as compared to a building built on the ground.

4.4. Regression Diagnostics and Model Validation

The seventh step in the development of regression models is to check the validity of the developed model. Regression models are developed under certain assumptions that must be met. The

distribution theory, confidence intervals, and tests of hypotheses are valid and have meaning only if the standard regression assumptions are satisfied (Chatterjee & Hadi, 2006). The assumptions of the regression analysis are described hereunder.

4.4.1. Assumptions of Regression Analysis

The basic assumptions of regression analysis, as outlined by Chaterjee and Hadi (2006) are:

1. **Assumptions about the form of the model:** the model that relates the dependent and independent variables is assumed to be linear in the regression parameters. The scatter plots between the predictor and response variable in the developed model was linear and this assumption is met. The transformation of the data set into natural logarithmic form helped linearize the relationship
2. **Assumptions about the errors:** the errors in the developed model are assumed to be independently and identically distributed normal random variables each with mean zero and a common variance σ^2 . This entails four assumptions
 - The errors have a normal distribution (the normality assumption);
 - The errors have mean zero;
 - The errors have the same but unknown variance. This is called the homoscedasticity assumption; the violation of this assumption is called heteroscedasticity and
 - The errors are independent to each other.

The above assumptions about the errors were checked by analyzing appropriate residual plots. Three residual plots were used to check the above assumptions. The first one is the residual scatter plot of the regression standardized predicted value versus the regression standardized residual. Figures 4-3 and 4-4 show the residual scatter plots of the BTC model and the developed model respectively. The pattern of the scatter plot showed randomness for both cases and hence the assumption that the errors have a normal distribution was met. Furthermore, the lack of pattern in the plot implies the variance of the errors is somehow the same and hence there is no problem of heteroscedasticity.

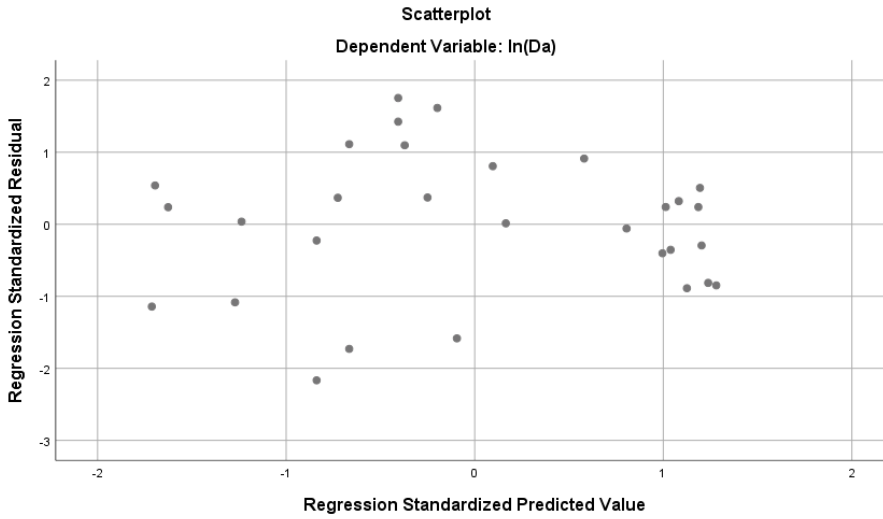


Figure 4-3 Residual Scatter Plot - BTC

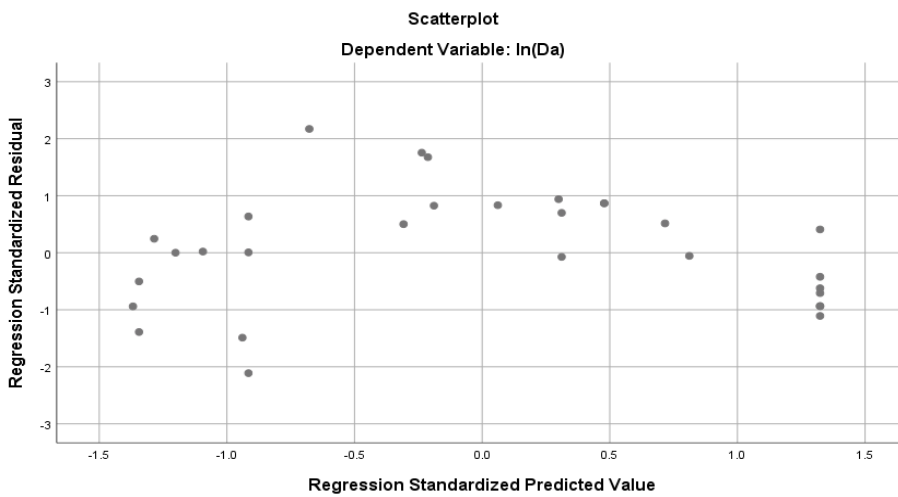


Figure 4-4 Residual Scatter Plot - Developed Model

The second residual plot used to check linearity and normality was the normal P-P plot of regression standardized residuals. The data points were found to be hovering around the straight line and this further affirms the assertion that the errors are normally distributed.

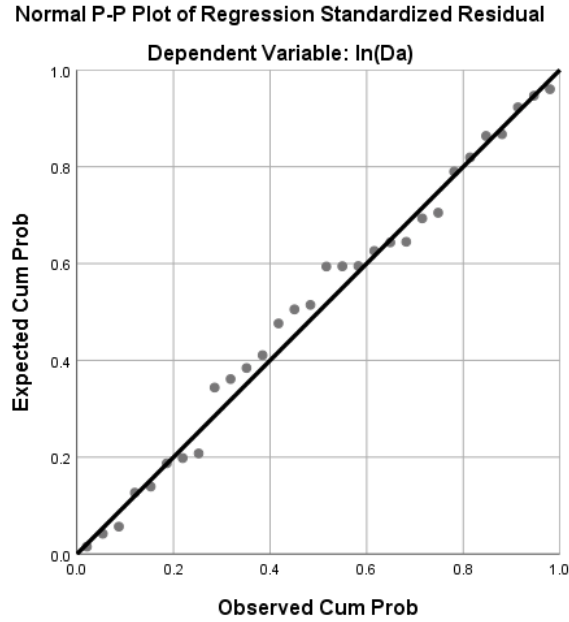


Figure 4-5 Normal P-P Plot of Regression Standard Residual - BTC

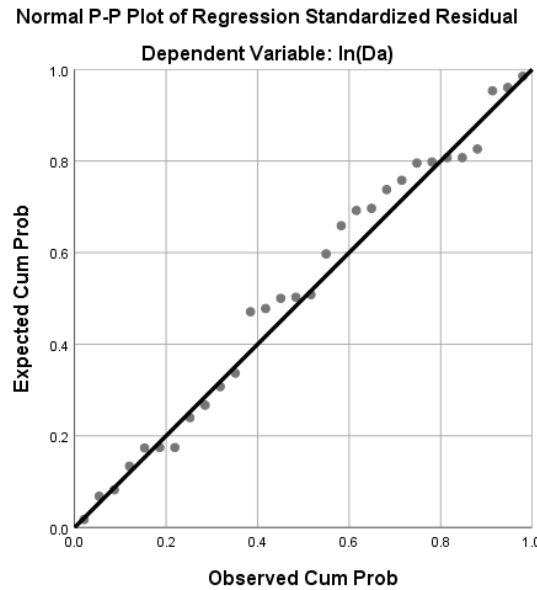


Figure 4-6 Normal P-P Plot of Standardized Residual - Developed Model

The third plot used to check the normality of the errors is the residual histogram as shown in figures 4-7 and 4-8 for the BTC and the developed models respectively. The resulting plot for both cases showed a pattern of normality and hence the normality assumption holds. Furthermore, the mean

of the errors in both the developed model and the BTC were very close to zero, while the standard deviation was close to 1.

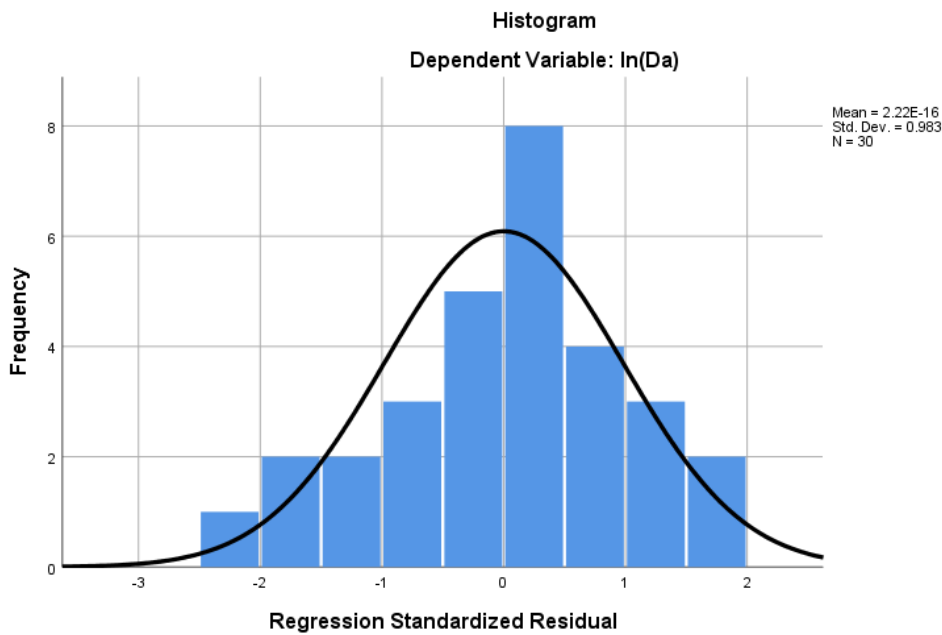


Figure 4-7 Residual Histogram - BTC

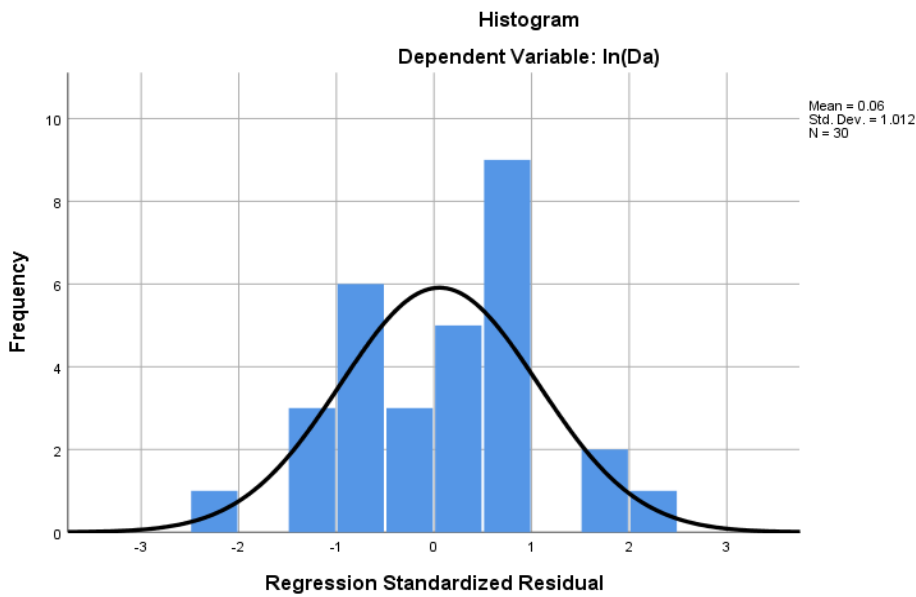


Figure 4-8 Residual Histogram - Developed Model

Accordingly, the assumptions about the errors all hold true and the

3. **Assumptions about the Predictors:** there are three assumptions about predictors

- The predictor variables are nonrandom.
- The values of the predictor variables are measured without error.
- The predictor variables are assumed to be linearly independent of each other. If this assumption is violated, the problem is referred to as the multicollinearity problem.

The first assumption about the predictors only holds true when the experimenter can set the values of the predictor variables at predetermined levels. It is clear that under non-experimental or observational situations this assumption will not be satisfied (Chatterjee & Hadi, 2006). The measurement error associated with the values of the predictor variables was overlooked in this study as values were measured with care. Evidently, the problem of multicollinearity is not assessed in the case of this study as there is only one predictor variable in the model.

4. **Assumptions about the observations:** All observations are equally reliable and have approximately equal role in determining the regression results and in influencing conclusions. In this regard, the project parameters included in the data set were believed to have the same influence and this assumption was not violated.

4.4.2. Model Validation

Further to the above regression diagnostics carried out to validate the model, this study tried to show the predictive ability of the models. In this regard, two standard error measures, namely Root Mean Square Error (RMSE) and Mean Absolute Percentage Error (MAPE), were calculated using Eqs. (3-2) and (3-3) respectively. The closer the error values are to zero, the more accurate the model forecasts are. As described above, out of the 34 project whose data had been collected under this study, 30 of them were used for the model development and 4 were reserved for the validation of the model. The actual duration of the projects and the calculated values based on the developed models and the model developed by BaTCoDA is shown in Table 4-17. The model developed by BaTCoDA is used to make inference with models developed under this study.

Table 4-17 Actual and Calculated Duration of Projects used for calibration.

S.N.	Project Name	Actual Duration	BTC	Developed	BaTCoDA
1	Addis Ketema Sub-city Office Building	1245	1,490.34	1,522.43	1,901.73
2	Yeka Sub-city Police Administration Office Building	763	1,005.45	671.27	1,129.94
3	Kotebe University College Building Project 2	702	639.55	701.36	621.09
4	Justice Bureau Building	803	1,100.61	1,116.34	1,273.52

Table 4-18 Predictive ability of models

Measure	BTC	Developed Model	BaTCoDA
Formula	$T = 10.48C^{0.252}$	$T = 39.71Ag^{0.367}$	$T = 7 \text{ to } 12^* C^{1/3}$
RMSE (days)	229.92	214.22	445.50
MAPE (%)	24.36	18.35	33.94

* an average value of 9 is taken for office buildings (assumption)

Although the standard error measure values are considerably high for all the above cases, the model developed based on gross floor area is found to be better in predicting duration. The model developed by BaTCoDA had a considerably high values of RMSE and MAPE which suggests that the lack of update on the model over the years had made it not applicable to predict duration.

5. Conclusion and Recommendation

5.1. Conclusion

The principal objective of this study was to develop a model that could be used to forecast the duration of public building projects in Addis Ababa, based on the information available to the estimator at pre-construction phase, in a reliable and practical way. In this regard, the study had put forth two hypotheses to realize its objectives.

- I. Verify the applicability of Bromilow's Time-Cost model for public building projects in Addis Ababa. and
- II. Verify whether incorporating influential factors other than cost could result in an enhanced best-fit model that can predict duration.

In this regard, data on 34 completed public buildings were used to test the hypotheses and the following conclusions are drawn from the study.

- Bromilow's 50-year old time-cost model (BTC) was found to be valid for the data set under the study. The corresponding null hypothesis was rejected and the formula based on the BTC was generated and is expressed as $T = 10.48C^{0.252}$.
- Construction cost, gross floor area, floor height, number of floors above and below ground were the parameters abundantly used to predict construction duration, based on the reviewed literature.
- Transformation of the data set into logarithmic form has resulted in an enhanced correlation between the dependent variable and the independent variables. Furthermore, a more linear relationship was achieved through the transformation.
- The inclusion of parameters other than cost to predict construction duration had proved to be fruitful. A model was tried to be developed based on project cost, average floor area, average floor height, number of floors above ground, total number of floors, gross floor area, total building height and building volume. The resulting model however, based on the stepwise regression analysis made on SPSS, was based only on gross floor area and is expressed as $T = 39.71Ag^{0.367}$.

- An inherent flaw of the model developed based on gross floor area is the fact that two buildings having the same gross floor area but different floor numbers would have the same duration based on the model. However, practically speaking a building having multiple floors would take relatively longer to build as compared to a building built on the ground. Hence, estimators shall consider this while predicting duration based on this model.
- The developed model based on gross floor area were found to be more statistically significant than the BTC model. The p -value and F -value for the BTC was 0.00019 and 18.403 respectively, while that of the developed model was 0.000076 and 21.734 respectively. Furthermore, the adjusted R^2 for the BTC was 0.397 while that of the developed model was 0.425 which implies the developed model was a better fit for the data set. Hence, the corresponding null hypothesis was rejected and the inclusion of parameters other than cost did in fact improve the statistical significance.
- The developed models were checked whether they were in line with the basic regression assumptions and both models didn't violate the assumptions.
- The developed models were validated by using two standard error measures of RMSE and MAPE which were found to be 229.92 days and 24.36% for the BTC and 214.22 days and 18.35% for the developed model respectively. This shows that the developed model based on gross floor area is superior to the BTC in terms of predictive ability and shall be used for prediction of construction duration of public buildings in Addis Ababa. However, the values of the standard error measures are considerably higher even for the developed model based on gross floor area and hence the model shall be continuously updated with the inclusion of more data.

5.2. Recommendations

Based on the findings of this study the following recommendations are forwarded to the respective project stakeholders.

Project Owners

The project owners for the public building projects i.e. Addis Ababa City Administration Construction Bureau (AACACB) and Federal Government Buildings Construction Project Office (FGBCPO):

- Shall use the developed model based on gross floor area to practically predict project duration in planning and bidding stages.
- Shall continuously update the formula with more project records to refine the model and improve the predictive ability.
- Shall use the formula to calculate duration while doing feasibility studies.

Contractors

- The contractors engaged in the construction of the buildings shall use the formula to check the duration set by their clients are reasonable.
- Shall use the formula in their construction schedule preparations.
- Shall update the model continuously and also come up with their own formula for different types of projects.

Future research

- Shall develop similar models for other sectors and locations to develop a comprehensive model for different applications.
- Shall incorporate other duration determinant factors to refine the model.

References

- Alemayehu, S., 2014. *Testing Regression Models to Estimate Costs of Road Construction Projects*, Addis Ababa: Addis Ababa University.
- Ameyaw, C., Arthur, Y. D. & Mensah, S., 2012. *Applicability of Bromilow's Time-Cost Model on Building Projects in Ghana*. Kumasi, Ghana, ResearchGate.
- Anshebo, A. Y., 2017. *Performance Assessment of Public Building Construction Projects in Addis Ababa*, Addis Ababa: Addis Ababa University.
- Assefa, A., 2008. *Time-Cost Relationships for Public Road Construction Projects in Ethiopia*, Addis Ababa: Addis Ababa University.
- Bayram, S., 2016. Duration Prediction Models for Construction Projects: In Terms of Cost or Physical Characteristics?. *KSCE Journal of Civil Engineering*, pp. 1-12.
- Bhokha, S. & Ogunlana, S. O., 1999. Application of artificial neural network to forecast construction duration of buildings at the predesign stage. *Engineering, Construction and Architectural Management*, 6(2), pp. 133-144.
- Bromilow, F. J., 1969. *Contract time performance: Expectations and the reality*. Australia, s.n.
- Chan, D. W. & Kumaraswamy, M. M., 1999. Modelling and predicting construction durations in Hong Kong public housing. *Construction Management and Economics*, 17(3), pp. 351-362.
- Chatterjee, S. & Hadi, A. S., 2006. *Regression by Example*. Fourth ed. New Jersey: John Wiley and Sons Inc..
- Dawson, R. & Zeevaart, S., 2009. *Construction Project Management Handbook*. Washington DC: Gannett Fleming, Inc..
- Dursun, O. & Stoy, C., 2012. Determinants of construction duration for building projects in Germany. *Engineering, Construction and Architectural Management*, Vol. 19(Issue. 4), pp. 444-468.

- Fellows, R. & Liu, A., 2008. *Research Methods for Construction*. Third ed. United Kingdom: Blackwell Publishing Ltd.
- FIDIC, 1999. *Conditions of Contract for Construction*. First ed. s.l.:International Federation of Consulting Engineers.
- Froby, T., 2008. *Multiple Linear Regression and Subset Selection*, Dallas: Department of Economics, Southern Methodist University.
- Heagney, J., 2012. *Fundamentals of Project Management*. Fourth ed. New York: American Management Association.
- Hoffman, G. J., Thal Jr., A. E., Webb, T. S. & Weir, J. D., 2007. Estimating Performance Time for Construction Projects. *Journal of Management in Engineering*, Volume 23, pp. 193-199.
- Jarkas, A. M., 2015. Predicting Contract Duration for Building Construction: Is Bromilow's Time-Cost Model a Panacea?. *Journal of Management in Engineering*, pp. 1-8.
- Kaka, A. & Price, A. D., 1991. Relationships between value and duration of construction projects. *Construction Management and Economics*, Volume 9, pp. 383-400.
- Karanci, H., 2010. *A Comparative Study of Regression Analysis, Neural Networks and Case-based Reasoning for Early Range Cost Estimation of Mass Housing Projects*, Turkey: Middle East Technical University.
- Kothari, C. R., 2004. *Research Methodology, Methods and Techniques*. Second ed. Jaipur: New Age International Publishers.
- Kumaraswamy, M. M. & Chan, D. W., 1995. Determinants of construction duration. *Construction Management and Economics*, Vol. 13(Iss. 3), pp. 209-217.
- Le-Hoai, L., Lee, Y. D. & Nguyen, A. T., 2012. Estimating Time Performance for Building Construction Projects in Vietnam. *KSCCE Journal of Civil Engineering*, 17(1), pp. 1-8.
- Lin, M.-C., Tserng, H. P., Ho, S.-P. & Young, D.-L., 2011. Developing a Construction-Duration Model Based on a Historical Dataset for Building Project. *Journal of Civil Engineering and Management*, 17(4), pp. 529-539.

- Love, P. E., Tse, R. Y. C. & Edwards, D. J., 2005. Time–Cost Relationships in Australian Building Construction Projects. *Journal of Construction Engineering And Management*, Volume 131, pp. 187-194.
- Mackova, D. et al., 2017. Construction-Duration Prediction Model for Residential Buildings in Slovak Republic Based on Computer Simulation. *International Journal of Applied Engineering Research*, Vol. 12(ISSN 0973-4562), pp. 3590-3599.
- MARTIN, J., BURROWS, T. K. & PEGG, I., 2006. *Predicting Construction Duration of Building Projects*. Munich, Germany, XXIII FIG Congress.
- Mensah, I., Adjei-Kumi, T., Nani, G. & Adinyira, E., 2017. *Modelling Construction Duration: A comprehensive review of literature*. Ghana, Researchgate.
- Mishra, S. B. & Alok, S., 2017. *Handbook of Research Methodology*. India: Educreation Publishing.
- Mizanur, R. M., Dai, L. Y., Khanh, H. D. & Hyun, C. Y., 2014. Development of Time-Cost Models for Building Construction Projects in Bangladesh. *KICEM Journal of Construction Engineering and Project Management*, 4(3), pp. 13-20.
- Moses, T., 2014. *Using Bromilow's Model and Other Regression Model to Predict the Duration of Feeder Road Projects in Ghana*, MSc. Thesis, Acra: Kuame Nkrumah University of Science and Technology.
- Nani, G., Mensah, I. & Adjei-Kumi, T., 2017. Duration Estimation Model For Bridge Construction Projects in Ghana. *J. Eng. Des. Technol*, 15(6), p. 754–777.
- Ng, S. T., Mak, M., Skitmore , R. M. & Varnam, M., 2001. The Predictive Ability of Bromilow's Time-Cost Model. *Construction Management and Economics*, 19(2), pp. 165-173.
- Nkado, R. N., 1995. Construction time-influencing factors: thecontractor's perspective. *Construction Management and Economics*, , 13(1), pp. 81-89.
- ODABAŞI, E., 2009. *Models for Estimating Construction Duration: An Application for Selected Buildings on the METU Campus*, Ankara: Middle East Technical University.

Oyedele, O. A., 2017. *A Study into the Factors Affecting Duration of Construction Projects in Nigeria*, Nigeria: ND Building Tech..

Petruseva, S., Pusic, D. C. & Pancovska, V. Z., 2016. *Model for Predicting Construction Time by Using General Regression Neural Network*. Luhacovice, Czech Republic, International Scientific Conference People, Buildings and Environment 2016 (PBE2016).

PMI, 2016. *Construction Extension to the PMBOK Guide*. Pennsylvania: Project Management Institute, Inc..

PMI, 2017. *or A Guide to the Project Management Body of Knowledge*. Sixth ed. Pennsylvania: Project Management Institute, Inc..

RICS, 2015. *Code of measuring practice*. 6th ed. London: Royal Institution of Chartered Surveyors (RICS).

Shehu, W. B. & Nurudeen, Y., 2014. *Arid Zone Journal of Engineering, Technology and Environment*., Vol. 10(2545-5818), pp. 1-11.

Skitmore, M. & Thomas, N. S., 2003. Forecast models for actual construction time and cost. *Building and Environment*, Vol. 38(Iss. 8), pp. 1075-1083.

Son, J., Khwaja, N. & Milligan, D. S., 2018. Planning-Phase Estimation of Construction Time for a Large Portfolio of Highway Projects. *Journal of Construction Engineering and Management*, Volume 145(4), pp. 1-11.

Sriana, T. & Hayati, K., 2015. Time-Cost Relationship Model on the Construction of Education Building in Aceh Province. *Journal of Asian Scientific Research*, 5(7), pp. 328-339.

Walliman, N., 2011. *Research Methods: The Basics*. First ed. New York: Routledge.

Walpole, R. E., Myers, R. H., Myers, S. L. & Ye, K., 2011. *Probability and Statistics for Engineers and Scientists*. 9th ed. Boston: Pearson Education, Inc..

Waziri, B. S., Bala, K. & Bustani, S. A., 2017. Artificial Neural Networks in Construction Engineering and Management. *International Journal of Architecture, Engineering and Construction*, 6(1), pp. 50-60.

Waziri, B. S. & Yusuf, N., 2014. Application of Bromilow's Time-Cost Model to Highway Projects in Nigeria. *Arid Zone Journal of Engineering, Technology and Environment*, Volume 10, pp. 1-11.

Yeom, D.-J. et al., 2018. Development of an Approximate Model During the Project Planning Phase for General Office Buildings. *Journal of Civil Engineering and Management*, 24(3), pp. 238-253.

Zegeye, A. et al., 2009. *Introduction to Research Methods*, Addis Ababa: Addis Ababa University.

Appendices

A: Data Set of the Study

Table A-0-1 Study Data Set

No	Project Name	D _c	D _a	C _c in METB	C _a in METB	A	FH	NoF _{ag}	NoF _{bg}	GFA	NoF	BH	V
1	Bole Sub-city	558	1760	227.13	234.77	1,589.1	3.4	12	1	20,658	13	45.70	72,621
2	Kirkos Sub-city	730	1317	297.22	280.25	1,589.1	3.4	12	1	20,658	13	45.70	72,621
3	Gulele Sub-city	558	1035	202.76	197.62	1,589.1	3.4	12	1	20,658	13	45.70	72,621
4	Addis Ketema Sub-city	558	1245	228.16	237.36	1,589.1	3.4	12	1	20,658	13	45.70	72,621
5	Arada Sub-city	730	1192	290.08	220.51	1,589.1	3.4	12	1	20,658	13	45.70	72,621
6	Kolfe Sub-city	730	1223	250.48	231.92	1,589.1	3.4	12	1	20,658	13	45.70	72,621
7	Yeka Sub-city	558	1095	240.61	220.34	1,589.1	3.4	12	1	20,658	13	45.70	72,621
8	Akaki Kality Sub-city	558	1095	229.01	225.90	1,589.1	3.4	12	1	20,658	13	45.70	72,621
9	Yeka Sub-city Police	530	763	60.14	54.56	554.6	3.2	4	0	2,218	4	14.30	7,931
10	Addis Ketema Sub-city Police	530	560	63.97	53.16	554.6	3.2	4	0	2,218	4	14.30	7,931
11	Polytechnic (Tegbared) Building	365	1095	53.94	37.76	583.2	3.6	8	1	5,249	9	34.14	19,909

12	Akaki TVET Building	365	740	18.58	19.07	583.2	3.2	4	0	2,333	4	14.30	8,339
12	Entoto Technical College Building	180	410	11.08	8.13	443.7	3.3	5	0	2,218	5	17.75	7,875
13	Kotebe University College Building Project 1	180	977	8.99	9.40	500.0	3.0	5	0	2,500	5	16.50	8,250
14	Kotebe University College Building Project 2	180	702	8.20	8.27	500.0	3.0	5	0	2,500	5	16.50	8,250
16	Bole Sub-city Fire and Rescue Agency Building	365	365	28.46	23.87	791.5	3.2	4	0	3,166	4	14.30	11,318
17	Addis Ketema Sub-city Fire and Rescue Agency Building	180	768	10.96	11.19	791.5	3.2	4	0	3,166	4	14.30	11,318
18	Gulele Sub-city Fire and Rescue Agency Building	365	949	36.51	25.56	791.5	3.2	4	0	3,166	4	14.30	11,318
19	Minilik Hospital Building	540	1580	183.32	188.64	1,132.0	3.5	8	1	10,188	9	33.41	37,820

20	Yekatit 12 Hospital Office Building	540	1580	178.48	197.61	1,132.0	3.5	8	1	10,188	9	33.41	37,820
21	Gandi Hospital Building	365	450	20.39	23.23	619.4	3.3	5	0	3,097	5	18.00	11,149
22	Zewditu Hospital Office Building	365	480	8.19	11.56	432.5	3.3	5	0	2,162	5	17.75	7,676
23	Justice Bureau Building	540	803	66.64	61.42	633.6	3.4	13	1	8,870	14	48.54	30,755
24	Environment Protection Office Building	485	1760	43.51	46.75	964.2	3.4	4	0	3,857	4	15.00	14,463
25	Addis Ababa Construction Bureau Laboratory Building	180	730	13.48	14.12	454.6	3.2	5	1	2,728	6	20.70	9,411
26	Education Bureau Building	540	1085	60.34	56.47	633.6	3.4	13	1	8,870	14	48.54	30,755
27	Ethiopian Sport Commision Office Building	540	1420	45.16	51.63	633.6	3.4	13	1	8,870	14	48.54	30,755
28	Urban Integrated Land Information System	540.00	1,273.00	166.26	160.42	900.0	3.0	13	2	13,500	15	46.50	41,850

	Development Office Building												
29	Ministry of Science and Technology Office Building	540.00	1,385.00	30.56	35.23	800.0	2.8	9	0	7,200	9	26.70	21,360
30	Industry Minister Office Building	540.00	1,547.00	30.48	33.76	800.0	3.0	11	0	8,800	11	34.50	27,600
31	Central Statistics Agency Office Building	555.00	1,273.00	23.65	27.35	731.6	3.4	7	1	5,853	8	28.38	20,764
32	Ethiopian Roads Authority Office Building	590.00	1,742.00	39.66	33.88	800.0	3.1	7	0	5,600	7	22.92	18,336
33	Ministry of Foreign Affairs Conference Hall Project	426.00	1,505.00	162.61	190.27	3,100.0	3.7	3	1	12,400	4	16.30	50,530
34	Mofa Car Parking Project	887.00	1,698.00	80.66	123.71	1,900.0	3.0	1	2	5,700	3	10.50	19,950