An Effort Estimation Model for Implementing ISO 9001 in Software Organizations

by

Silas Rahhal

School of Computer Science
McGill University, Montréal, CANADA

October 1995

A THESIS SUBMITTED TO THE FACULTY OF GRADUATE STUDIES AND RESEARCH IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE

ISBN 0-612-08045-5

Abstract

The adoption of software development principles and methodologies embodying best practices and standards is essential to achieving quality software products. Many organizations world-wide have implemented quality management systems that comply with the requirements of the ISO 9001 standard, or similar schemes, to ensure product quality. Meeting the requirement of ISO 9000 can be costly in time, effort and money. The effort primarily depends on the size of the organization, and the status of the quality management system. The focus of this thesis is on the development of an effort estimation model for the implementation of ISO 9001 in software organizations. In determining this effort, a survey of 1190 registered organizations was conducted in 1995, of which 63 were software organizations. A statistical regression model for predicting the effort was developed and validated. The proposed effort estimation model forms a foundation for building and comparing related effort estimation models for ISO 9000 and other process improvement frameworks.

Résumé

L'adoption de principes de développement de logiciel selon les meilleures pratiques et standards est essentielle pour attendre la création de logiciels de qualité. Plusieurs organisations internationales ont implanté des systèmes de gestion de qualité respectant les exigences de ISO 9001, ou de normes similaires pour assurer la qualité des produits.

Le respect des exigences d'ISO 9000 peut être coûteux en argent, effort et temps, les principaux facteurs qui déterminent l'effort sont la taille de l'organisation et l'importance accordée au système de gestion de la qualité. Le but de cette thèse est le développement d'un modèle d'estimation de l'effort pour implantation de ISO 9001 dans l'entreprise de logiciel. Pour determiner cet effort, un questionnaire a été soumis à 1190 entreprises en 1995. 63 de ces organisations étaient des entreprises de logiciel.

Un modèle de regression statistique a été développé et validé pour prédire cet effort. Le modèle proposé représente une base pour créer et comparer des modèles d'estimation d'effort similaires pour ISO 9000 et autres modèles d'amélioration de processus.

Acknowledgments

I would like to thank Professor Nazim H. Madhavji, my supervisor, for his support, supervision and intellectual academic guidance in the process of carrying out this research.

I am also thankful to Khaled El Emam, Won-Kook Hong, Tilmann Bruckhaus, Josée Turgeon, and Dirk Höltje of the McGill Software Engineering team, for their fruitful suggestions and comments, and the School of Computer Science, McGill University, for the pleasant environment it provides.

I would like to thank the management of Bell Sygma Telecom Solutions for their support during my educational leave.

I would like to thank all the 462 organizations that have responded to the questionnaire, and provided the data necessary to conduct this research.

Special thanks to my loving wife Pauline Al-Jiries, and our two children Timothy James, and Karen Jane for all their love and support during my studies.

I am also thankful to my parents, and my brothers Kassan and Danny who assisted me and provided comments and suggestions.

Above all, thanks be to God who has given me grace and strength throughout these years.

Contents

List of Figures

List of Tables

List of Equations

1 Introduction

Software quality, and software process improvement, are two important goals in software engineering. An essential step in achieving these goals, is the construction of quality management systems and process infrastructure, that embodies effective software engineering and management practices.

The International Organization for Standardization (ISO)[1] has published the ISO 9000[2] series of standards on quality (ISO 9001, ISO 9002, and ISO 9003), and organizations world-wide have implemented quality management systems that comply with the requirements of these standards, to ensure product quality.

Meeting the requirement of ISO 9000 can be costly in time, effort and money. The focus of this thesis is on the development of an effort estimation model for implementing ISO 9001 in software organizations.

In this chapter we describe ISO, ISO 9000, and registration cost. In addition, we define the research problem and present the research method and research contribution.

[1] "ISO" is a word, derived from the Greek isos, meaning "equal". This line of thought from "equal" to "standard", led to the choice of using the word "ISO" by the International Organization for Standardization.

[2] Throughout this thesis, whenever the term "ISO 9000" is used, it is intended to mean all three ISO 9000 quality system models, ISO 9001, ISO 9002 and ISO 9003.

1.1 ISO

The International Organization for Standardization (ISO) is a world-wide non-governmental federation of national standards bodies (ISO member bodies), founded in 1946 with an objective to facilitate the international trade through the development, promotion, coordination and unification of industrial standards.

The International Organization for Standardization is headquartered in Geneva, Switzerland, and at present comprises some 100 members, one from each country.

The work of the International Organization for Standardization results in international agreements that are published as international standards. All standards developed by the organization are voluntary; no legal requirements are imposed on any country to adopt the ISO standards. ISO has published nearly nine thousand international standards covering all industrial fields except electrical and electronic engineering which are the responsibility of the International Electrotechnical Commission (IEC).

1.2 ISO 9000

ISO 9000 is a series of international standards that specifies quality system requirements and quality management guidance for generic product categories: hardware, software, processed material and services.

The ISO 9000 series on quality were first published in 1987, and were revised in 1994 to clarify and better explain the requirements of the standard. ISO 9000 standards have been adopted by many nations and the European Community (EC), as the key standard for quality assurance.

ISO 9000 establishes a basic set of quality system requirements necessary to ensure that the process is capable of consistently producing products that meet the customer's expectations. Products and Services are the result of activities or processes that transform inputs into outputs, and the quality of products and services is directly related to the quality of the process that produces them. Therefore, product quality improvement necessitates process quality improvement.

The ISO 9000 standards are generic; they are not written for any particular business. The standards provide *a model, not specifications.* The standards describe the minimum requirements of *what must be done, not how it should be done.*

The underlying essence of the ISO 9000 series of standards is that an organization's quality system must be: (i) documented; (ii) controlled; (iii) auditable; (iv) monitored; (v) improved; and (vi) effective. In addition, management commitment and employee involvement are of prime importance [Sch94].

The ISO 9000 approach, as defined by the standard, can be synthesized by the following key points [Gal94]:

1. ***Say what you do:*** This implies a well documented quality system covering all phases and processes.
2. ***Do what you say:*** This implies that the documented processes and quality related activities are assigned to the appropriate groups, understood, executed, monitored and controlled to ensure the conformance of the enacted process to the documented process.
3. ***Record what you did:*** This implies that records are kept of the execution and outcome, of all quality related activities.

Conforming to the ISO 9000 quality assurance standards is becoming essential to succeed in an increasingly competitive global marketplace. ISO 9000 registration provides a competitive advantage over non-ISO 9000 competitors. Furthermore, the adoption of ISO 9000 by the European Community (EC) and many other countries as their standard for quality assurance has placed a new marketplace pressure on producers world-wide wishing to trade and compete on the national and international markets.

1.3 ISO 9000 Registration and Cost

ISO 9000 registration, (known as ISO 9000 certification in Europe) is granted to an organization when a third-party accredited registrar assesses the organization's quality system and determines that it conforms to the requirements of the ISO 9000 standard. The assessment is done through a registration audit of the facility being registered. The auditors will report any observed nonconformity, and based on their judgment as to the seriousness of the nonconformities (minor or major), recommend for, or against registration.

Registering a facility to ISO 9000 reduces or eliminates the need for individual purchasers or customers to perform their own assessment of the supplier's quality system. In addition, it is an attestation to the world that the quality system conforms to the requirements of the ISO 9000 standard.

Meeting the requirements of ISO 9000 can be costly in time, effort and money [Sch94], [HGJ94]. The cost primarily depends on how much the current quality system conforms to the requirements of the ISO 9000 standard. It requires the allocation of funds and resources for the quality function, and may lead to disruptions and additional training if significant changes are made to the existing processes.

A survey by Deloitte & Touche and Quality Systems Update [CIS93], indicates that up to 80% of the total cost incurred by registered companies is internal, while the remaining is external. Internal cost is defined as the total time-spent (effort) in preparation for meeting the requirements of ISO 9000, from start to completion and successful registration audit. This includes, definition and development of all procedures, instructions, quality manual and training material preparation; staff training;

audit activities, process review, corrective and preventive action; ISO project coordination and administration activities. External cost includes registration fees, subsequent surveillance audit fees, and consultancy fee.

The cost of the actual registration audit (external costs) is predictable, quotable, contractual and fixed. They depend on the size and type of the facility to be registered. However, the costs of developing and implementing a quality system (internal costs) that meets the requirements of the ISO 9000 standard are less predictable and difficult to estimate.

1.4 Problem Definition

Although thousands of companies world-wide have achieved ISO 9000 registration, there has been a lack of work dealing with the estimation of the effort required for implementing ISO 9000. The existence of an ISO 9000 effort estimation model would allow organizations that are contemplating registration to ISO 9000 to predict the effort required for its implementation. This is an *a priori* requirement in corporate decision making, regardless of the maturity level of the organization [hum89], or the current state of the quality system, and to what extent it complies to the requirements of the ISO 9000 standard. Therefore, this thesis addresses this problem, and focuses on this important aspect of estimating the required effort for implementing ISO 9001 in software organizations.

1.5 Research Method

In order to solve the described problem, a viable approach is to build an effort estimation model for implementing ISO 9001. This entails the design of a research instrument, data collection, data validation, data analysis, model testing, and model validation. Therefore, we have addressed these issues extensively in our research to ensure methodological rigor and validation.

We have combined scientific research methods from Software Engineering [Boe81], [BW84], [Bas91], [Pfl95], [Alb79], [Str89], and Social Sciences [KJ86], [AS85], [Fow84], [HJ78] for the conception, design, preparation, execution, data collection, data validation, and data analysis. The research method addressed by these authors consists of seven steps:

1. Establish objectives
2. Plan for required data and resources
3. Research instrument design, development and pre-test
4. Execution of the research instrument
5. Data collection and validation
6. Sample analysis
7. Validation and results

Our research necessitated the collection of data from ISO 9000 registered organizations. Therefore, we surveyed organizations that have achieved registration to ISO 9001, ISO 9002, and ISO 9003. The effort estimation model we present is based on responses received from ISO 9001 registered software organizations. The remaining data is for future research and comparison between estimation models of different ISO 9000 quality system standards [RM95].

1.6 Research Contribution

The originality of this thesis and its contribution to software engineering research and practice, is evident in the proposed ISO 9001 effort estimation model for software organizations. We have thoroughly examined the published literature, and to our knowledge, currently, there are no effort estimation models for ISO 9000. The research value of the proposed model is that it forms a foundation for building and comparing related effort estimation models for ISO 9000 and other process improvement frameworks.

The practical value of this thesis is that it provides organizations with a model for predicting the effort required for meeting the ISO 9001 International Standard, and thus can form a basis for making business decisions concerning the affordability of the implementation effort.

1.7 Thesis Organization

The next chapter describes background material on national and international standards, process improvement, and capability determination models. Chapter three presents an overview of cost estimation and related surveys. Chapter four describes the research design approach, research method, validity, and execution. Chapter five describes the sample data analysis, characteristics, and compliance profile. Chapter six details the statistical approach, model assumptions, variables, and regression. Chapter seven presents the research results, and final estimation model. Chapter eight provides a comparison and discussion. Finally, chapter nine presents a summary and conclusion.

2 Background

This chapter presents a review of ISO 9000. In addition, it describes other related certification and capability evaluation schemes, such as TickIT, SEI CMM (Software Engineering Institute Capability Maturity Model), Trillium, and the proposed SPICE (Software Process Improvement and Capability dEtermination) International Standard.

2.1 ISO 9000

In reviewing ISO 9000, it is important to have a historical, present and future perspective of this international standard. The following few pages will cover the ISO 9000 history, ISO 9000 growth in North America, ISO 9000 future, ISO 9000 structure and ISO 9000 standard elements.

2.1.1 ISO 9000 History

During the past few decades, various national and multinational standards were developed in the quality arena for military, nuclear power industry, commercial, and industrial use. In 1959 the US Department of Defense (DoD) established MIL-Q-9858 quality management program, and in 1968 the North Atlantic Treaty Organization (NATO) adopted the tenet of the DoD program in the NATO's Allied Quality Assurance Publications (AQAP) AQAP1, AQAP4 and AQAP9 series of standards.

In 1979, the UK's British Standards Institution (BSI) developed the first commercial quality assurance system standards from its predecessors. These standards were designated the BS 5750 series, parts 1, 2 and 3.

In 1979, the International Organization for Standardization formed Technical Committee 176 (TC 176) on Quality Management and Quality Assurance to address the world-wide trend toward quality, and the growing confusion in international trade resulting from differing national quality system requirements.

In 1987, based on the work of TC 176, the International Organization for Standardization published the ISO 9000 series of standards on quality. The series, among other things, comprises three quality system models:

- ISO 9001 Quality systems - Model for quality assurance in design, development, production, installation and servicing [ISO9001].
- ISO 9002 Quality systems - Model for quality assurance in production and installation and servicing [ISO9002].
- ISO 9003 Quality systems - Model for quality assurance in final inspection and test [ISO9003].

In 1994, the first revision to the ISO 9000 series was completed. The 1994 revision contained no new major requirements, but rather was aimed at clarifying and better explaining the requirements of the standard.

2.1.2 ISO 9000 Growth

ISO 9000 registration is rapidly becoming a business imperative. According to the latest Mobil Survey, there are approximately 70,000 companies registered world-wide [Bau95]. Many governments are moving towards ISO 9000 for their procurements; most large companies have already done so, and now, small and medium-sized companies are also getting in on the act.

ISO 9000 certificates awarded in Canada and the United States has nearly doubled in 1994 over the accumulated total reported in the previous seven years (1987- 1993) [CIS95]. We present in Figure 2-1 and Figure 2-2 the ISO 9000 cumulative registration growth in Canada and the United States.

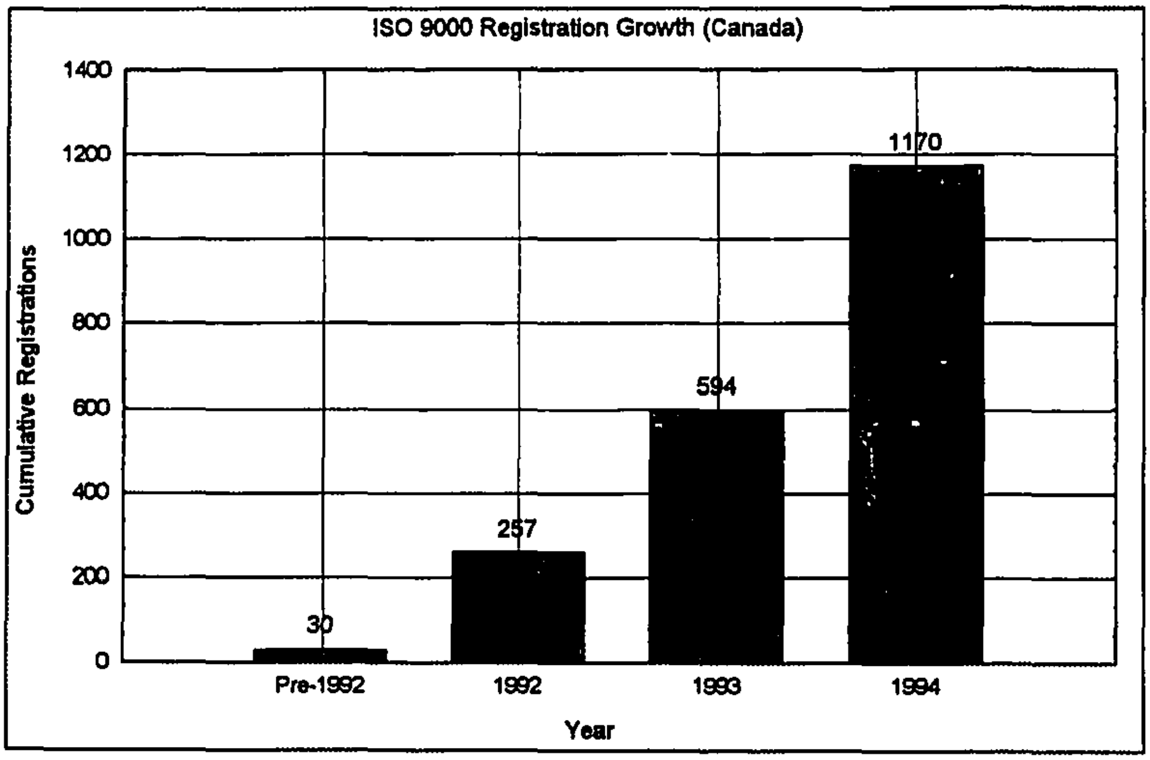


Figure 2-1: Cumulative ISO 9000 registrations reported in Canada

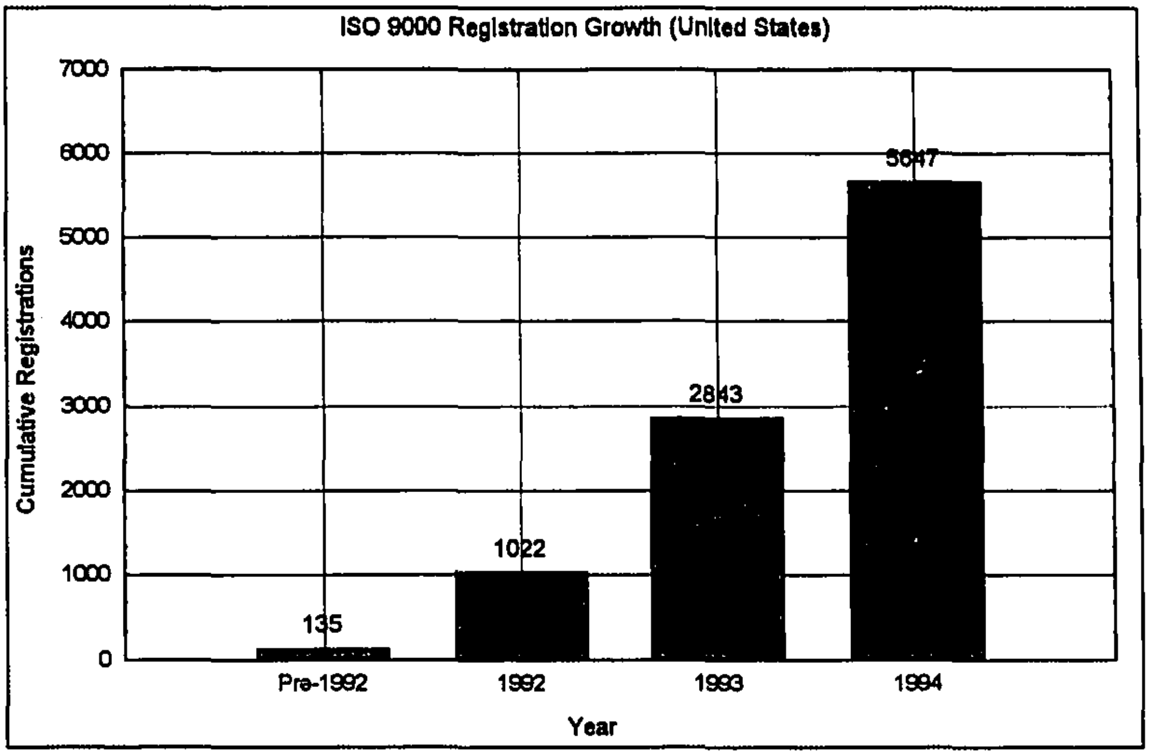


Figure 2-2: Cumulative ISO 9000 registrations reported in the United States

2.1.3 ISO 9000 Future

The ISO 9000 series of standards are not static documents but will continue to be revised. An ad hoc task force was commissioned to prepare a strategic plan for ISO 9000 series architecture, numbering and implementation. The task force report, which has become known as "*Vision 2000 - A strategy for International Standards' Implementation in the Quality Arena During the 1990's*" [MCJ91] calls for a number of modifications, including providing additional guidance on applying the series to the four generic product categories that were identified. The four generic product categories are:

- Hardware
- Software

- Processed Material
- Services

Also four strategic goals were set for the ISO 9000 series and related standards. They are:

- Universal Acceptance
- Current Compatibility
- Forward Compatibility
- Forward Flexibility

As ISO 9000 becomes more universally adopted and accepted, and as more government bodies world-wide begin to use ISO 9000 series, the market will eventually dictate the future of ISO 9000 - and market facts speak loudly in favor of the standards. The ISO 9000 series has been adopted by over 90 countries, and for many governments, the only decision left may be, when, not if, ISO 9000 should be the basis of quality assurance in their procurement activities.

2.1.4 ISO 9000 Structure

The ISO 9000 series of standards on quality covers, quality management, quality assurance and quality systems.

The basic ISO 9000 series comprises two types of standards: *guidance standards* and *conformance standards*.

Guidance standards are descriptive documents, not prescriptive requirements. That is, companies do not register to them, but rather use them as guidance to one of the conformance standards. (An example of a guidance standard is ISO 9000-3:1991, which provides guidelines for the application of the ISO 9001 to software [ISO9000-3]).

Conformance standards Specify requirements that must be performed. Companies are required to conform to the requirements specified in the conformance standards. The three ISO 9000 conformance standards are: ISO 9001, ISO 9002 and ISO 9003.

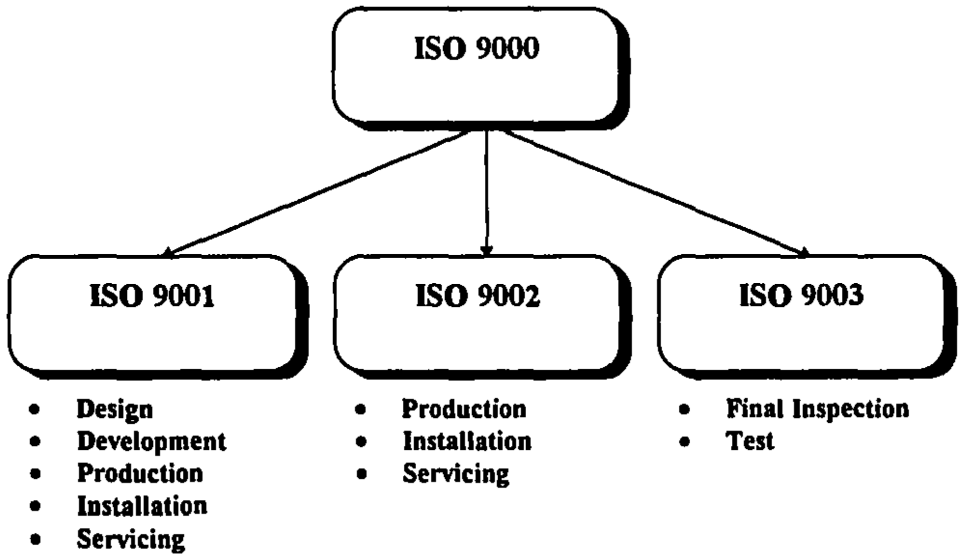


Figure 2-3: ISO 9000 structure and series of Quality System standards

Software products are inherently complex and challenging to scope, develop, implement, verify, validate, and maintain. They can quickly reach a high level of design complexity [Coa94]. In addition, the process of development and maintenance of software is different from that of most other types of industrial products. In such a rapidly evolving technology field it is therefore necessary to provide additional guidance for quality systems where software products are involved. This has led to the publication of the guideline standard ISO 9000-3 which facilitates the application of ISO 9001 to organizations developing, supplying and maintaining software.

2.1.5 ISO 9000 Standard Elements

The ISO 9000 series of standards (Revision 1994), specifies 20 standard elements. ISO 9001 requires conformance to all of the 20 elements, while ISO 9002 and ISO 9003 require conformance to 19 and 16 of the elements respectively. Also, some of the required elements in ISO 9003 are less stringent than the corresponding elements in ISO 9001.

Table 2-1, lists the 20 elements of ISO 9000 series, and indicates the extent to which they apply to each of ISO 9001, ISO 9002 and ISO 9003.

Clause no.	ISO Clause (element) title in ISO 9001	ISO 9001	ISO 9002	ISO 9003
4.1	Management responsibility	✓	✓	⊂
4.2	Quality system	✓	✓	⊂
4.3	Contract review	✓	✓	✓
4.4	Design control	✓	×	×
4.5	Document and data control	✓	✓	✓
4.6	Purchasing	✓	✓	×
4.7	Control of customer-supplied product	✓	✓	✓
4.8	Product identification and traceability	✓	✓	⊂
4.9	Process control	✓	✓	×
4.10	Inspection and testing	✓	✓	⊂
4.11	Control of inspection, measuring, and test equipment	✓	✓	✓
4.12	Inspection and test status	✓	✓	✓
4.13	Control of nonconforming product	✓	✓	⊂
4.14	Corrective and preventive action	✓	✓	⊂
4.15	Handling, storage, packaging, preservation and delivery	✓	✓	✓
4.16	Control of quality records	✓	✓	⊂
4.17	Internal quality audits	✓	✓	⊂
4.18	Training	✓	✓	⊂
4.19	Servicing	✓	✓	×
4.20	Statistical techniques	✓	✓	⊂
Key:	✓ = Comprehensive requirement ⊂ = Less comprehensive requirement than ISO 9001 × = Element not present			

Table 2-1: ISO 9000 elements

2.2 TickIT

TickIT is a scheme devised by IT professionals and supported by the UK Department of Trade and Industry (DTI), and the British Computer Society. The objective of the TickIT scheme is to establish a relevance of the ISO 9001 standard to the production of software. The TickIT scheme applies to the software industry and it requires that the assessment of quality management systems be carried to the requirements of ISO 9001 through the application of ISO 9000-3 guidelines [Tic92].

TickIT certification is common in the European Community, and specially in the United Kingdom. The TickIT scheme and registration differ from ISO 9001 software registration in the following areas:

1. The existence of a TickIT guide *"TickIT guide to Software Quality Management System Construction and Certification using EN290001"* which is used as a guidance standard and contains ISO 9000-3, a purchaser's guide, a supplier's guide, and an auditor's guide.
2. A TickIT auditor must have three to five years experience in software development.
3. A TickIT auditor must successfully pass the continuous assessment and written examination of TickIT lead auditor training course, and an oral exam.

2.3 Capability Maturity Model

The Capability Maturity Model (CMM) for software is a process maturity framework, developed by the Software Engineering Institute (SEI) with assistance from the MITRE Corporation. This effort was initiated in response to a request by the U.S. Federal Government, to provide a method for assessing the capability of its software contractors [CMM93].

The CMM v1.1 is composed of five maturity levels. Each maturity level (except level one), is composed of several *Key Process Areas*. Each *Key Process Area* is organized into *Common Features*. *Common Features* specify *Key Practices*. *Key Practices*, when implemented, help satisfy the goals of the key process area.

The CMM software process maturity levels are:

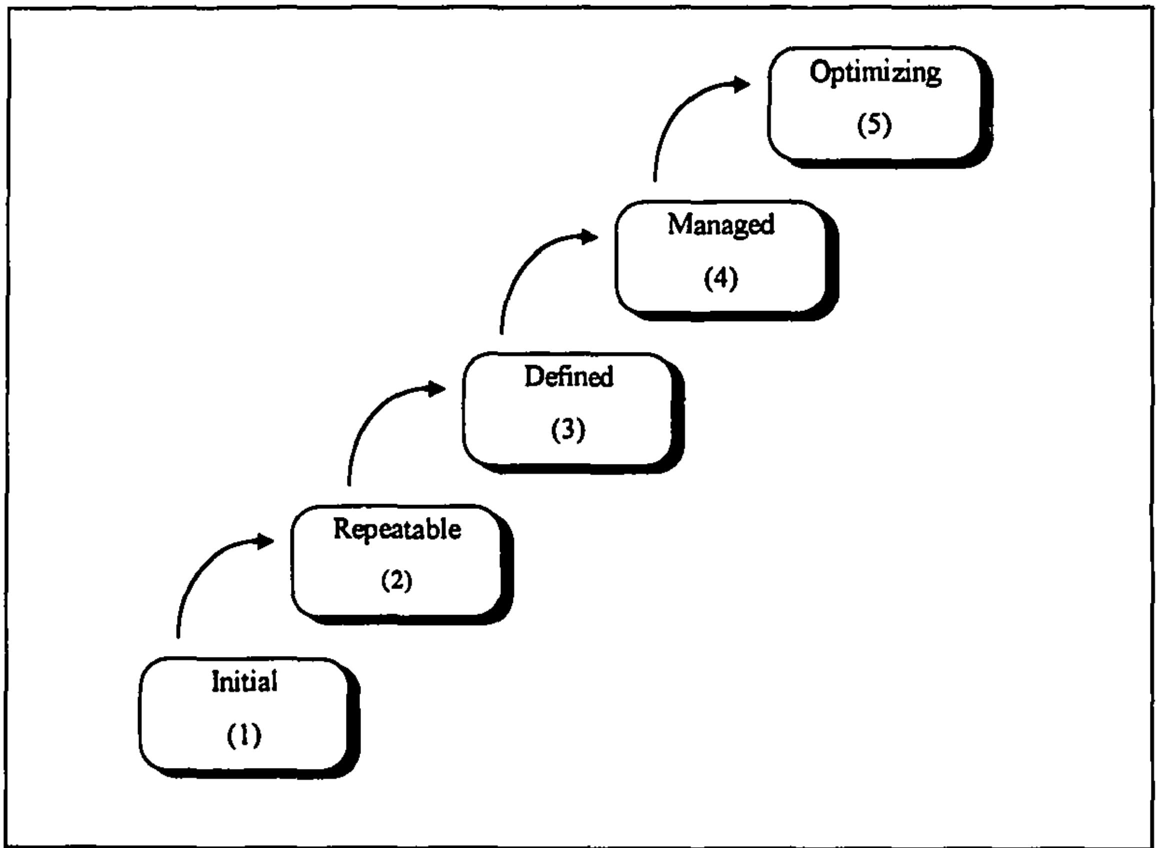


Table 2-2: The five levels of Software Process Maturity

Each maturity level provides a layer in the foundation for continuous process improvement. Continuous process improvement is based on many small, evolutionary steps rather than revolutionary innovations [Ima86].

2.4 Trillium

Trillium is a model for telecom product development and support process capability. The primary objective of *Trillium* is to assess the product development and support capability of prospective and existing suppliers of telecommunication or information technology-based products. *Trillium* may also be used as a reference benchmark in an internal capability improvement program [Tri94].

Trillium is based on the Software Engineering Institute (SEI) *Capability Maturity Model* (CMM) version 1.1. It was primarily designed for embedded software systems, although it can be applied to Management Information Systems (MIS).

The *Trillium* model architecture consists of *Capability Areas*, *Roadmaps* and *Practices*. A software development organization is measured by an assessment of the practices achieved within each defined capability area. As a result of the assessment, a capability profile of the organization is determined.

Similar to the SEI CMM version 1.1, *Trillium* places an organization on a scale that spans 5 levels. The *Trillium* levels that characterize the capability areas of an organization are: (1) Unstructured; (2) Repeatable and project oriented; (3) Defined and process oriented; (4) Managed and integrated; and (5) Fully integrated.

The *Trillium* model consists of eight *Capability Areas*. Each *Capability Area* contains practices at multiple *Trillium* levels that are incorporated within one or more roadmaps. Table 2-3 presents a mapping of the capability areas, roadmaps, and the spanning level of practices.

Capability Area	Roadmaps	Spanning Level of Practices 2	3	4	5
1. Organizational Process Quality	1. Quality Management	■	■	■	
	2. Business Process Engineering	■	■		
2. Human Res. Dev. & Management	1. Human Res. Dev. & Management	■	■	■	
3. Process	1. Process Definition	■	■	■	
	2. Technology Management	■	■	■	■
	3. Process Improv. & Engineering	■	■	■	■
	4. Measurements	■	■	■	■
4. Management	1. Project management	■	■	■	
	2. Subcontractor Management	■	■		
	3. Customer-Supplier Relationship	■	■		
	4. Requirements Management	■	■	■	
	5. Estimation	■	■		
5. Quality System	1. Quality System	■	■	■	■
6. Development Practices	1. Development Process	■	■	■	
	2. Development Techniques	■	■	■	■
	3. Internal Documentation	■	■		
	4. Verification & Validation	■	■	■	
	5. Configuration Management	■	■	■	■
	6. Re-Use	■	■	■	■
	7. Reliability Management	■	■	■	
7. Development Environment	1. Development Environment	■	■	■	■
8. Customer Support	1. Problem Response System	■	■		
	2. Usability Engineering	■	■	■	
	3. Life-Cycle Cost Modeling	■	■	■	
	4. User Documentation	■	■	■	
	5. Customer Engineering	■	■		
	6. User Training	■	■		

Table 2-3: Trillium Capability areas, Roadmaps, and practices..

2.5 SPICE

The International Organization for Standardization (ISO) and International Electrotechnical Commission (IEC) created a new working group (WG10) to develop an international standard on software process assessment. WG10 created a special project called SPICE (Software Process Improvement and Capability dEtermination).

SPICE is an emerging ISO-sponsored standard and includes the following assessment methods and models: the CMM (SEI), TRILLIUM (Bell Canada), ISO 9000, Bootstrap (Esprit), HealthCheck (BT), SAM (BT), SQPA (HP), and STD (Compita).

The proposed SPICE standard will deal with software processes (e.g., development, management, customer support, quality, etc.) but will be also concerned with people and technology. It is intended to be applicable to different business sectors, application domains, and organization sizes, as well as being culturally independent [SPN95].

SPICE can be used by organizations in different modes:

1. Capability determination mode (2nd party), where an assessment is made of a potential software supplier to determine the supplier's capabilities.
2. Self-assessment mode (1st Party), where an organization assesses itself with an objective of determining its capabilities, and abilities to undertake a new project.
3. Process improvement mode, as a result of an assessment, to help improve the software development process of an organization.

The proposed SPICE standard consist of seven separate products:

1. Introductory Guide (IG): This provides a high level overview of the SPICE structure and how the SPICE products fit together, and provides guidance on selection and use.
2. Baseline Practices Guide (BPG): Identifies the set of basic practices that are fundamental to good software engineering. The guide describes what activities are required, not how they are to be implemented.

3. Assessment Instrument (AI): A generic tool for the extraction of data relating to the processes undergoing assessment. In addition, it helps in detecting the presence or absence of baseline practices.

4. Process Assessment Guide (PAG): This guide specifies the assessment method and defines how an assessment is conducted using the Assessment Instrument and the Baseline Practices. It also sets out the basis for rating, scoring and the process capability profile.

5. Process Improvement Guide (PIG): Provides guidance for process improvements based on the results of a process assessment.

6. Process Capability Determination Guide (PCDG): Provides guidance on determining an organization's capability and risks based on the results of a process assessment.

7. Assessor Training and Qualification Guide (ATQG): Provides guidance for the development of assessor training programmes and defines procedures for the qualification of assessors.

2.6 Inference

The existing and emerging national and international standards, models, and frameworks emphasize the importance of process definition, process implementation, process assessment, and process improvement. This is specially important today, because of the exponential increase in software and hardware complexity, and technological advances. The world-wide movement towards attaining one or the other of the described standards and schemes, reemphasizes the importance of estimating the effort required of such an implementation. This thesis is a step towards solving this problem, with a focus on ISO 9001 for software organizations.

3 Cost Estimation

An estimate, although a judgment using expertise and knowledge, should be based on a method using a set of rules that are defined, recognized, quantifiable, and verifiable [Lon87].

An estimation method is successful when:

1. It produces estimates that are within an acceptable limit of the final cost.
2. The estimation method is easy to use, and the set of rules used are reasonably understandable.
3. The estimation method is supported by tools and documented procedures. Tools can increase the effectiveness, accuracy and speed of the estimation process.
4. The estimation method is repeatable and verifiable.

3.1 Estimation Models

It is important to recognize that there are no methods currently available that accurately estimate the cost of producing 10,000 software lines of code, or source instructions, similar to the methods available for estimating the cost of producing 10,000 units of a specific telephone set, or 10,000 units of floppy disk drives [Boe81]. The primary reasons are that software source instructions are not a uniform commodity and require the creativity, cooperation and understanding of individuals.

The issue of software development cost estimation models has been addressed in the last two decades, and as a result many cost estimation models have been proposed. A review of software estimation models such as, Farr and Zagorski Model, Wolverton Model, ESD Model, RCA Model (Process-S), Walston and Felix Model, Putnam Model (SLIM), Boehm Model (COCOMO), and Albrecht's (Function Point) is provided by [Lon87]. However, there has been little research done on the development of estimation models that are related to the process.

The ISO 9000 standard, and the SEI Capability Maturity Model, and other schemes are concerned with the development, documentation, implementation and improvement of the process of an organization. This thesis builds a foundation on which future research can be based, and provides a method whereby organizations can estimate the effort required for implementing process work.

3.2 TickIT Registration Costs

The TickIT campaign office of IT World Limited Management Consulting has conducted several TickIT case studies in 1993 [Tic93], and has reported on the benefit and cost of the TickIT registration.

The case studies provided information on the effort and cost involved in achieving TickIT certification for 5 companies in the UK. The reported cost per head of staff ranged between £50 and £3,416. While the certification overhead costs averaged over 3 years, ranged between £30 and £50. Some companies involved in the study have reported annual savings that surpassed the initial investment made. This implies the existence of element(s) other than the size of an organization, that affects the cost of registration.

3.3 CMM Software Process Improvement Costs

The SEI has reported on the cost and business value resulting from the CMM-Based Software Process Improvement (SPI) efforts of five organizations [HCR94]. The reported amount of money spent on software process improvement activities varied greatly from one organization to another. The cost for SPI activities per year ranged from $49,000 to $1,203,000 with $245,000 as the median. Even after normalizing the cost incurred by the number of "software engineers" or "software professional" as the denominator, significant variances were reported. The cost per software engineer ranged from $490 to $2000 with $1375 as the median.

Although the SEI has reported a substantial increase in the number of first time assessments of software organizations, there has been no published or known models to estimate the effort or cost of implementing the CMM-based Software Process Improvement model. Therefore, the development of an effort estimation model for ISO 9001 sets the foundation for future process related models [Pau94].

3.4 Other ISO 9000 Surveys on Cost

A few surveys and case studies have been carried out by a number of organizations on different aspects of ISO 9000 [SPQ94], [CIS93]. The primary focus in these studies was "the benefit (tangible or otherwise) of ISO 9000 registration." In one survey, 74% of respondents felt that the benefits of ISO 9000 out-weighed the costs.

Another survey, which was sent to 72 organizations in 1993 [ATT94] requested data on the effort used in becoming ISO 9000 compliant. Although the reported data from the 24 respondents in this survey does provide some insight into the effort required for implementing ISO 9000, it does not propose an estimation model that organizations can use to predict their effort. Table 3-1: Staff, presents the reported findings:

Size of the Organization (Number of Employees)	Effort (years)	Median Effort (years)
< 500	0.4 - 5.7	1.9
500 - 2000	1.8 - 12.5	5
> 2000	6 - 22.5	17.5

Table 3-1: Staff effort to become compliant (source: AT&T)

3.5 Summary

Organizations today are faced with the problem of estimating the effort required of implementing ISO 9000 and other such frameworks, similar to the problem of predicting the size of a software system and its development effort. The background work described highlights the lack of an appropriate cost estimation model for ISO 9000. In order to build such a model empirically, it is important to do research design and execute this design appropriately.

4 Research Design

This chapter defines research design and describes the research methods available. Also discussed are: research design quality, validity issues, research method selection, and the survey research method used in our study.

The sample data analysis and the statistical approach used for model building and validation are presented in chapter five and six respectively.

4.1 Research Design Definition

A research design is an action plan "blueprint" that specifies a series of steps that guide the researcher in the process of determining and designing the questions to be studied, collecting relevant data, analyzing data and validating results. The research design is the logic that links the data to be collected and the conclusions to be drawn, to the initial questions of a study [Yin94]. A good research design avoids the situation in which the evidence does not address the initial research problem.

4.2 Available Research Methods

Scientific research methods and experiments require a great deal of design and planning if they are to provide meaningful and useful results. Researchers use different research methods depending on the type of analysis that is required. Some of the most common research methods used in software engineering and social sciences are:

1. Randomized experiments
2. Quasi-experiments
3. Pre-experiments

Randomized experiments: are ideally suited for the task of cause analysis. Often they must be conducted in laboratory setting with proper controls of a treatment group and a control group. The strength of randomized experiments is their internal validity. The greater the control the experimenter has, the greater the internal validity of the experiment. The results obtained from such laboratory settings are difficult to generalize due to the unusual settings that may affect the behavior of the subjects, thus the weakness of randomized experiments stems from the fact that the subjects are well controlled.

Quasi_experiments: Quasi experiments are research designs without randomly assigned treatment and comparison groups. Quasi-experiments can assess causes and effects and do not have the drawback of the difficulty to generalized.

Pre-experiments: Pre-experimental research design offers the possibility of gathering data from subjects in their natural settings. Pre-experimental research design is useful when issues of internal validity are not a major concern. There are a wide variety of circumstances when pre-experimental research design is enormously useful, adequate, and sufficient [KJ86].

4.3 Research Design Quality

The criteria for judging the quality of research design and research study were first published by Campbell and Stanley [CS63] and referred to these as threats to internal and external validity, respectively. Cook and Campbell [CC79] subsequently suggested that actually four types of validity checks should be distinguished: 1) Construct Validity; 2) Internal Validity; 3) External Validity; 4) Statistical Conclusion Validity.

1. *Construct Validity:* Is the degree to which the causal construct (independent variable) and the affected construct (dependent variable) accurately reflect or measure the constructs of interest.

2. *Internal Validity:* Is a concern on the extent to which conclusions can be drawn and inferences made on the causal effects of one variable on another.

3. *External Validity:* Is the issue of generalization of the research results that are based on a sample to a larger universe.

4. *Statistical Conclusion Validity:* Is the extent to which the variables demonstrate relationships not explainable by chance or some other standard of comparison. There are two ways in which this can happen: Type I errors, or false positive - that is, concluding there is a relationship between two variables when there is none - and Type II errors, or false negatives - that is, failing to detect a relationship that in fact exists in the population [KJ86].

A review of MIS research has revealed that the issue of research instrument validation has often been inadequately addressed [Str89]. This concern calls for methodological rigor in experimental design, validation, inferences and conclusion. Therefore, we have addressed these issues extensively in our research, to ensure methodological rigor and validation is carried fully in our research design and execution. We have borrowed from other disciplines, such as social and physical sciences, rigorous techniques and requirements for validating experimental design, and research instruments.

4.4 Research Method Selection

The choice of a research method is a far-reaching decision that a researcher must make. It has implications on the validity of the whole research. Robert K. Yin, in his comparison of research strategies [Yin94], provided a table, Table 4-1, for identifying the likely research that should be used under different conditions. If the research questions focus mainly on "what" questions, either of two possibilities arises. First, some types of "what" questions are exploratory, and may require the use of exploratory study with pertinent hypotheses and proposition. However, the second type of "what" questions, is actually a form of "how much" or "how many" - for example, "what was the score of yesterday's hockey game ?" This second type of "what" and other forms of questions like "who", "where" "how much" and "how many" favor survey research strategies.

In addition, to the arguments presented below, our research focuses on contemporary events that require no control over behavioral events. Therefore, the most appropriate research method for this study is pre-experimental survey research.

Strategy	**Form of research question**	**Required control over behavioral events ?**	**Focuses on contemporary events ?**
Experimental	how, why	yes	yes
Survey	who, what, where, how many, how much	no	yes
Archival analysis	who, what, where, how many, how much	no	yes / no
History	how, why	no	no
Case Study	how, why	no	yes

Table 4-1: Relevant Situations for Different Research (Source Yin94)

Randomized experiments are highly specialized tools that are suited to the task of causal analysis. They require great control and manipulation of the variables, random assignments, treatment groups and control groups. This would not have been possible, since our objective is to observe the effort spent by organizations that have achieved registrations, and should not consider collecting data from those that have not achieved registration. Therefore, randomized experiments are not suitable for our research.

Quasi-experiments are research designs similar to randomized experiments but do not have randomly assigned treatment and comparison groups. Their objective is causal analysis and comparisons of the effect of some treatment versus no treatment. For this reason, they do not fulfill our research objective, and therefore, are not suited for our research.

Pre-experimental survey research was clearly the most appropriate research method, because the primary objective of our research was not to describe causal effects, nor make conclusions about the causal effect of one variable on another, but rather to answer questions about the distribution of the characteristics of populations in their natural settings [KJ86]. Such research is known as survey research.

The choice of the data collection method in a survey research has implications for response rates, question form and format, survey costs and resources. Surveys can be conducted in several ways:

- Face-to-face interviews
- Self administered questionnaire
- Telephone survey
- Computerized response telephone survey

In choosing a data collection method for our survey research, we choose the self administered mail questionnaire. A full coverage of traditional data collection methods utilizing face-to-face interviews and self-administered mail questionnaires is given in [MK71], [H+78].

4.5 Survey Research Method

The research method described below, which involves the steps outlined in section 1.5, was used for the entire 1190 survey questionnaires mailed. However, the model presented in this thesis is based on the responses received from ISO 9001 software registered organizations.

4.5.1 Establish Objectives

The primary objective of our research was to build a model that would permit us to predict the required effort for implementing ISO 9001 in software organizations.

4.5.2 Plan for Required Data and Resources

Accurate and effective specification and planning of the required data are crucial to the success of the survey research, because, errors or inadequacies in data requirement specification will be amplified during the later steps of the survey research.

In our research, the dependent variable that we wish to predict is the *effort* required for implementing ISO 9000. We avoided the use of dollar estimates, due to variance in wages paid by different organizations. Effort is predicted in person months, where each person month consists of 152 hours of working time. This has been found to be consistent with practical experience [Boe81], [Kem87], and other software engineering and software development estimation models [Lon87].

The other required data for our research was the independent variable(s) that would have a measurable effect on the dependent variable. There are many independent variables that influence the effort required for implementing ISO 9001, such as size of organization, management commitment, management and employee capabilities, employees involvement, business complexity, degree of non-compliance to the standard, number of sites included in the scope, etc. Our research and a review of many ISO 9000 literature [Sch94], [Pea94], [CIS93], [Tic92], [RB93], and [ATT94], suggests that the two primary independent variables are:

1. The size of the organization (number of employees).
2. Degree of non-compliance (percentage of the ISO 9001 requirements that the organization has still to satisfy to become fully compliant).

Size of the organization

Although there are numerous ways to measure the size of an organization (annual sales volume, production capacity, assets, number of employees, etc.). The ISO 9000 standard is concerned with the quality management system and the process that is carried by those in the organization. Therefore, the measure of the size of the organization for this research is with respect to the number of employees.

Degree of non-compliance

The degree of non-compliance is a measure (percentage) of the ISO 9001 requirements that are not being met. (For example, if the organization had no quality management system in place, no defined or documented processes, no quality records, and no employee quality training program, then the organization would be near 0% compliance, or said alternately, the degree of non-compliance would be near 100%.)

Sampling design

The concept of sampling is to take part of some population to represent the whole. The reliability and validity of survey data depends in part on the sampling design used. Our objective in sampling design is to increase the validity and reliability of the survey data, and reduce the sampling error and sampling bias.

The most desirable sample for almost every survey is a random sample, thus the probability of any one organization being included in the sample is precisely equal to the probability of including any other organization. Such a sample has statistical properties that allow the researcher to make *inferences* about the population, set *confidence intervals*, and report *statistical significance*.

In sampling design we begin with the identification of the population to be surveyed. Then, we specify the sample unit, or the smallest entity that will provide one response. The sample frame is a list or a set of directions indicating all the sample units in the population.

The population of all ISO 9000 registered organizations in North America as of December 1994, were listed in different ISO 9000 directories. The sample selection was done randomly and without replacement. This is known to be the preferred random sampling method.

The sample size depends on the degree of confidence level required, resource allocated for the study, time and budget constraints. The sample size selected for this research amounted to 1190 survey recipients. Figure 4-1, shows the breakdown of the sample size by ISO 9000 quality standard. The number of surveys mailed per standard is shown in squared brackets.

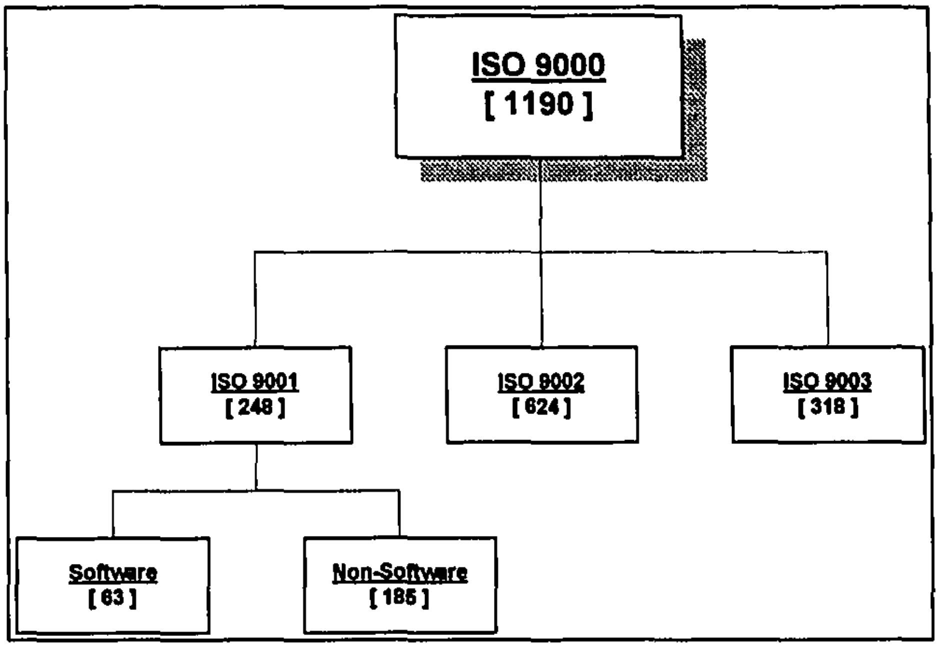


Figure 4-1: ISO 9000 Survey Breakdown

The identification of registered ISO 9001 software organizations was possible through the use of the Standard Industrial Code and the registration scope. The scope identifies the business functions that were audited and registered to the ISO standard. This was an important selection criterion for software organizations in our research.

Information Source

The population of registered organizations surveyed[3], were selected from three pools of published ISO 9000 directories. There were no exclusions or unlisted registrations reported. The three directories used were:

1. CEEM Information Services (ISO 9000 Registered Company Directory)

2. QMI (Directory of Registered Companies)

[3] Surveyed companies were registered to ISO 9000 by different U.S. and Canadian registrars: (listed in alphabetical order), ABS QE, AGA, ASCERT, AT&T QR, AV Qualite, BSI, BVQI, DNV, INTERTEK, KEMA, KPMG, LRQA, LSG QSR, NSAI, QMI, SGS ICS, SGS ICS Canada, UL, UL & BSI.

3. Litton Systems Canada Limited Quality System Registrars (LSL QSR Directory of Registered Quality Systems).

Surveyed Industries

Companies are classified into industrial sectors by the Standard Industrial Code (SIC). This code enables us to perform inter-industry and intra-industry comparisons. Each industrial sector has its own SIC number(s). Surveyed industries[4] were:

Surveyed Industries	Surveyed Industries
Agriculture, forestry and fishing	Wholesale Trade
Mining	Retail Trade
Construction	Business Services (Excluding Software)
Manufacturing	Software
Transportation, Communication and Public Utilities	Other industries

Table 4-2: Surveyed Industries

4.5.3 Research Instrument Design, Development and Pre-test

The design and development of the research instrument is an important element in the survey process. The research instrument used was the survey questionnaire. The quality of the questionnaire and the way it is presented affect the quality of responses and the response rate. Therefore, several checklists and guidelines were used to ensure the correctness of the questionnaire's construct [KS82], [AS85].

Survey Questionnaire

The questionnaire, [Appendix B], was constructed and organized in a logical sequence that avoided sudden breaks and dramatic changes to the respondent's frame of reference [SB82], [MD90]. In addition, conditional and unconditional branching that

[4] The Standard Industrial Codes may differ from one country to another, and are in the process of being recoded.

tends to increase errors and reduce the reliability and validity of the results were avoided [Opp66].

The use of diagrams, Figure 4-2, to establish common understanding of the questions, and capture several pieces of data, was instrumental in achieving proper responses.

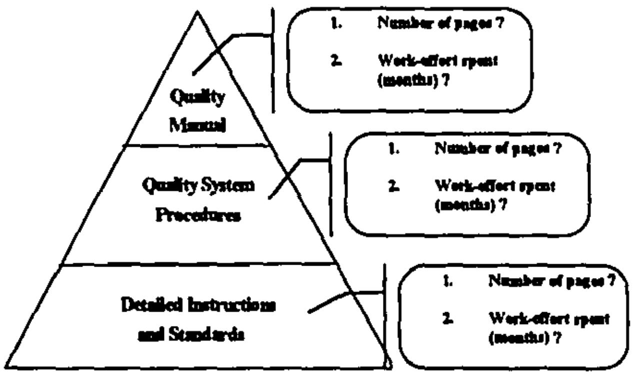


Figure 4-2: Typical quality system documentation hierarchy.[5]

The length of the questionnaire is vital for an adequate response rate. The longer the questionnaire, the fewer are the respondents. Several design iterations and tests were made to maintain a reasonable length of the questionnaire. This resulted in a questionnaire that takes approximately 5 to 10 minutes to complete.

Pre-coding was accomplished by mailmerging the recipient's name, company, Standard Industrial Code, and ISO 9000 standard onto the questionnaire. This step facilitated the identification, classification, data validation and tracking of responses.

[5] *A typical quality system documentation hierarchy consists of three document structure. The first tier, called the Quality Manual; describes the company policy, objectives, and plan for quality. The second tier, called the Quality System Procedures, describes the who, what, when, where, and why of an activity or function. The third tier, called Detail Instructions and Standard are desk/work instructions which provide details on how specific tasks are performed, including forms, reports and other.*

Validation and review of the survey questionnaire were done by one quality system registrar, one professor, and three researchers in software engineering. Similarly, pre-tests were made on twelve organizations who had achieved ISO 9000. This led to further refinement of the survey questionnaire.

Finally, the last page of the questionnaire included a note thanking the respondent for completing the questionnaire, and urging a prompt response.

Cover letter

The cover letter or "letter of transmittal" is an important piece of the entire survey package. Recipients are more likely to read and respond to a personally addressed and hand signed letter, than they are to a "general" letter that is unaddressed and obviously bulk printed. Thus the researcher can increase the response rate by including a personalized letter. The cover letter must explain the project and win the cooperation of the recipient.

The cover letter used, was personalized by name to each ISO 9000 coordinator, and individually hand signed. Recipients were informed of the survey objective, and were assured of the complete confidentiality of all information obtained. In addition, participants were advised that they will receive a summary of the research results.

Table 4-3 presents the checklist that was used in preparing the cover letter of the ISO 9000 survey questionnaire. [Appendix 'A'].

	Question to be answered	How it was answered
1	What is it about ?	ISO 9000 effort estimation
2	Who wants to know ?	Researchers at McGill University
3	Why do they want this ?	Gain insight into effort required for implementing ISO 9000
4	Why was I picked ?	Registered to ISO 9000
5	How important is this ?	Important to organizations, industry, and research
6	Will this be difficult ?	No
7	How long will this take ?	Five to Ten minutes
8	Will it cost me anything ?	No additional cost
9	Will I be identified ?	Complete Confidentiality
10	How will this be used ?	At aggregate level for effort estimation model building.
11	What is in it for me ?	Respondents will receive survey results
12	When should I do this ?	Mail before deadline

Table 4-3: Cover letter checklist

4.5.4 Execution of the Research Instrument

The enormous work involved in preparing each survey package, led to three separate mailings:

- First mailing Thursday, February 9, 1995: 320 Surveys
- Second mailing Friday, February 10, 1995: 405 surveys
- Third mailing Monday, February 13, 1995: 465 surveys

The survey package included, the cover letter, the survey questionnaire, and a self-addressed return envelope (stamp not included). Several measures were taken to improve the response rate, they include:

1. Use of authentic McGill University letter head and envelopes
2. Use of quality laser printing and graphics for the questionnaire and the cover letter

3. Personal hand signature of each cover letter

4. Printing the name of the ISO 9000 management representative/coordinator on each cover letter and envelop

5. Inclusion of a self addressed return envelope

6. Commitment to provide all respondents with the research results

7. Carefully packaging and quality controlled

4.5.5 Data Collection and Validation

Considerable amount of data was generated by the 462 responses received from the survey. Although the ISO 9001 responses from software organizations amounted to 28 responses, the data collection process included all 462 responses.

Data Entry

Responses were dated and given a sequence as they were received. Then, the data entry was made into a computer using Microsoft Excel workbook as the repository. Two provisions were made to prevent and minimize data entry errors:

- Data entry sessions were limited to about an hour, thus eliminating fatigue that could result in data entry errors.
- Data entry sessions were followed by a printout of the data entered. The printouts were compared manually to the original responses received.

Not all the data gathered from this questionnaire has been used in the model building reported in this thesis. Only the data resulting from questions 1, 4 and 7 were used. The remaining data is for further analysis of issues outside the scope of this thesis.

Data Accuracy

Errors can occur at all stages of the survey process [PS81]. Some errors will have negligible effect on the results, while others may have serious consequences. The objective of the data collection method is to reduce the number of errors and minimize their effect.

A response error is simply the difference between the true and the observed measurement. The error may have originated from the respondent, or caused by the computer operator during data input, or caused by the data analyst during the data analysis [Jol86].

The measures taken to ensure the accuracy of the data, included the creation of several macros and formulas to automatically calculate totals and flag out-of-range conditions as the entries were being made. These measures proved to be quite effective in identifying discrepancies.

Out of the four hundred and sixty-two responses received, nine responses had totaling errors. In five of the cases the errors were obvious, and were easily corrected. However, the remaining four cases required us to contact the respondents and revisit the data provided before the cases could be accepted.

Data omissions

Although there are several statistical methods that deal with missing data values, (e.g., casewise deletion, mean replacement value, pairwise deletion), we valued every response, and therefore identified all responses with data omissions.

A total of twenty-one responses out of the four hundred and sixty-two were identified as having missing data. Only one of the twenty-one responses with missing data belonged to the ISO 9001 software responses. The missing information was the size of the organization. After contacting the responding organization, and obtaining the missing information, we were able to include that response in our analysis. The other data omissions were dealt with similarly.

Response Validation

Response validation is an important step in the research and data collection process. The validation consists of checking the responses for correctness, consistency and completeness [BW84]. During the two weeks following the receipt of the survey responses, we randomly selected 44 responses to be validated (approx. 10% of all responses received). In this validation process, we aimed at three findings: (i) If the individual was aware of the questionnaire, and was involved in completing it. (ii) The understanding of the questions. (iii) Validating the provided data. The validation step concluded that no major variations were present, and thus increased our confidence in the research method and the collected data.

4.6 Research Validity

The issues of research instrument validation and validity threats have been rigorously validated. In summary, we present the following validity checks.

Internal validity raises the question of whether the observed effects could have been caused by, or correlated with, a set of unhypothesized and/or unmeasured variables. The design of the research instrument through the iterative process, which measured the effect of one variable on another, as well as the review, validation and pretest, which has resulted in more refinements of the research instrument, made strong evidence that the observed effects are strongly related to the hypothesized variables. Thus, the issue of internal validity has been met.

The evaluation and selection of the size of the organization and the degree of non-compliance to be the two independent variables, in measuring the effort required for implementing ISO 9000, has been identified to accurately reflect or measure the constructs of interest, thus confirming high degree of construct validity [Sch94], [Pea94], [CIS93], [Tic92], [RB93], and [ATT94].

The issue of external validity is dealt with extensively in this chapter, and in chapter 5. The survey covered the entire population of ISO 9001 software registered organization, whether small, medium or large. In addition the histograms that are presented in chapter 5, are an additional proof that the responses received were representative of the entire population. Thus, permitting us to generalize the research results, which would confirm a high degree of external validity.

The issue of statistical conclusion validity is dealt with in chapters 6 and 7. Statistical conclusion validity assesses the mathematical relationships between variables, incorrect conclusions concerning covariation, null hypothesis validation, etc. This has been the focus of the statistical analysis, model comparison techniques, Type I error checking, Type II error checking, and standard error. Thus the issue of statistical conclusion validity has been equally met.

In conclusion, we reiterate that the validity of research results, are dependent on the validity of the research method, research design, and instrument validation.

5 Sample Analysis

In this chapter we present a detailed analysis of the collected data, including an overview of the responses received, size of responding organizations, distribution of responses, and the degree of non-compliance.

5.1 Response Overview

The number of ISO 9001 software registered organizations in North America has nearly doubled in the past 18 months. At the time of the survey, there were approximately 63 ISO 9001 software registered organization in Canada and the United States [CIS95]. All 63 organizations were included in the survey. The number of responses received from surveyed software organization was 28 (44.4%), compared to the overall survey response of 462 (38.8%). This response is considered favorable compared to other ISO 9000 surveys [SPQ94], [Wol94].

The number of responses received from software organizations is considered statistically significant because it represents 44.4% of the population of ISO 9001 software registered organizations in North America.

5.2 Organization Size Profile

The distribution of responses received from ISO 9001 software organizations is of interest and importance, because, if responses were received only from small size organizations, then the model would not be useful to medium size or large organizations, and therefore can not be generalized. Similarly, if responses were received only from large organizations, then the model would be of little importance to small or medium size software organizations.

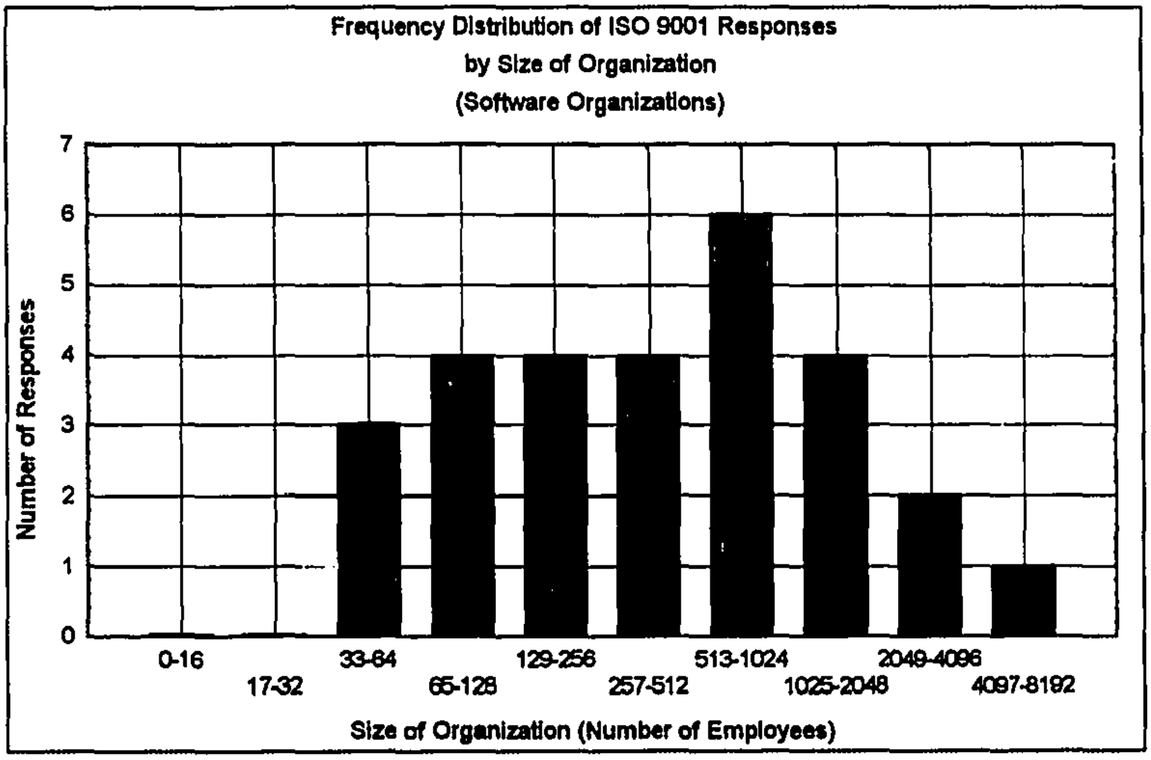


Figure 5-1: ISO 9001 (Software) responses by size of organization

The frequency distribution graph Figure 5-1, depicts the ISO 9001 responses from software organizations on a logarithmic scale; the use of logarithmic scale is not biased to any clustering method. Evidently the histogram is near a normal distribution, implying that the ISO 9001 responses received from software organizations are representative of the population, and are not biased towards small, medium, or large organizations.

Similarly, Figure 5-2, depicts the ISO 9001 responses received from all industries including software organizations on a logarithmic scale. Evidently the histogram follows a normal distribution, implying that the ISO 9001 responses received from all industries are also representative of the population, and are not biased towards small, medium, or large organizations [RM95].

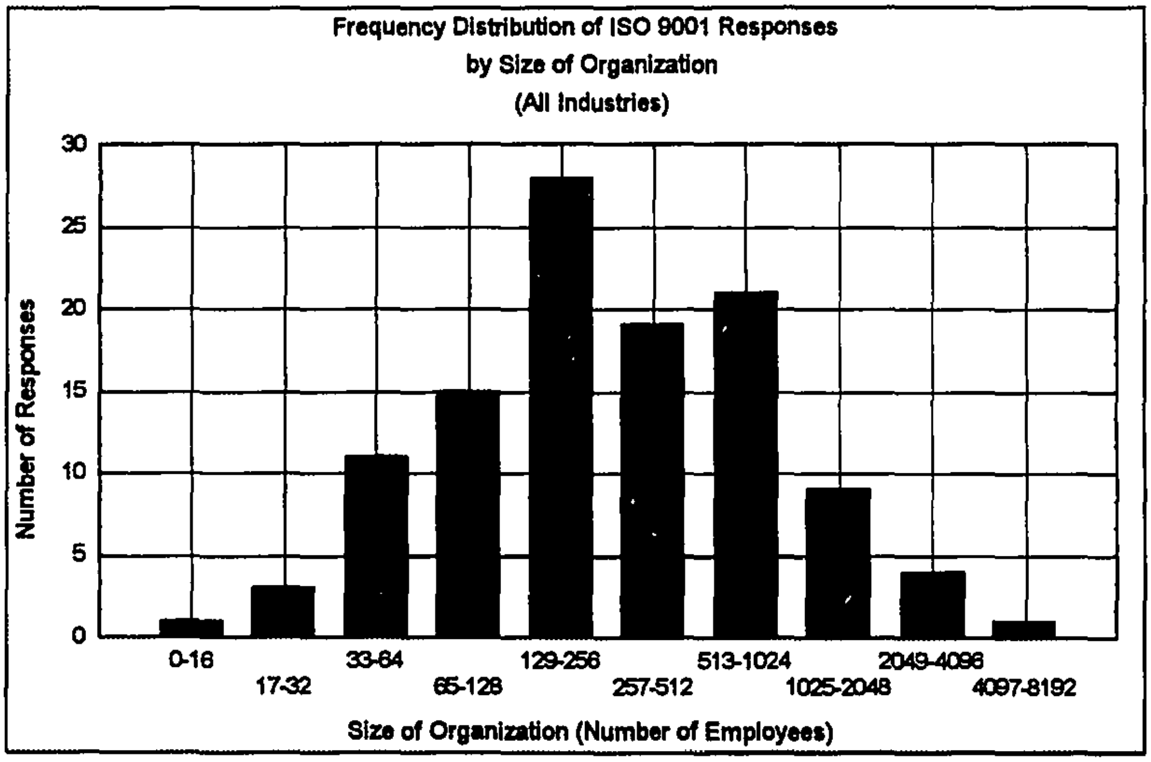


Figure 5-2: ISO 9001 (All Industries) responses by size of organization

5.3 Non-Compliance profile

The non-compliance profile of each organization was obtained from the survey questionnaire and equates to (100 - degree of compliance). This is a key element in the effort estimation model. The non-compliance profile indicates the status of the quality management system prior to the commencement of any ISO 9000 work.

The average degree of non-compliance of all responses received, provides an indication of the non-compliance level of the industry, or at least of the population of ISO 9001 registered software organizations. We shall also present a comparison between software and non-software organizations.

Figure 5-3 presents a plot of the degree of non-compliance of ISO 9001 *software* organization that responded to the survey in ascending order by size of organization.

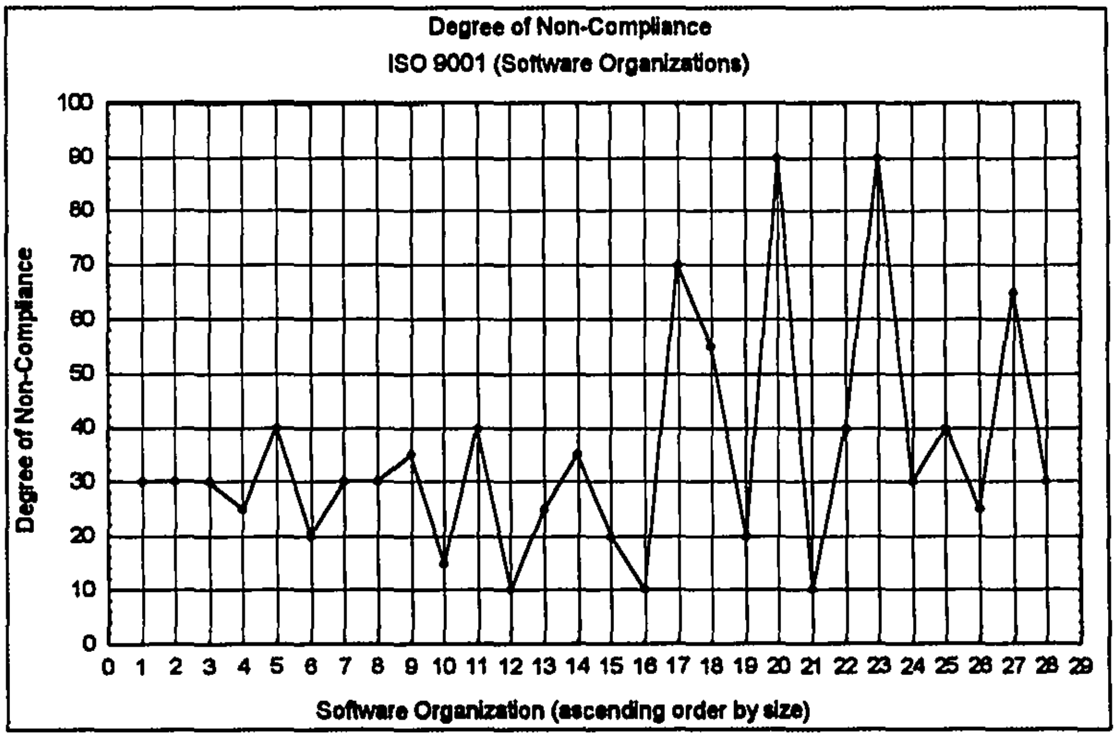


Figure 5-3: Degree of Non-Compliance ISO 9001 (Software Organizations)

In keeping with the objective of this research, we do not make any conclusions on the evidence presented in Figure 5-3, which may suggest that small size organizations tend to have a lesser degree of non-compliance than large organizations.

Similarly, Figure 5-4, presents a plot of the degree of non-compliance of the eighty-four ISO 9001 *non-software* organization that responded to the survey, in ascending order by size of organization. This plot does not show any difference between small or large organization.

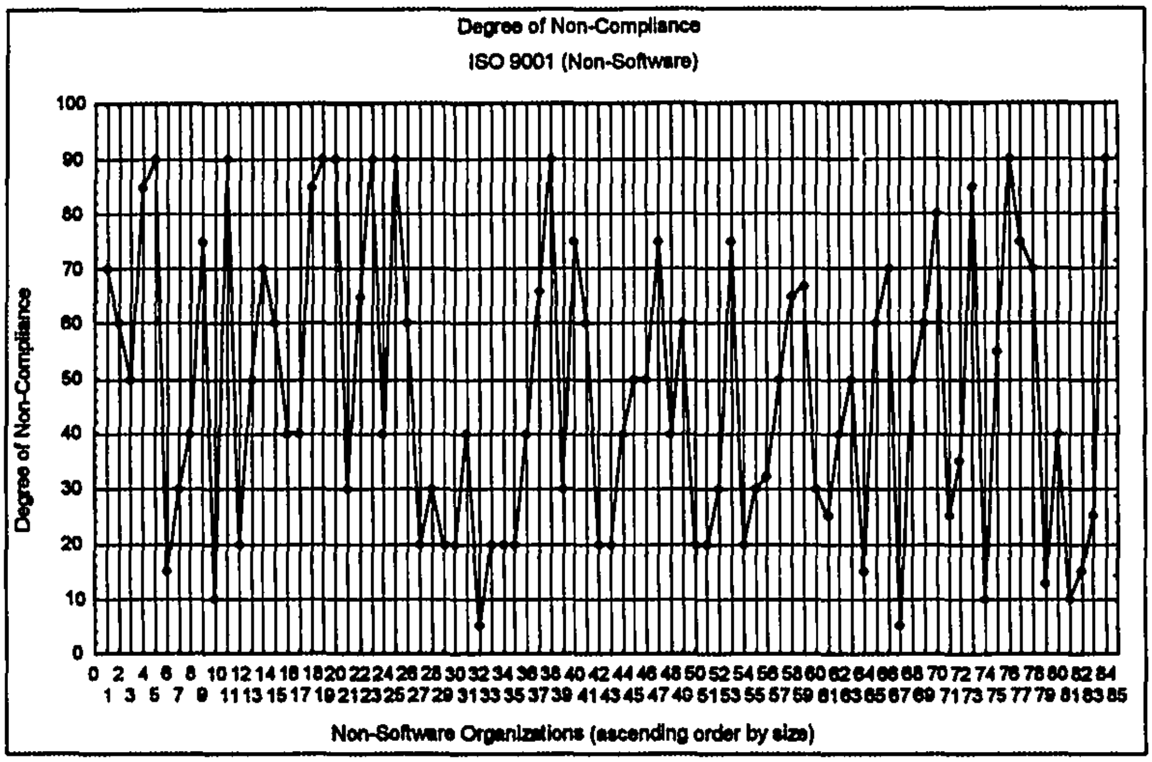


Figure 5-4: Degree of Non-Compliance ISO 9001 (Non-Software)

5.4 Compliance by Industry

This section presents two plots that contrast software organizations versus non-software organizations. First, Figure 5-5, presents the average degree of non-compliance of the ISO 9001 software registered organizations versus the non-software ISO 9001 registered organizations. The average degree of non-compliance was calculated as:

$$\frac{\sum(\text{Degree of NonCompliance})}{\text{Total number of organizations}}$$

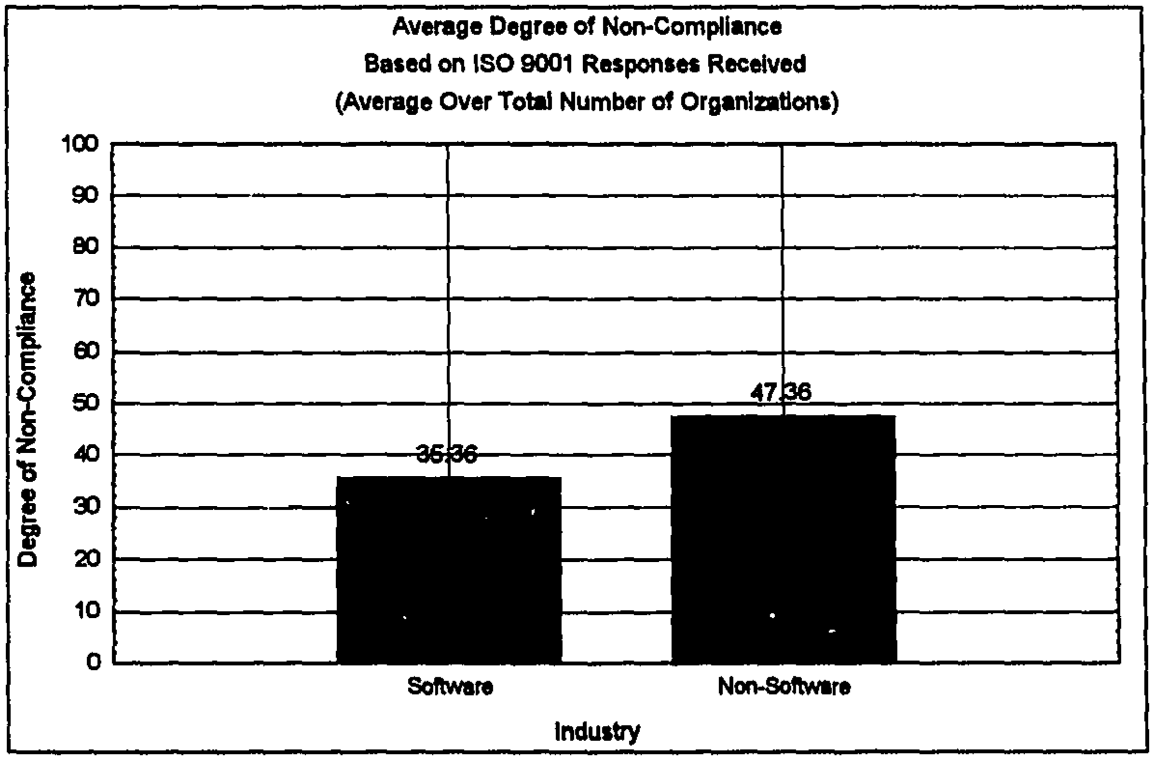


Figure 5-5: Average Degree of Non-Compliance (Number of Organizations)

This graph would suggest (not conclusive) that software organizations in general, have lesser degree of non-compliance than non-software (i.e. better).

Other analysis of the degree of non-compliance were performed, and contrasted with non-software industries. We present a similar graph in Figure 5-6, of the average degree of non-compliance of ISO 9001 software registered organizations versus the non-software ISO 9001 registered organizations surveyed. In this graph, the average degree of non-compliance was calculated using the size of the organizations (number of employees):

$$\frac{\Sigma(\text{ Degree of NonCompliance})}{\Sigma(\text{ Size of Organizations})}$$

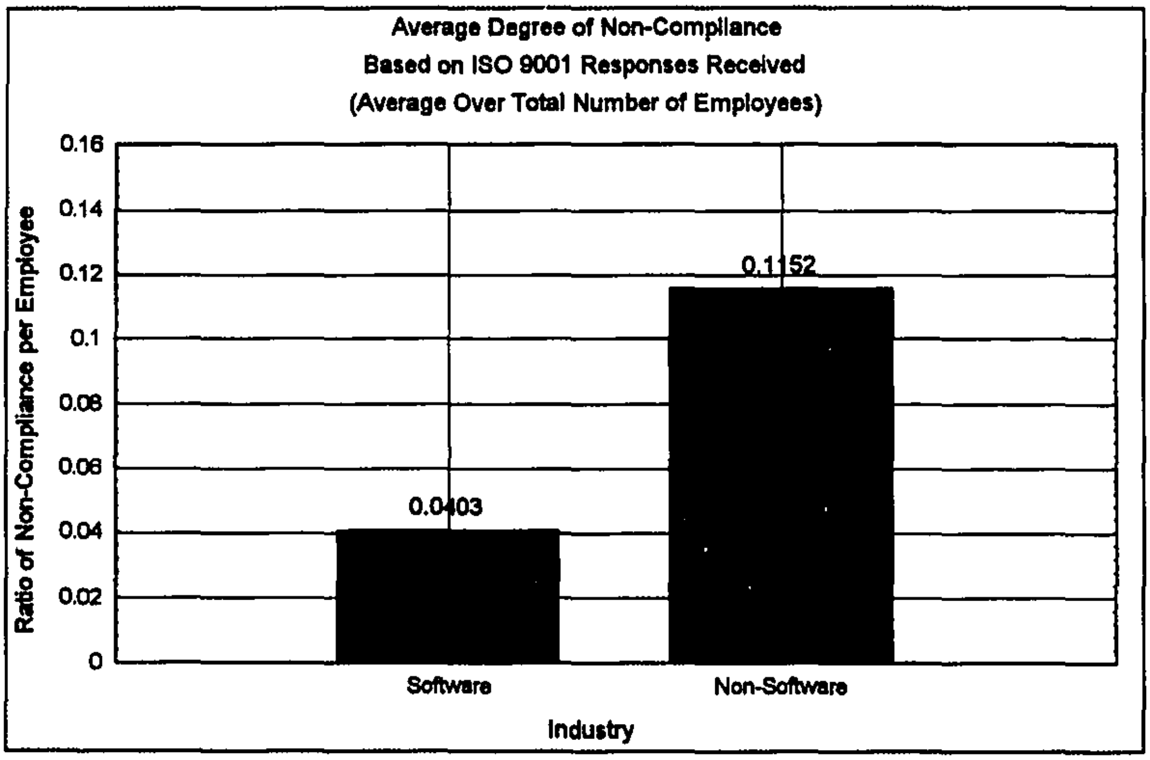


Figure 5-6: Average Degree of Non-Compliance (Size of Organizations)

The data shows that software organizations have a lesser degree of non-compliance per employee, than other industries (based on the data obtained from the survey).

5.5 Summary

The sample data analysis presented in this chapter indicates that responses received from software organizations, are representative of the population of ISO 9001 software registered organizations in North America. Similarly, the degree of non-compliance is not static, but different organizations have different degrees of non-compliance. Therefore, the effort estimation model that will be presented will be useful to small, medium and large size software organizations.

Finally, the data collected from software and non-software organizations, showed that software organizations in general, were more compliant to the ISO 9001 standard prior to registration than non-software organizations.

6 Statistical Approach

In this chapter, we describe the statistical approach for building the effort estimation model, the model assumptions, the regression data variables, the regression principle and the potential seven different regression models for estimating the effort of ISO 9001 implementation in a software organization.

The science of statistics involves making inferences (draw conclusions) about a population based on the information contained in a sample drawn from that population. A statistical methodology that relates variables, called regression analysis, studies the effect of independent variable(s) upon a dependent variable; the objective of the regression analysis is to build a regression model or prediction equation.

In building a regression model, a number of concerns must be addressed [BO90], [MBM94], [Ped82], [SPA81]. The first is model adequacy, or explanatory power of the independent variable(s) in accounting for the variability of the dependent variable. This is typically measured by $\mathbf{R}^2$ (the multiple coefficient of determination). The larger the $\mathbf{R}^2$, the larger the proportion of the total variation that is explained by the regression model. However, a large value of $\mathbf{R}^2$ is not the only measure of a good model, because the size of $\mathbf{R}^2$ does not necessarily indicate whether a regression model will accurately describe, predict, and control the dependent variable. Other criteria that can be used to measure the model aptness include, MSE (the mean square error), s (standard error of estimate), lengths of the prediction intervals, C statistic and PRESS statistic. In addition, the validity of the regression model is dependent on certain inference assumptions.

6.1 Model Assumptions

The theory of estimation and the validity of linear regression models are dependent on certain assumptions about the distribution and behavior of the residuals or error terms. If these are seriously violated, a large $\mathbf{R}^2$ may be of little importance.

These assumptions, which are referred to as inference assumptions are: (1) Homoscedasticity of the error terms (that is, constant variance around the regression line); (2) Independence; and (3) Normality [BO90], [Ped82].

These assumptions are typically verified with the aid of diagnostic plots. The most common plots are the normal probability plot, and the residual plot. The residual plot can be used to assess the validity of the *homoscedasticity* assumption. A residual plot that "fans out", indicates that the error terms are increasing in absolute value as the horizontal plot criterion increases. A residual plot that "funnels in", suggests that the error terms are decreasing in absolute value as the horizontal plot criterion increases. However, a residual plot with a "horizontal band appearance", indicates that the error terms are relatively constant in absolute value. Whenever the error terms increases "fans out" or decreases "funnels in" the residual plots suggest that the homoscedasticity assumption is violated, and we would say that a nonconstant error variance exists. When this and other assumptions are violated, transformation of variables is often attempted as a remedy.

Validation of the *independence* assumption can be employed using the Durbin-Watson test of the error term. The independence assumption holds if the error terms are independent and display a random pattern. The independence assumption requires that any one value of the dependent variable y be *statistically independent* of any other value of y.

Validation of the *normality* assumption can be accomplished using histograms of the residuals, chi-square test, Kolmogorov-Smirnov test, or normal probability plot. The normal probability plot is a good graphical technique for examining the validity of the normality assumption. If the normality assumption holds, and if the model has the correct functional form, then this plot should have a straight-line appearance. Although a mild departure from the inference assumption does not seriously hinder the ability to make statistical inferences [BO90], [SPA81], a substantial departure from a straight-line appearance is a violation of the normality assumption.

6.2 Regression Data

The data employed in a regression analysis can be *observational* or *experimental*. In addition, the data can be *time series data* or can be *cross-sectional data*. Data is called *observational* when the values of the different independent variables $x_1, x_2, \ldots, x_n$ cannot be controlled nor set, but can be measured without error. If the values of the different independent variables are set before the experiment, and then the values of the dependent variables are observed, we say that the data is *experimental*. When data is observed in time order, we call it *time series data*. Time series data often consists of yearly, quarterly, or monthly observations, although any other time period may be used. When data is observed at one point in time, the data is called *cross-sectional*.

The regression data employed in this research is *Observational* and *Cross-sectional*, the two independent variables that are used in the prediction of the effort required to implement ISO 9001 in an organization are:

1. The size of the organization (number of employees).
2. The degree of non-compliance (percentage of the ISO 9001 requirements that the organization still needs to satisfy to become fully compliant).

We begin by looking at regression models that employ a single independent variable.

6.3 Simple Regression

A simple linear (or straight line) regression model employs two parameters: *the slope* and *the intercept*. We can compute the "least squares point estimates" of these parameters, and use these estimates to compute point predictions of the dependent variable.

If we assume that our model has only one independent variable, namely the size of the organization, and the dependent variable is the effort required to implement ISO 9001 in an organization, then the simple linear regression model would be of the form:

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$$

Equation 6-1: Simple Linear Regression Model

The y-intercept β_0 is the mean value of the dependent variable when the value of the independent variable is zero.

The slope β_1 is the change in the mean value of the dependent variable that is associated with a one-unit change in the value of the independent variable. If β_1 is positive, the mean value of the dependent variable increases as the value of the independent variable increases, and If β_1 is negative, the mean value of the dependent variable decreases as the value of the independent variable increases.

The error term ε_i, describes the effects on y_i of all factors other than the value of x_i of the independent variable x.

x_i = size of organization$_i$

y_i = effort spent by organization$_i$

Although the size of the organization is an important variable in predicting the required effort for implementing ISO 9000, we nevertheless have argued that the effort required by an organization to reach the goal of ISO 9000 registration may also be dependent on the status of the organization's quality management system prior to their commencement on this initiative. Therefore, an alternate regression model that employs two independent variables may be needed.

6.4 Multiple Regression

Regression models that employ more than one independent variable are called *multiple regression models.* Similar to simple linear regression models, we can compute the "least squares point estimates" of these parameters, and use these estimates to compute point predictions of the dependent variable.

Our most basic multiple regression model would have two independent variables, namely the size of the organization, and the degree of non-compliance. The dependent variable is the effort required to implement ISO 9001 in a software organization, this would lead us to a model of the form:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \varepsilon_1$$

Equation 6-2: Multiple Linear Regression Model

In Figure 6-1, we present a plot of y (effort in months) versus x_1 (size of the organization). This plot indicates that there is a strong positive linear correlation between the number of employees in an organization and the effort it takes to achieve ISO 9001 registration. Similarly, in Figure 6-2, we present a plot of y (effort in months) versus x_2 (degree of non-compliance). This plot indicates that there is some positive linear correlation between the degree of non-compliance and the effort it takes to achieve ISO 9001 registration in a software organization. Therefore, as the number of employees increases, so does the effort required. Similarly, as the degree of non-compliance increases, so does the effort required. Thus, there is positive and linear correlation between x_1 and y, and between x_2 and y.

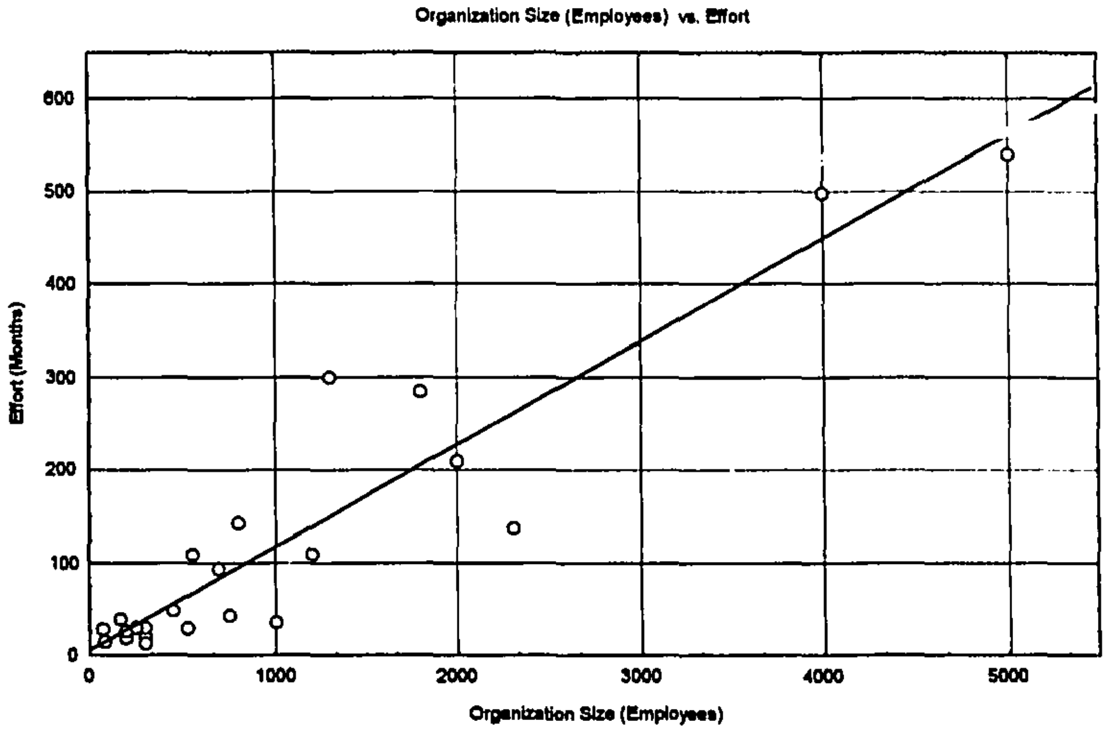


Figure 6-1: Organization Size versus Effort (Software)

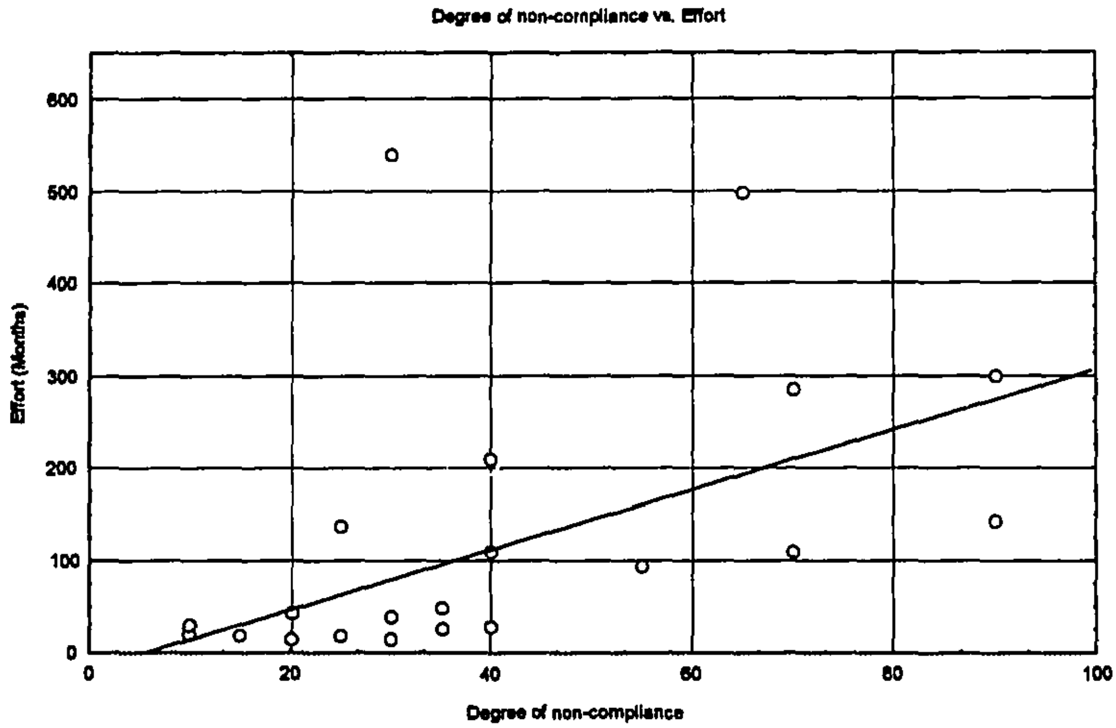


Figure 6-2: Degree of Non-Compliance versus Effort

6.4.1 Variable Interaction

Multiple regression models often contain *interaction variables*. We say that there is *interaction* between two independent variables if the relationship between the mean value of the dependent variable and one of the independent variables is dependent upon the value of another independent variable. This appears to be true in our case since the relationship between the mean value of the dependent variable y (effort) and x_1 (Size of the organization) is also dependent on x_2 (degree of non-compliance). Thus, if we were to plot y versus x_1 for different levels of x_2 (i.e. $x_2 = 10, 40$, etc.), we can see that the slope of the straight line when $x_2=10$ (degree of non-compliance at 10%), is smaller than the slope of the straight line when $x_2=40$ (degree of non-compliance at 40%). Thus, implying that the higher the degree of non-compliance for an organization from the requirement of the ISO 9001 standard, the higher the effort required to implement the ISO 9001 requirements.

Therefore, because of the above described interaction between the independent variables x_1 and x_2, it may be necessary to use a model that includes the *cross-product* term x_1x_2 [BO90]. However, the *cross-product* term contributes somewhat overlapping information for describing the dependent variable, thus requiring special care. A regression model that employs x_1 , x_2 , and x_1x_2, the *cross-product* term is shown below.

$$y = \beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_3x_1x_2 + \varepsilon_i$$

Equation 6-3: Multiple regression model with cross-product term.

6.4.2 Multicollinearity

Multicollinearity is said to exist when the independent variables in a regression model are interrelated or are dependent on each other. The consequence of multicollinearity in a regression analysis vary from slight to quite substantial, depending on the extent of the multicollinearity. Extreme cases of multicollinearity can cause the least squares point estimate to be far from the true values of the regression parameters. In fact, some of the least squares point estimates may have the wrong sign.

Perhaps the most serious effect of multicollinearity is that it can hinder our ability to assess the importance of an independent variable, and often causes the t statistics obtained to be smaller (in absolute value), thus causing some of the correlated independent variables to appear as less important.

The best solution to the multicollinearity problem is by eliminating redundant independent variables in the regression model [BO90]. However, we shall delay this approach, and retain the model that includes the *cross-product* term in the list of potential regression models, for further analysis and comparison.

6.5 Potential Effort Estimation Models

We have presented three forms of potential linear regression models. The first form uses a single independent variable, the second uses two independent variables, and the third form uses two independent variables and a cross-product term.

Table 6-1, presents the seven potential regression models, and the equations of the *best fitting*, or, *Least-Squares Line* that minimizes the *sum of squared residuals* ***SSE***, in the ISO 9001 data for software organizations.

In the following chapter, we compare the seven regression models using several statistical techniques to determine the most appropriate model for predicting the effort required for implementing ISO 9001 in software organizations.

No.	Form	Model
1	$y = \beta_0 + \beta_1 x_1 + \varepsilon$	$y = 3.1670 + .106875x_1$
2	$y = \beta_0 + \beta_2 x_2 + \varepsilon$	$y = -7.9415 + 2.96501x_2$
3	$y = \beta_0 + \beta_3 x_1 x_2 + \varepsilon$	$y = 15.8526 + .00224x_1x_2$
4	$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon$	$y = -53.1508 + .1003x_1 + 1.7553x_2$
5	$y = \beta_0 + \beta_1 x_1 + \beta_3 x_1 x_2 + \varepsilon$	$y = 4.0992 + .053059x_1 + .001277x_1x_2$
6	$y = \beta_0 + \beta_2 x_2 + \beta_3 x_1 x_2 + \varepsilon$	$y = 30.4514 - .51790x_2 + .00234x_1x_2$
7	$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1 x_2 + \varepsilon$	$y = -28.9425 + .0728x_1 + 1.0172x_2 + .0007x_1x_2$

x_1: Size of organization (Number of employees)

x_2: Degree of non-compliance

x_1x_2: (Size of organization * Degree of non-compliance)

Table 6-1: The seven potential ISO 9001 regression models

7 Research Results

This chapter compares the seven potential regression models presented in chapter 6, selects the best model, validates the inference assumptions, applies data transformation, and validates the final model. The criteria that are used to compare the usefulness of the regression models include: the multiple coefficient of determination, R^2; the mean square error, S^2 (and standard error, s); length of prediction intervals; corrected (or adjusted) R^2; the C statistic; and the ***PRESS*** statistic.

7.1 Model comparison

The first criterion we employed in comparing the regression models is R^2, the multiple coefficient of determination. The larger the R^2, the larger is the proportion of the total variation that is explained by the regression model. However, we must balance the magnitude of R^2, or in general the "goodness" of any criterion, with other factors such as the length of the prediction interval. A second criterion that we employed in comparing the various regression models was s, the length of the prediction interval. The smaller the prediction interval, the more precise the prediction, thus a shorter prediction interval is more desirable. This prediction interval, or standard error is a better criterion than R^2 for comparing regression models because adding any independent variable to a regression model will increase R^2 but adding an "unimportant" independent variable can increase s.

The SAS Statistical package SAS/STAT Version 6, by SAS Institute Inc., and STATISTICA for Windows Release 4.3B by StatSoft Inc. were used for the advance statistical and regression analysis of this research.

We present in Table 7-1, the seven potential regression models from chapter 6, which were built using different combination of the two independent variables x_1 and x_2 and the cross-product term x_1x_2.

Model No.	Model
1	$y = 3.1670 + .106875x_1$
2	$y = -7.9415 + 2.96501x_2$
3	$y = 15.8526 + .00224x_1x_2$
4	$y = -53.1508 + .1003x_1 + 1.7553x_2$
5	$y = 4.0992 + .053059x_1 + .001277x_1x_2$
6	$y = 30.4514 + -.51790x_2 + .00234x_1x_2$
7	$y = -28.9425 + .0728x_1 + 1.0172x_2 + .0007x_1x_2$

Table 7-1: The potential seven ISO 9001 regression models.

In comparing the seven regression models, we calculated the SST, total variation; SSE, the unexplained variation; SSM, the explained variation; df, the degrees of freedom (number of observations - number of parameters); R^2, coefficient of determination; S^2, mean square error; s, standard error of estimate; F(Model), the overall F-Statistic of the model; and the P-level of significance. The calculated value of each regression model is shown in Table 7-2.

No.	Model	SST	SSE	SSM	df	R^2	S^2	s	F	P
1	$y = \beta_0 + \beta_1x_1 + \varepsilon$	516422	70117	446305	26	.86422	2696	51.93	165	.00000
2	$y = \beta_0 + \beta_2x_2 + \varepsilon$	516422	410079	106343	26	.20592	15772	125.58	6.7	.01528
3	$y = \beta_0 + \beta_3x_1x_2 + \varepsilon$	516422	60991	455431	26	.88189	2345	48.43	194	.00000
4	$y = \beta_0 + \beta_1x_1 + \beta_2x_2 + \varepsilon$	516422	34522	481900	25	.93315	1380	37.16	174	.00000
5	$y = \beta_0 + \beta_1x_1 + \beta_3x_1x_2 + \varepsilon$	516422	34863	481559	25	.93249	1394	37.34	172	.00000
6	$y = \beta_0 + \beta_2x_2 + \beta_3x_1x_2 + \varepsilon$	516422	58703	457719	25	.88632	2348	48.45	97	.00000
7	$y = \beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_3x_1x_2 + \varepsilon$	516422	29645	486777	24	.94259	1235	35.14	131	.00000

Table 7-2: R^2, s^2, s, of the seven ISO 9001 regression models

A review of each model is given below.

Model 1: $(y = \beta_0 + \beta_1 x_1 + \varepsilon)$

This model uses a single independent variable (size of the organization), and does not account for the degree of non-compliance in an organization. Thus, an organization that has a well defined and documented software development process and procedures that is near conformance to the ISO 9001 standard, is treated equally to an organization that has an ad-hoc and undefined software development process. Also, the standard error of estimate *s*, is relatively high, thus yielding a large prediction interval. Therefore, it is not considered a good regression model, and thus rejected.

Model 2: $(y = \beta_0 + \beta_2 x_2 + \varepsilon)$

This model also uses a single independent variable (degree of non-compliance), and does not account for the size of the organization in predicting the effort required to implement ISO 9001 in an organization. Therefore, it is not considered a good regression model, and thus rejected.

Model 3: $(y = \beta_0 + \beta_3 x_1 x_2 + \varepsilon)$

This model uses the cross-product term (size of organization * degree of non-compliance), and has a better R^2, than the previous two models, and also has a lower standard error of estimate, however, the proportion of the total variation in the observed values of the dependent variable are not well explained and uses a single β_1, slope value. Therefore, it is not considered a good regression model, and thus rejected.

Model 4: $(y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon)$

This model uses the two independent variables, and has a high R^2, thus a high proportion of the total variation in the observed values of the dependent variable are well explained.

Also, the model has a relatively low standard error of estimate. Therefore, we consider this model to be a potential final model.

Model 5: $(y = \beta_0 + \beta_1 x_1 + \beta_3 x_1 x_2 + \varepsilon)$

This model uses one independent variable, and the cross-product term, it is very similar in its coefficient of determination, R^2, and standard error of estimate, s, to model no. 4. However, this model requires further analysis of *multicollinearity*. Multicollinearity is said to exist among the dependent variables, when the independent variables in a regression model are interrelated or are dependent on each other. Therefore, we shall consider (for now) this model in the list of potential final models.

Model 6: $(y = \beta_0 + \beta_2 x_2 + \beta_3 x_1 x_2 + \varepsilon)$

This model uses one independent variable, and the cross-product term, is similar to model no. 3. Therefore, it is not considered a good regression model, and thus rejected.

Model 7: $(y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1 x_2 + \varepsilon)$

This model uses the two independent variables, and the cross-product term, and has a high R^2, thus a high proportion of the total variation in the observed values of the dependent variable is well explained. Also, the model has the lowest standard error of estimate. However, this model requires further analysis of *multicollinearity*, this is similar to model no. 5. Therefore, we shall consider (for now) this model in the list of potential final models.

Further analysis and comparison of the above models were performed to achieve higher confidence in the selection of the final model. Thus we have used advance model comparison methods using Corrected (Adjusted) R^2, C statistic, and ***PRESS*** statistic.

No.	Model	Corrected R^2	C statistic	$PRESS$ statistic
1	$y = \beta_0 + \beta_1 x_1 + \varepsilon$	.85900	32.76	82347
2	$y = \beta_0 + \beta_2 x_2 + \varepsilon$	.17538	307.98	475835
3	$y = \beta_0 + \beta_3 x_1 x_2 + \varepsilon$	.87735	25.37	127692
4	$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon$	.92780	5.94	51337
5	$y = \beta_0 + \beta_1 x_1 + \beta_3 x_1 x_2 + \varepsilon$	.92709	6.22	106653
6	$y = \beta_0 + \beta_2 x_2 + \beta_3 x_1 x_2 + \varepsilon$	.87723	25.52	153557
7	$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1 x_2 + \varepsilon$	.93541	4	99009

Table 7-3: Corrected R^2, C and PRESS Statistic of the seven models.

The advance regression model comparison procedures used eliminated model no. 5 because it had a large ***C*** and ***PRESS*** statistic, thus we were left with model no. 4 and model no. 7. The ***C*** statistic which was obtained in model no. 7 equals exactly the number of parameters in the model, and the ***C*** statistic that was obtained in model no. 4 was relatively close to the number of parameters in the model. Statistical theory states that it is desirable to have a model with a value of ***C*** statistic equal or close to the number of parameters in the regression model.

Since model 7 has the smallest ***C*** statistic and the smallest standard error of estimate, ***s***, we might be tempted to conclude that this model is the best model to use to predict the effort required to implement ISO 9001 in an organization. However, having considered the ***PRESS*** statistic for each model, the results suggest that model no. 4, is probably as good in predicting the effort required to implement ISO 9001 in an organization, and does not possess the threats of multicollinearity. Thus we shall conclude with the selection of model no. 4 as the final regression model.

As a final step, and before validating the inference assumptions of the regression model, we shall present an example using the predictive variations of all seven models. This is presented in Table 7-4.

No.	Model	Size of organization (Employees)	Degree of non-compliance %	Predicted Effort (months) Rounded
1	$y = \beta_0 + \beta_1 x_1 + \varepsilon$	250	40	30
2	$y = \beta_0 + \beta_2 x_2 + \varepsilon$	250	40	111
3	$y = \beta_0 + \beta_3 x_1 x_2 + \varepsilon$	250	40	38
4	$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon$	250	40	42
5	$y = \beta_0 + \beta_1 x_1 + \beta_3 x_1 x_2 + \varepsilon$	250	40	30
6	$y = \beta_0 + \beta_2 x_2 + \beta_3 x_1 x_2 + \varepsilon$	250	40	33
7	$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1 x_2 + \varepsilon$	250	40	37

Table 7-4: Effort prediction of each regression model

As can be seen in Table 7-4, the predicted value of model no. 2 is completely outside the range of the predictions provided by the other models. This was also reflected by the low R^2 value of this model (Table 7-2).

Having compared all seven models, we shall focus our attention on the validation of model no. 4, which was determined based on the statistical comparison procedures we employed, as the best and most desirable model.

7.2 Assumption Validation

The validity of the regression model, regression results, and statistical hypothesis depend on the three assumptions that we refer to as the *inference assumptions* and have discussed in Chapter 6.

We shall start with the first inference assumption, homoscedasticity or constant variance of the residual. To construct a residual plot, we compute the residual for each observation. Thus we compute $e_i = y_i - \hat{y}_i$ for $y_1, y_2, \ldots, y_n$, where y_i is the observed value, and $\acute{y}_i$ is the predicted value. Then we plotted the values of the predicted variables in the model against the residual scores.

Figure 7-1 shows the predicted versus the residual scores plot of the ISO 9001 data, using regression model no. 4. The residual plot seems to suggest that the error terms are increasing in absolute value as the predicted values increases. Thus suggesting that the homoscedasticity (constant variance) assumption is slightly departed from the norm and violated.

In order to remedy the suggested violation of the regression assumption, and the increasing error variance that casts some doubts on the assumption of homoscedasticity of the residual, we need to apply established statistical transformation techniques on the ISO 9001 data. However, before applying those remedies we will validate the remaining two inference assumptions.

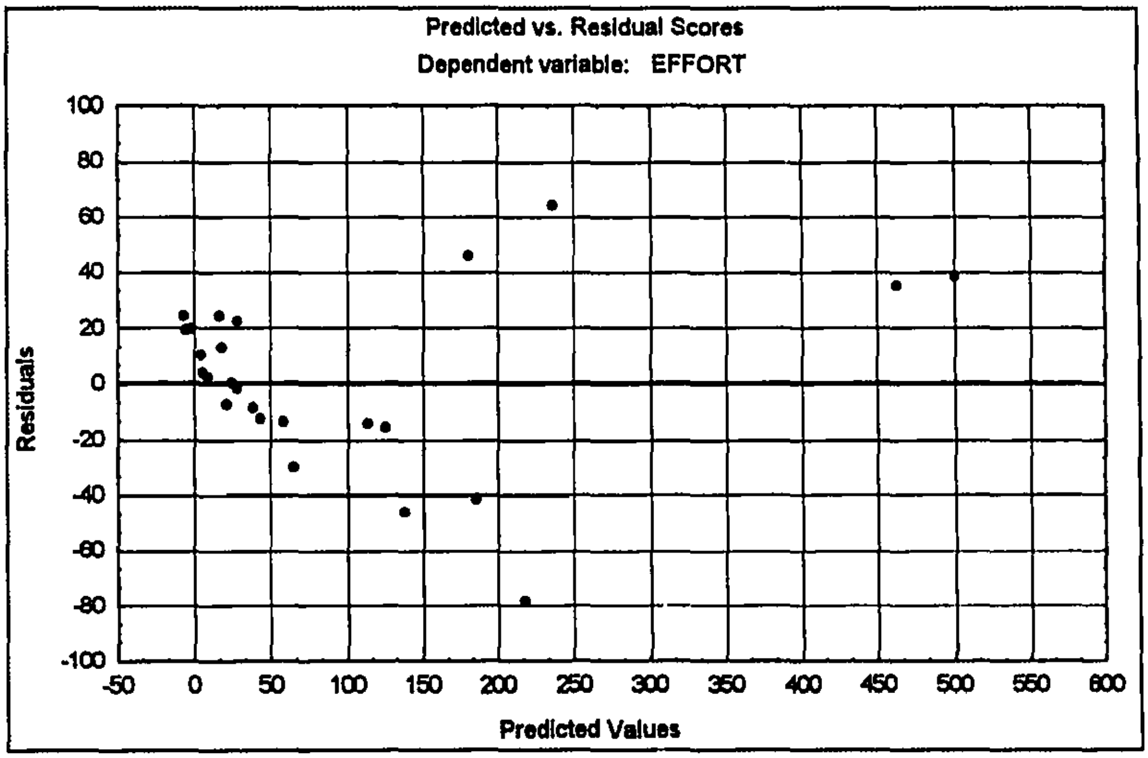


Figure 7-1: Predicted versus residual score plot (Model no. 4)

The second inference assumption we need to validate is the independence assumption. A common validation technique that is employed uses the Durbin-Watson test of the error term. The independence assumption holds if the error terms are independent and display a random pattern. The independence assumption requires that any one value of the dependent variable *y* be *statistically independent* of any other value of *y*.

We will employ the Durbin-Watson test for (first-order) positive and negative autocorrelation. Therefore, we start testing the null hypothesis for positive autocorrelation

H_0 : The error terms are not autocorrelated

versus the alternative hypothesis

H_1 : The error terms are positively autocorrelated

Durbin and Watson have shown that, based on setting the probability of a Type I error equal to α,

1. If $d < d_{L,\alpha}$ we reject H_0.
2. If $d > d_{U,\alpha}$ we do not reject H_0.
3. If $d_{L,\alpha} \leq d \leq d_{U,\alpha}$ the test is inconclusive.

We computed the Durbin-Watson statistic using SAS and STATISTICA, and looked up the critical values for the Durbin-Watson d statistic (α = .05), and we obtained the following values:

Model	Durbin Watson d	$d_{L,\alpha/2}$	$d_{U,\alpha/2}$	α	Autocorrelation
no. 4	1.153020	1.26	1.56	.05	YES

Table 7-5: Durbin-Watson d statistic for positive autocorrelation

This test showed that $d < d_{L,\alpha}$ is true, therefore, we reject H_0 in favor H_1, and conclude that the error terms are positively autocorrelated.

Similarly, we test the hypothesis for negative autocorrelation.

H_0 : The error terms are not autocorrelated

versus the alternative hypothesis

H_1 : The error terms are negatively autocorrelated

Durbin and Watson have shown that, based on setting the probability of a Type I error equal to α,

1. if $(4 - d) < d_{L,\alpha}$ we reject H_0.
2. if $(4 - d) > d_{U,\alpha}$ we do not reject H_0.
3. If $d_{L,\alpha} \leq (4 - d) \leq d_{U,\alpha}$ the test is inconclusive.

We computed the Durbin-Watson statistic using SAS and STATISTICA, and looked up the critical values for the Durbin-Watson d statistic (α = .05), and we obtained the following values:

Model	Durbin-Watson d	$d_{L,\alpha/2}$	$d_{U,\alpha/2}$	α	Autocorrelation
no. 4	1.153020	1.26	1.56	.05	None

Table 7-6: Durbin-Watson d statistic for negative autocorrelation

The Durbin-Watson statistic suggests that there is no negative autocorrelation in model no. 4, but there is a chance of positive autocorrelation. Therefore, the inference assumption of independence is equally violated.

The last inference assumption that we need to validate is the normality assumption. We shall use the normal probability plot to validate this assumption. The normality assumption holds if the plot has a straight-line appearance. Substantial departure from a straight-line appearance may indicate a violation of the normality assumption.

Figure 7-2 shows the normal probability plot of the ISO 9001 data, using regression model no. 4. The normal probability plot seems to suggest a slight departure from the desired straight-line appearance..

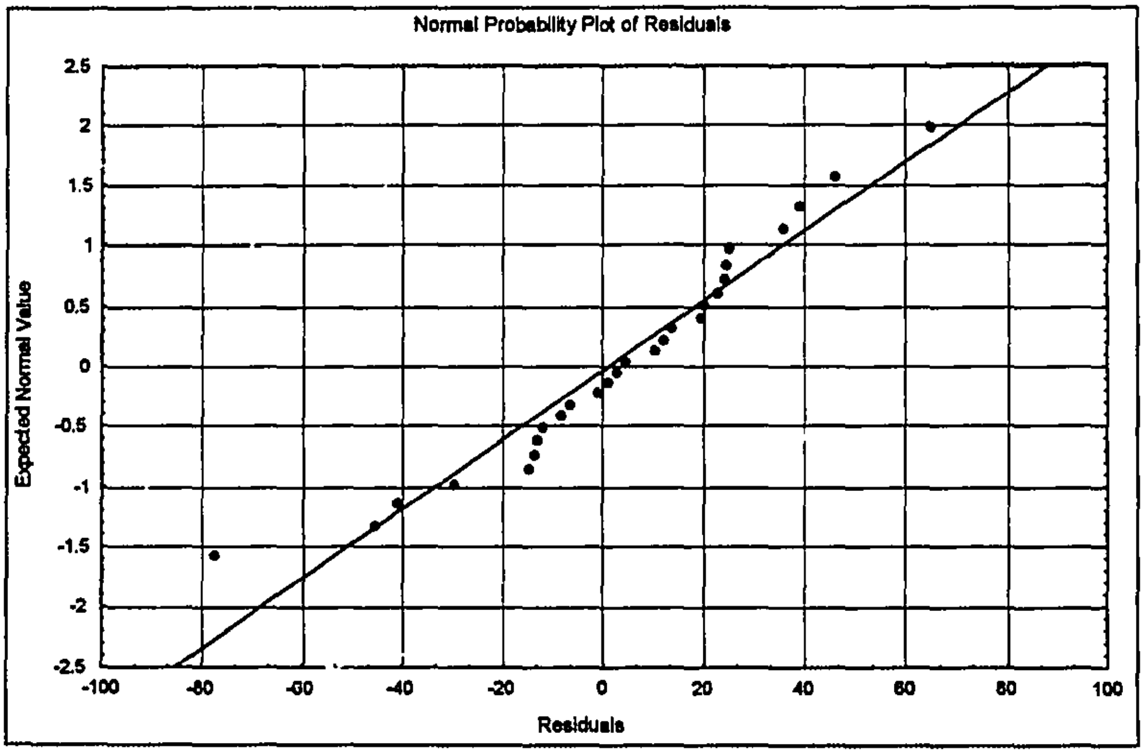


Figure 7-2: Normal probability plot of residuals (Model no. 4)

Having validated the inference assumption using the residual plot, normal probability plot and Watson-Durbin test of independence, it became evident that the homoscedasticity, independence and normality assumptions are all violated. To remedy the violations of the regression, data transformation is suggested [MBM94], [BO90], [Ped82].

7.3 Data Transformation

Data transformation techniques are applied as remedies when the inference assumptions are violated, and specially when the homoscedasticity and normality assumptions are violated. As we have seen in section 7.2 - Assumption Validation, the residual plot seemed to "fan-out", and the independence assumption showed positive autocorrelation.

We have thus attempted the logarithmic transformation to the independent and dependent variables of the ISO 9001 data [MBM94], [Ped82]. This transformation is particularly useful in that linearity of the original relationship is preserved while stabilizing the error variance.

Having applied the data transformation, we used the same criteria as before, and started with seven new regression models. We combined several procedures and techniques to arrive at the final model. The criteria we used to compare the usefulness of the overall regression model include: the multiple coefficient of determination, R^2; the mean square error, S^2 (and standard error, *s*); length of prediction intervals; corrected (or adjusted) R^2; the *C* statistic; and the ***PRESS*** statistic.

Similar to the previous model comparison, the analysis carried on the transformed data indicated that model no. 4 was the best model. The final model uses two independent variables, size of organization, and degree of non-compliance. It has an R^2 of .908, thus a high proportion of the total variation in the observed values of the dependent variable are well explained. Also, the model has a low standard error of estimate of .367, a *P*-level of .000000, a *C* statistic of 2.0257, and a ***PRESS*** statistic of 4.10422. The final model is based on the natural logarithm and is of the form:

$$\mathrm{Ln}\,(y) = \beta_0 + \beta_1 \mathrm{Ln}\,(x_1) + \beta_2 x \, \mathrm{Ln}\,(x_2) + \varepsilon.$$

7.4 Effort Estimation Model

The **Effort Estimation Model for Implementing ISO 9001 in Software Organizations** is presented in Equation 7-1, and the predicted versus observed values graph is in Figure 7-3.

$$Ln(y) = -2.793 + .692 * Ln(x1) + .74 * Ln(x2)$$

Equation 7-1: Effort estimation model for implementing ISO 9001 for software

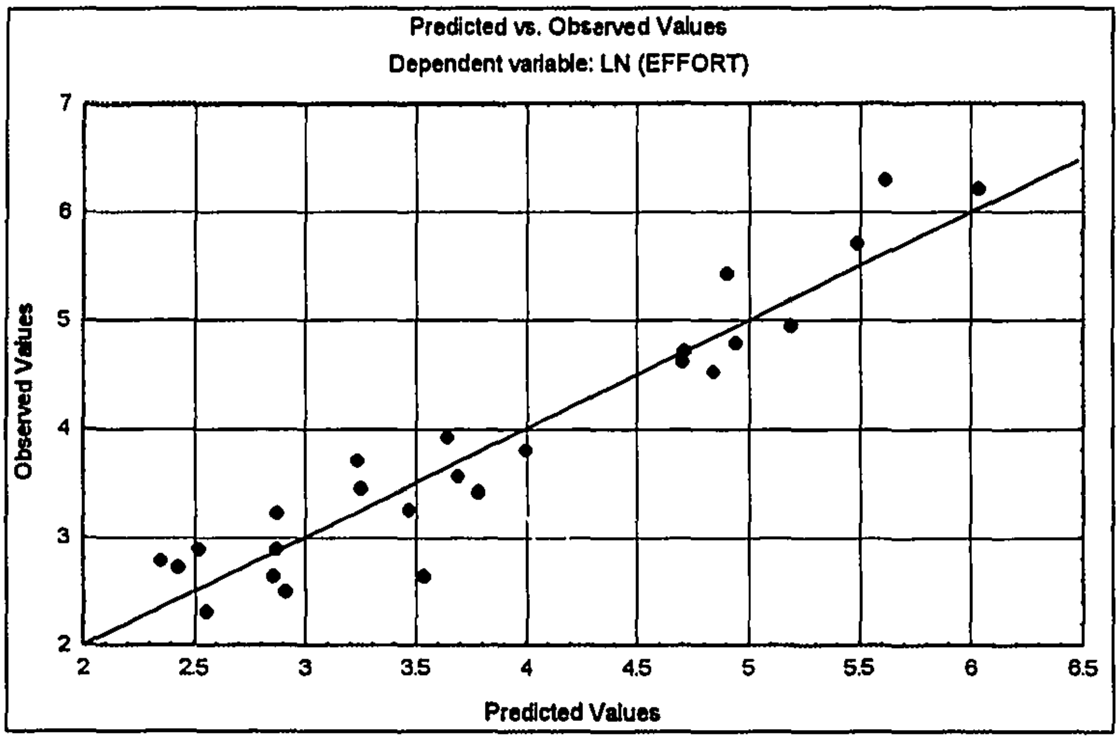


Figure 7-3: Predicted versus observed values (effort estimation model)

The validity of the regression model depends on the validity of the inference assumptions. We shall start with the first inference assumption, homoscedasticity or constant variance of the residual. Figure 7-4: Predicted versus residual scores (transformed data), shows the predicted versus the residual scores plot of the transformed ISO 9001 data. The residual plot clearly indicates that the residuals are well behaved and have a horizontal band appearance. Thus we can conclude that the assumption of homoscedasticity holds.

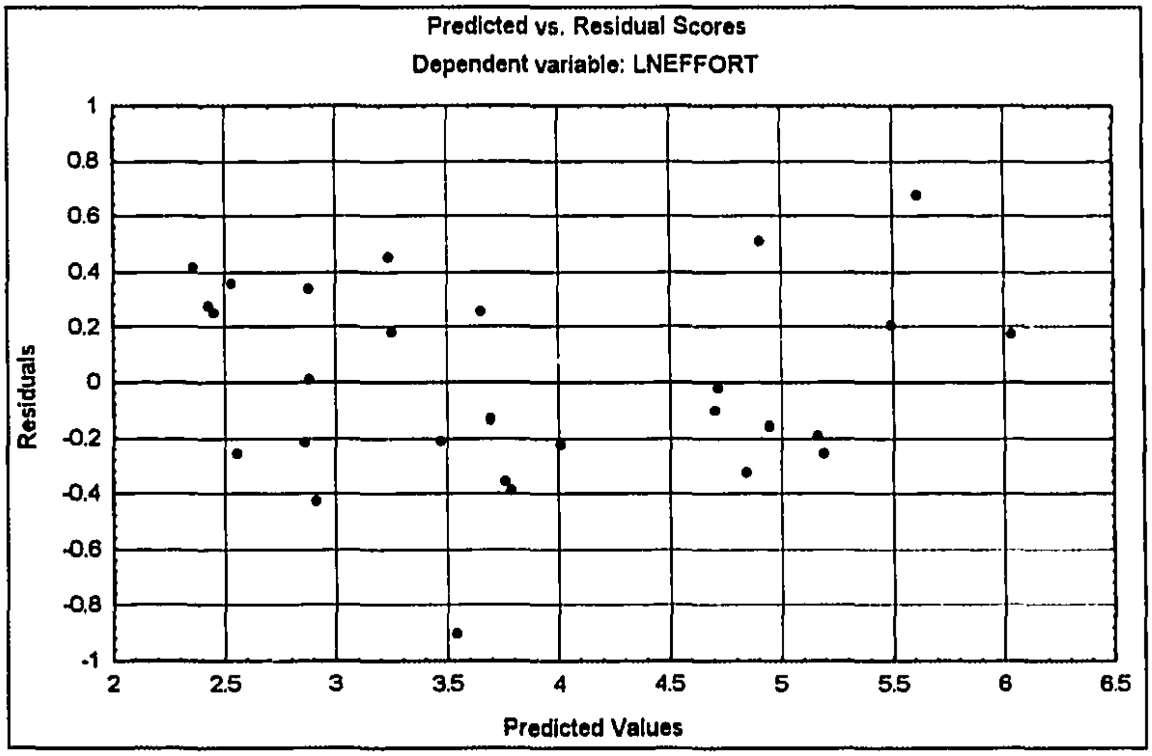


Figure 7-4: Predicted versus residual scores (transformed data)

The second inference assumption we need to validate is the independence assumption. Again we use the Durbin-Watson test of the error term. We will employ the Durbin-Watson test for (first-order) positive and (first-order) negative autocorrelation. Thus, we test the null hypothesis

H_0 : The error terms are not autocorrelated

versus the alternative hypothesis

H_1 : The error terms are positively autocorrelated

Durbin and Watson have shown that, based on setting the probability of a Type I error equal to α,

1. If $d < d_{L,\alpha}$ we reject H_0.

2. If $d > d_{U,\alpha}$ we do not reject H_0.

3. If $d_{L,\alpha} \leq d \leq d_{U,\alpha}$ the test is inconclusive.

We computed the Durbin-Watson statistic using SAS and STATISTICA, and looked up the critical values for the Durbin-Watson d statistic ($\alpha = .05$), and we obtained the following values:

Model	Durbin Watson d	$d_{L,\alpha/2}$	$d_{U,\alpha/2}$	α	Autocorrelation
Transformed	1.5937	1.26	1.56	.05	None

Table 7-7: Durbin-Watson d statistic for positive autocorrelation

Similarly we test the hypothesis for negative autocorrelation.

H₀ : The error terms are not autocorrelated

versus the alternative hypothesis

H_1 : The error terms are negatively autocorrelated

Durbin and Watson have shown that, based on setting the probability of a Type I error equal to α,

1. if $(4 - d) < d_{L,\alpha}$ we reject H_0.

2. if $(4 - d) > d_{U,\alpha}$ we do not reject H_0.

3. If $d_{L,\alpha} \leq (4 - d) \leq d_{U,\alpha}$ the test is inconclusive.

We computed the Durbin-Watson statistic using SAS and STATISTICA, and looked up the critical values for the Durbin-Watson d statistic ($\alpha = .05$), and we obtained the following values:

Model	Durbin Watson d	$d_{L,\alpha/2}$	$d_{U,\alpha/2}$	α	Autocorrelation
Transformed	1.5937	1.26	1.56	.05	None

Table 7-8: Durbin-Watson d statistic for negative autocorrelation

Therefore, we do not reject the null hypothesis of no correlation. That is, there is no evidence of positive or negative (first-order) autocorrelation.

The last inference assumption we need to validate is the normality assumption. We shall use the normal probability plot to validate this assumption. The normality assumption holds if the plot has a straight-line appearance.

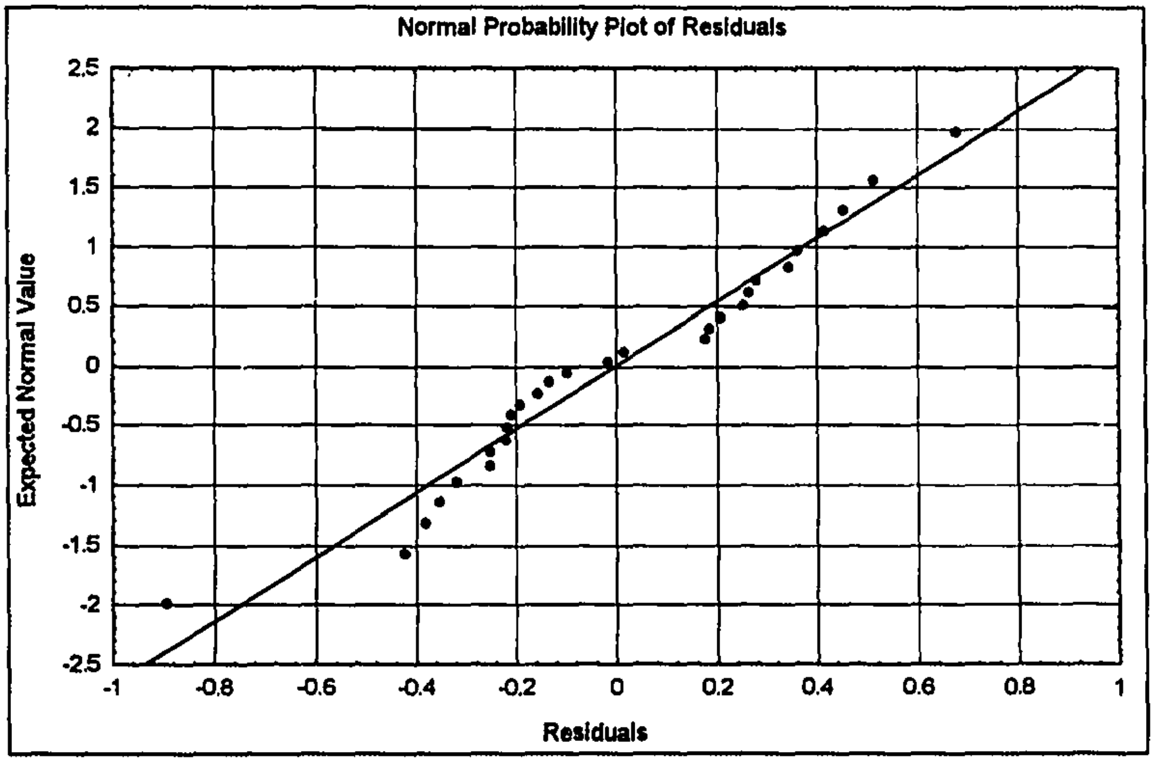


Figure 7-5: Normal probability plot (Transformed data)

Figure 7-5 shows the normal probability plot using the transformed data. The normal probability plot has a better fit, and can assume normality. Thus we can conclude that the data transformation has suitably equalized the variance.

Finally, we shall perform an ***F***-TEST for the overall model. The ***F***-test is related to the utility of the overall regression model. We shall test the null hypothesis

H_0: $\beta_1 = \beta_2 = 0$

which says that none of the independent variables x_1, x_2 affects y, versus the alternative hypothesis

H_1: At least one of β_1, β_2 does not equal zero

Which says that at least one of the independent variables x_1, x_2 affects y.

We can reject H_0 in favor of H_1 by setting the probability of a Type I error equal to α if and only if $F(\text{model}) > F^{(k-1,n-k)}_{[\alpha]}$

We set α equal to .05, and we know that K (the number of parameters in the model) is equal to 3, and n (the number of observations) is equal to 28.

The F (model) = 123 and $F^{(3-1,28-3)}_{[.05]} = 3.39$, therefore, we can reject H_0 in favor of H_1 by setting α equal to .05 and we can conclude that at least one of the independent variables x_1, x_2 significantly affects y.

With this, we conclude our test of this regression model, and confidently confirm that it meets the tests for linearity, normality, independence and homoscedasticity of the residuals assumptions, and that it has a large R^2 which indicates that a large proportion of the total variation is explained by the regression model.

7.5 Model Validation

Regression model validation is an important step in determining the usefulness of any regression model. A regression model is usually validated using a data set that is different from the data set used to build the model. Generally, we check the model's predictive ability on the different data set, which might be newly collected data or a hold-out sample from the original data. Another way to validate a model is by comparing it to theoretical expectations and earlier empirical findings. If new data can not be obtained or if the original data set is not large enough to allow for a hold-out sample, the ***PRESS*** residual and statistic can be used for validation.

We have performed extensive validation on the regression model, and carried out validation using new data which was not used in building the model. In addition, we validated the prediction interval and power of the model using the old and new data.

7.5.1 Validation Using New Data

Often enough, the original data is not large enough to warrant the use of a hold-out sample, because that reduces substantially the number of observations required for model building. In order to satisfy both conditions, we used the entire data set in building the model, and then obtained a new data set for the model validation.

The new data set was acquired in June 1995 using the same survey questionnaire, data collection and validation methods described in Chapter 4. We shall present the use of the estimation model on the new data acquired, and predict the effort for implementing ISO 9001 for software organization no. 1.

Software organization no. 1, is a relatively large software organization, with offices across Canada, and employs over 1000 employees. The number of employees covered by the ISO 9001 registration were 175. The organization's quality manager responsible for ISO 9001, estimated that the degree of non-compliance prior to their commencement on implementing ISO 9001 was 35%. The elapsed time from start to registration was 6 months, and the actual effort that was taken to implement ISO 9001 was calculated as follow:

Two staff members working 70% of the time on ISO 9001 for 6 months in region A.

Two staff members working 70% of the time on ISO 9001 for 6 months in region B.

Two staff members working 70% of the time on ISO 9001 for 6 months in region C.

One staff member working 100% of the time on ISO 9001 for 6 months at corporate office.

Therefore, the actual total effort taken to implement ISO 9001 in organization no. 1 was 31.2 months.

Our estimate using the estimation mode and based on the data provided by this organization was 30.3 months, calculated as follow

$$\begin{aligned}
\mathrm{Ln}(y) &= -2.793 + .692 * \mathrm{Ln}(x1) + .74 * \mathrm{Ln}(x2) \\
\mathrm{Ln}(\mathrm{Effort}) &= -2.793 + .692 * \mathrm{Ln}(175) + .74 * \mathrm{Ln}(35) \\
&= -2.793 + .692 * 5.165 + .74 * 3.555 \\
&= 3.412
\end{aligned}$$

Therefore, if Ln (Effort) = 3.412, then the effort is $e^{3.412}$ = 30.3 months. We can observe that the predicated effort by the model, was very close to the actual effort reported.

Table 7-9 presents the predicted and observed effort of the other organizations. We can see that the prediction capability of the model was very close to the actual, except for organization no. 4, which was a new software development and support function that had no process in place.

Organization number	Elapsed time (months)	Size of organization	Degree of non-compliance	Predicted effort (months)	Actual effort (months)	Residual (months)
1	6	175	35%	30.3	31.2	.9
2	5	108	15%	11.6	13	1.4
3	9	170	30%	26.5	27	.5
4	12	45	100%	25.8	36	10.2
5	16	100	70%	34.4	37	2.6

Table 7-9: New data obtained and predicted effort

7.5.2 Prediction Interval Validation

The validation of the model predictive ability is of prime importance. We present in Table 7-10 the predicted versus actual effort observed and the lower95% and upper95% prediction interval of the 28 software organizations we surveyed.

A 95% prediction interval indicates that we are 95% confident that the estimated effort will be in a given range. In estimating efforts, the upper bound of the prediction interval is more important than the lower bound, since organizations are generally more concerned with estimation overruns rather than coming under in the estimated effort.

Having examined Table 7-10, we can be confident in saying that the effort estimation model presented in Equation 7-1 is a valid and useful model for estimating the effort of implementing ISO 9001 in software organizations.

Organization	Observed	Predicted	Lower 95%	Upper 95%
1	14.00	17.33	12.69	23.68
2	25.00	17.74	13.72	22.93
3	300.00	243.78	177.50	334.81
4	18.00	17.79	14.16	22.35
5	15.00	11.32	8.68	14.75
6	110.00	111.73	86.62	144.10
7	40.00	25.44	21.42	30.21
8	540.00	275.37	201.16	376.94
9	15.00	11.66	9.25	14.70
10	10.00	12.82	9.99	16.44
11	30.00	43.98	37.77	51.22
12	16.00	10.56	8.02	13.89
13	498.00	414.83	300.94	571.83
14	18.00	12.54	9.94	15.81
15	31.00	25.72	18.61	35.55
16	140.00	180.44	143.51	226.88
17	120.00	140.47	109.06	180.93
18	35.00	40.06	28.24	56.84
19	50.00	38.62	32.29	46.20
20	144.00	173.64	127.50	236.48
21	100.00	110.48	89.53	136.34
22	30.00	42.74	36.01	50.73
23	92.00	126.76	104.51	153.74
24	44.00	54.98	45.09	67.03
25	12.00	18.37	14.97	22.55
26	14.00	34.21	29.40	39.81
27	226.00	135.89	108.97	169.46
28	26.00	32.21	27.15	38.21

Table 7-10: Predicted versus actual effort observed

8 Comparison and Discussion

The world-wide movement towards ISO 9000 registration, and other standards and schemes, accentuate the importance of estimating the effort required for implementing ISO 9000. In this chapter, we discuss the significance of our results, and compare them with other related work in the field of software engineering, with a focus on software organizations.

The importance of our work stems from the fact that this first published effort estimation models dealing with ISO 9001 for software. Our previously published generic effort estimation model [RM95], is not specific for software, but rather deals with estimating the effort of ISO 9001 for all industries (generic).

In comparing our work, two types of comparisons are done. First, we compare the software ISO 9001 estimation model versus the generic ISO 9001 estimation model. Second, we provide a comparison of the prediction power and interval of our model versus other software engineering cost estimation models.

Software Model versus Generic Model

The ISO 9001 estimation model for software uses two independent variables: The size of the organization, and the degree of non-compliance; while the ISO 9001 generic model uses the cross-product term: (Size of organization*degree of non-compliance).

The estimation model in Equation 8-2 was built using data from 28 ISO 9001 software registered organizations; while the estimation model in Equation 8-3 was built using data from 112 ISO 9001 organization from all industries.

ISO 9001 effort estimation model (software)
Ln (Effort) = -2.793 + 0.692 * Ln (Size of organization) + 0.74 * Ln (Degree of non-compliance)

Equation 8-2: ISO 9001 Effort Estimation Model for Software

ISO 9001 effort estimation model (Generic)
Ln (Effort) = -1.422 + 0.55697 * Ln (Size of organization * Degree of non-compliance)

Equation 8-3: ISO 9001 Effort Estimation Model for All industries [RM95]

In order to compare the two models, we first plot each model separately at various degrees of non-compliance. This shows the impact of the degree of non-compliance on the predicted effort required. The higher the degree of non-compliance the more effort is required to meet the ISO 9001 standard.

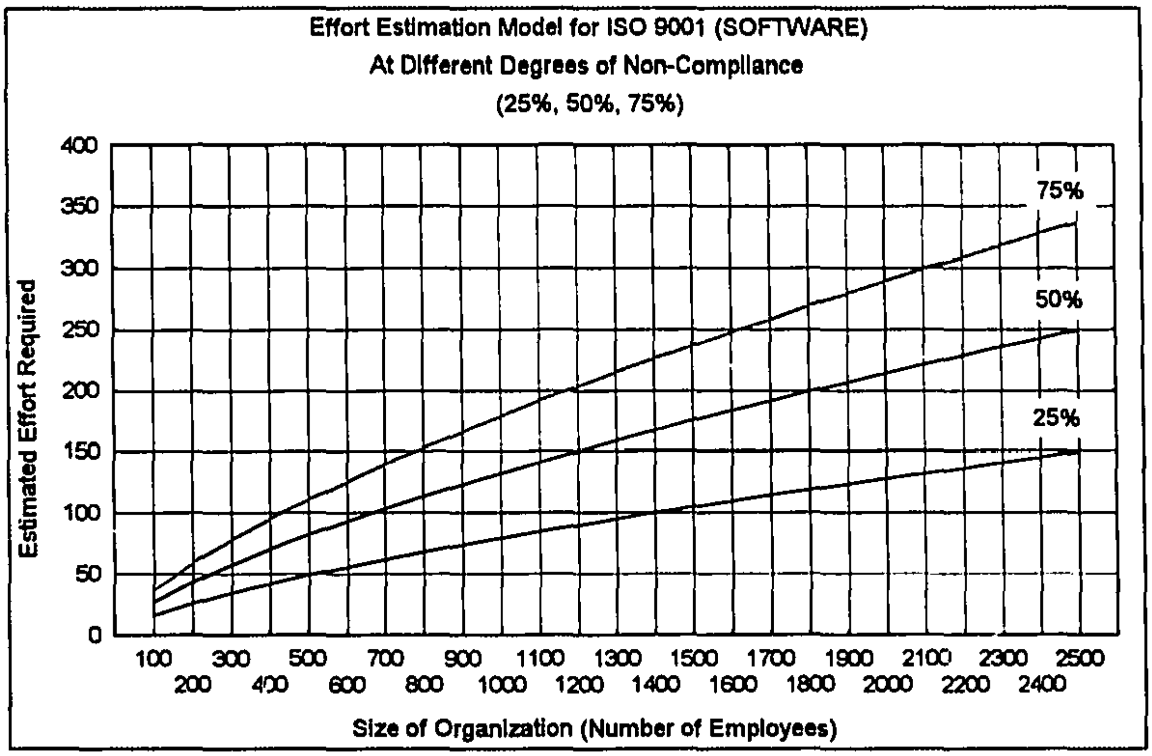


Figure 8-6: Estimated effort (Software Model)

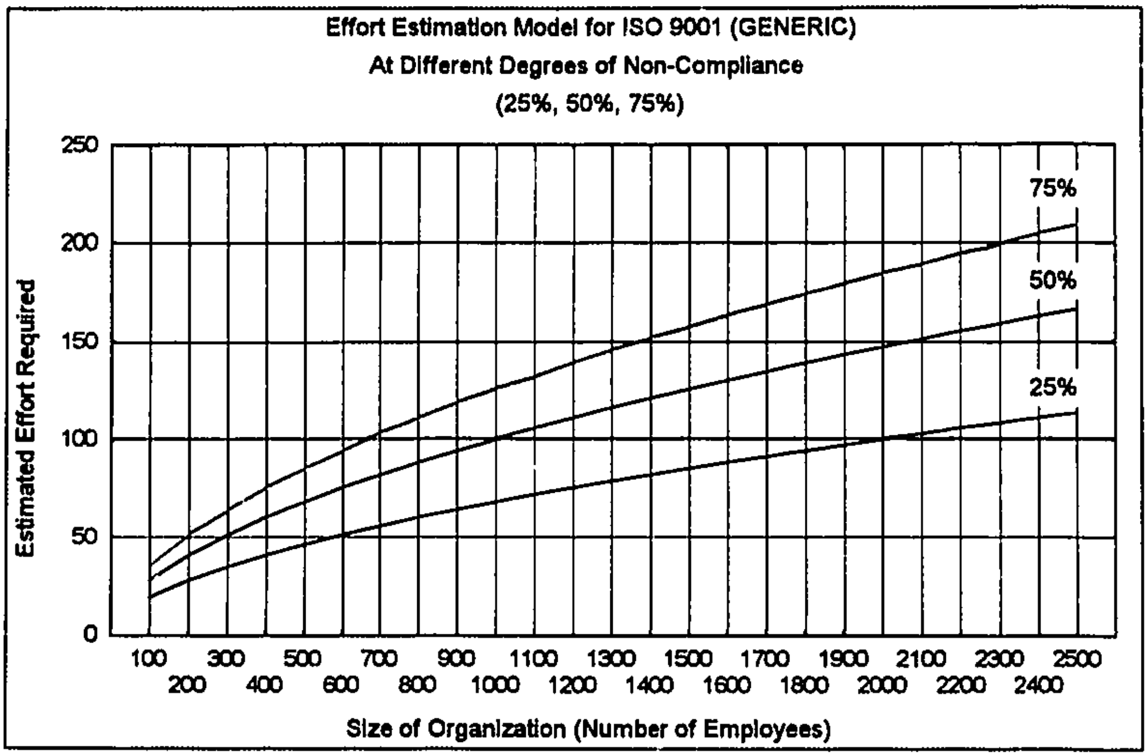


Figure 8-7: Estimated effort (Generic Model)

In Figure 8-8, we compare the two models (software versus generic), at a degree of non-compliance equal to 50% (50% degree of non-compliance was chosen arbitrarily). Evidently the effort required for a software organization that currently meets 50% of the ISO 9001 standard requirements, is more than the effort required for a similar non-software organization, specially if the size of the organization is more than 150 employees.

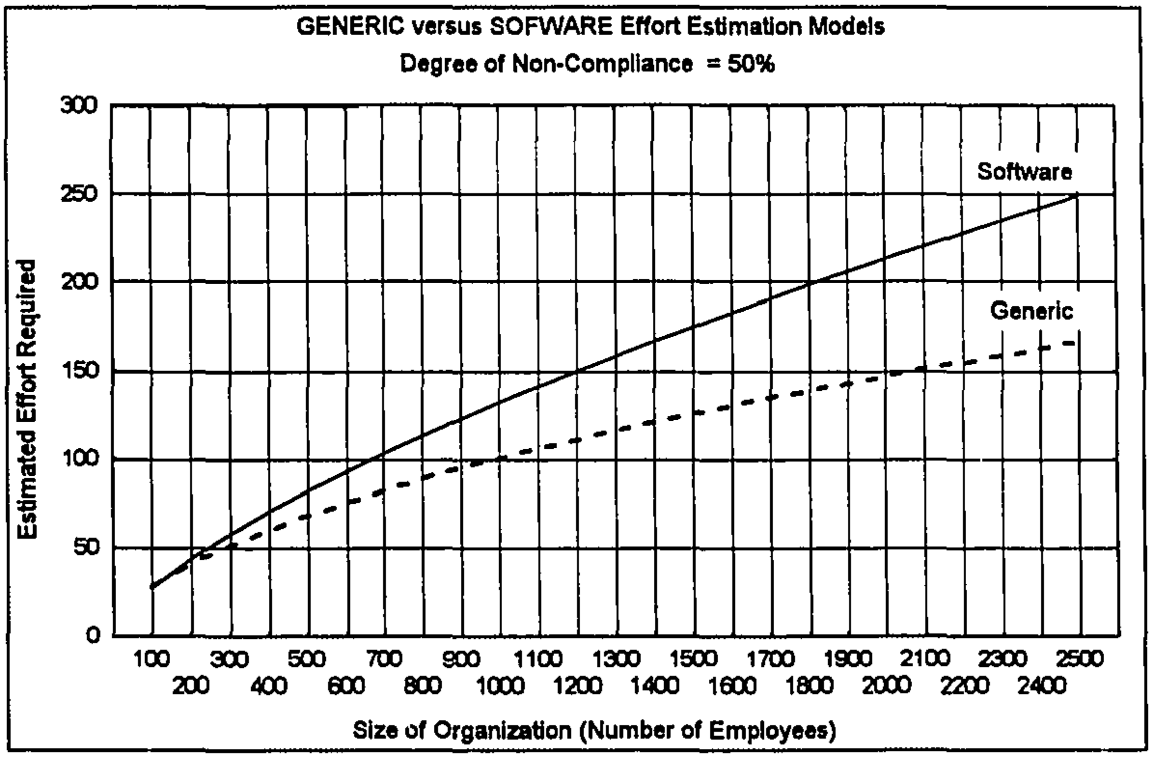


Figure 8-8: Software versus Generic industry Estimated Effort

This important discovery leads to many questions, suggestions, and further research. This research has provided objective evidence, that the reported effort for implementing ISO 9001 in a software organization is higher than most other industries. There are many probable reasons why software organizations require more effort. We list a few: (i) Software products are inherently complex and challenging; they can quickly reach a high level of complexity [Coa94]; (ii) software is a relatively new field that has seen an exponential increase in technology, and is moving faster than any other field; (iii) software is a very competitive field, users' expectations have been set high; (iv) software is dynamic and not tangible; (v) software development requires high degree of skill, knowledge and experience at all levels of the development process; (vi) software development principles and methodologies require time to be developed and practiced.

Future research, can focus on this area, to determine possible causes of this outcome.

Prediction Interval Comparison

Many of the estimation models in software have a wide prediction interval, specially when the 90% or higher confidence is desired. The prediction interval can be 2 to 3 times the predicted value [MBM94].

In comparison, our results have indicated that the "goodness" of our estimation model and its prediction interval is comparable and initially better than those used in practice. The worst upper95% prediction interval we encountered on the 28 observations we had, amounted to 1.41 times the predicted effort value [Refer to chapter 7] . Therefore, in comparison with other published models, this model is considered useful, and with relatively good prediction interval.

9 Summary and Conclusion

Software quality, and software process improvement, are two important goals in software engineering. The adoption of software development principles and methodologies embodying best practices and standards are essential to achieving quality software products and improving the software development process.

The International Organization for Standardization, has published the ISO 9000 series of standards on quality. The ISO 9000 series includes three generic standards: ISO 9001, ISO 9002, and ISO 9003. The ISO 9001 standard is the most comprehensive. Software development organizations involved in the design, development, testing, installation, support and maintenance of software, wishing to register to ISO 9000, are required to comply with the requirements of the ISO 9001 standard.

Meeting the requirements of the ISO 9001 standard, can be costly in time, effort and money. The effort primarily depends on, the size of the organization, and the status of the quality management system. The ability to estimate the effort required for meeting the requirements of ISO 9001, is an *a priori* requirement in corporate decision making. Our work focused on the development of an effort estimation model, for implementing ISO 9001 in software organizations. In determining this effort, a survey of 1190 registered organizations was conducted in 1995 of which 63 were software companies. The data collected from software companies, was used in the development of a statistical regression model for predicting the effort of implementing ISO 9001 in software organizations. This estimation model forms a foundation, for building and comparing related effort estimation models for ISO 9000, and other process improvement frameworks.

The existence of an ISO 9001 effort estimation model, would allow software organizations that are contemplating registration, to predict the required effort, and determine the affordability of such a decision. In addition, the availability of such a tool, will assist in the financial and resource planning of this important corporate decision, thus reducing the risk factor, by providing an estimate that is based on statistical model, rather than a simple prediction.

Registering a facility to any ISO 9000 standard, reduces or eliminates the need for individual purchasers or customers to perform their own assessment of the supplier's quality system. In addition, it is an attestation to the world that the quality system conforms to the requirements of the ISO 9000 standard.

This work also resulted in two other first-time discoveries regarding the software industry. First, the data provided from software and non-software organizations showed that software organizations in general, were meeting more of the requirements of the standard prior to the implementation of ISO 9001 than non-software organizations. Second, the effort estimated for complying with the requirements of the ISO 9001 standard for software companies is more than the effort required for non-software companies.

These initial findings are of interest to the software engineering and research community, and merit further research and analysis.

Finally, there are at least two other possible areas where the proposed effort estimation model may be applicable and useful.

1. The model may be used for estimating the effort required for implementing ISO 9001 for software organizations using the TickIT scheme. The TickIT scheme

certifies software organizations using ISO 9001 through the application of ISO 9000-3.

2. As per comparison and mapping between the SEI CMM and ISO 9001 [Pau94], an organization with level 2 practices, and with some level 3 practices, such as training program, peer reviews and software product engineering, is somewhat equivalent to a software organization with ISO 9001 registration. Therefore, the ISO 9001 effort estimation model for software may provide an estimate for an organization wishing to attain level 2 under the CMM model.

Our objective is to perform further refinement and fine tuning of the estimation model. This important step is only achievable if more data is collected from other organizations. The more data available for building a model, the more significant is the model.

Bibliography

[Alb79] Albrecht, A.J. "Measuring application development productivity," in Proceedings of the IBM Applications Development Symposium, GUIDE/ SHARE (Monterey, Calif., Oct 14-17) IBM, 1979, pp. 83-92.

[AS85] Alreck, P.L. and Settle. R.B, *The Survey Research Handbook*, 1985 Richard D. IRWIN INC.

[ATT94] Using ISO 9000 to Improve Business Processes, AT&T Corporate Quality Office, July 1994

[Bas91] Basili, V.R. and Selby, R.W "paradigms for Experimentation and Empirical Studies in Software Engineering" Reliability Engineering and System Safety, 1991.

[Bau95] Bauer, L. "ISO 9000 for Governments: not if, but when", Standards Council of Canada, Consensus / Winter 1995.

[Boe81] Boehm, B.W., *Software Engineering Economics*, Englewood Cliffs, New Jersey. Prentice-Hall, 1981

[BO90] Bowerman, B. L. and O'Connell, R.T., "Linear Statistical Models an Applied Approach", second edition, 1990, Duxbury press

[BW84] Basili, V. R. and Weiss, D. M. "A Methodology for Collecting Valid Software Engineering Data," IEEE Transactions on Software Engineering, Vol. SE-10, No. 6, November 1984. pp. 728-738

[CC79] Cook, T.D., and Campbell, D.T. Quasi-Experimentation: Design and analysis issues for field settings. Chicago: Rand McNally, 1979.

[CIS93] Deloitte & Touche and Quality Systems Update. "ISO 9000 Survey", September 1993, CEEM Information Services, [CIS95] Quality Systems Update, "Statistical profile of ISO 9000 Registered Companies USA and Canada February 1995", CEEM Information Services.

[CMM93] Software Engineering Institute SEI/CMM Capability Maturity Model V1.1 Carnegie Mellon University. SEI-93-TR-25

[Coa94] Coallier, F. "How ISO 9001 Fits into the Software World," IEEE Software Vol. 11, No. 1, January 1994, pp. 98-100

[CS63] Campbell, D.T., and Stanley, J.C. "Experimental and quasi-experimental designs for research on teaching." In L. Gage (ed.) *Handbook of research on teaching.* Chicago: Rand McNally, 1963.

[Fow84] Fowler, F.J. jr, "Survey Research Methods,"Sage Publication Inc. 1984.

[Gal94] Gianluigi Galdiera "Impact of ISO 9000 on Software Maintenance" Institute of Advanced Computer Studies, Department of Computer Science, University Of Maryland At College Park. PP. 228 - 230

[HCR94] Herbsleb, J. and Carleton, A. and Rozum, J. and Siegel, J. and Zubrow, D. "Benefits of CMM-Based Software Process Improvement: Initial Results" CMU/SEI-94-013 August 1994

[HGJ94] Hockman, K.K. and Grenville R. and Jackson S. "Road Map to ISO 9000 Registration", Quality progress, May 1994 pp. 39-42

[HJ78] Hoinville, G. and Jowell, R. and associates (1978) Survey Research Practice. London, Heinemann

[Hum89] Humphrey, W.S., *Managing the Software process*, SEI series in Software Engineering, Addison-Wesley, 1989

[Ima86] Imai M., 1986, Kaizen - The Key to Japan's Competitive Success, McGraw-Hill Publishing Company

[ISO9001] ISO 9001:1994 Quality Systems - Model for quality assurance in design, development, production, installation and servicing

[ISO9002] ISO 9002:1994 Quality systems - Model for quality assurance in production, installation and servicing

[ISO9003] ISO 9003:1994 Quality systems - Model for quality assurance in final inspection and test

[ISO900-3] ISO 9000-3:1991 Quality management and quality assurance standards - part 3: Guidelines for the application of ISO 9001 to the development, supply and maintenance of software

[Jol86] Joliffe, F. R. (1986) Survey Design and Analysis. Ellis Horwood Limited 1986

[Kem87] Kemerer, C.F. "An Empirical Validation of Software Cost Estimation Models," Communications of the ACM, Vol 30 No. 5, May 1987, pp. 416-429

[KJ86] Kidder, L.H and Judd, C.M. "Research Methods in Social Relations", 1986 CBS College Publishing

[KS82] Kalton, G. and Schuman, H. (1982) The effect of the question on survey responses: a review, J.R. Statist. Soc.

[Lon87] Londeix, B. "Cost Estimation for Siftware Development" 1987, Addison-Wesley Publishing Company

[MBM94] Matson, J.E. and Barrett, B.C. and Mellichamp, J.M. "Software Development Cost Estimation Using Function Points," IEEE Transactions on Software Engineering, Vol 20. No. 4, April 1994, pp.275-287

[MCJ91] "Vision 2000: A Strategy for International Standards' Implementation in the Quality Arena During the 1990s" ISO 9000 Compendium, 3rd Edition 1993.

[MD90] Maxwell, S. E. and Delaney H. D. "Designing Experiments and Analyzing Data," 1990 Wadsworth Publishing Company

[MK71] Moser, C. A. and Kalton, G. (1971) Survey Methods in Social Investigation. London, Heinemann Educational

[Opp66] Oppenheim, A.N. (1966) Questionnaire Design and Attitude Measurement. London, Heinmann.

[Pau94] Paulk, M.C., "A Comparison of ISO 9001 and the Capability Maturity Model for Software." 1994, CMU/SEI-94-TR-12

[Pea94] Peach, R.W., ed. 1994 *The ISO 9000 Handbook*, 2nd ed. Fairfax, Virginia: CEEM Information Services.

[Ped82] Pedhazur, E.J., "Multiple Regression in Behavioral Research", second edition 1982 CBS College Publishing.

[Pfl95] Pfleeger, S.L., "Experimental Design and Analysis in Software Engineering." ACM SIGSOFT, Software Engineering Notes Vol. 20. No. 1, January 1995, pp. 22-27

[PS81] Platek, R. and Singh, M.P. (1981) Cost Benefit Analysis of Controls in Surveys. In: Krewski, D., Platek, R. and Rao, J.N.K. (eds.) Current Topics in Survey Sampling. New York, Academic Press, pp. 105-136.

[RB93] Rabbitt, J.T. and Bergh, P.A. "The ISO 9000 Book, A Global Competitor's Guide to Compliance & Certification," 1993, Quality Resources

[RM95] Rahhal, S. and Madhavji, N. H. "An Effort Estimation Model for Implementing ISO 9001", International Software Engineering Standards Symposium (ISESS95), August 21-25, 1995, pp. 278-286.

[SB82] Sudman, S. and Bradburn, N.M (1982) Asking Questions. A Practical Guide to Questionnaire Design. San Francisco, Jossey-Bass

[Sch94] Schmauch, C.H. *ISO 9000 for Software Developers*, ASQC Quality Press 1994

[SPA81] Summers, G.W. and Peters, W.S. and Armstrong C.P. *Basic Statistics in Business and Economics*, 3rd edition 1981 Wadsworth publishing company

[SPN95] Drouin, J.N., "The SPICE Project: An Overview," Software Process Newsletter, No. 2, Winter 1995, pp.8-9.

[SPQ94] "A Survey of the Surveys on the Benefits of ISO 9000", Software Process, Quality & ISO 9000, Vol. 3, No. 11, November 1994.

[Str89] Straub, D. W. "Validating Instruments in MIS Research", MIS Quarterly/June 1989 pp. 147-168

[Tic92] Department of Trade and Industry and the British Computer Society. 1992. TickIT, Guide to Software Quality Management System Construction and Certification using EN29001. DISC TickIT Office 2 Park Street London W1A 2BS

[Tic93] TickIT Case Studies. TickIT Campaign office of IT World Limited, January 1993, 18 Buckingham Gate London SW1E 6LB

[Tri94] Trillium, Model for Telecom Product Development and Support Process Capability, Release 3.0, Bell Canada

[Wol94] Wolak, J., "ISO 9000 a Software Market.", Quality, March 1994 pp.44-45

[Yin94] Yin, R. K. "Case Study Research: Design and Methods," 1994, Sage Publication Inc.

Appendix A - ISO 9000 Cover Letter

Appendix B - ISO 9000 Questionnaire

McGill

School of Computer Science
McConnell Engineering Bldg, Suite 318
McGill University
3480 University Street
Montréal, QC H3A 2A7

Software Engineering Group
Silas F. Rahhal
Tel. : (514) 398-5928
Fax : (514) 398-3883
E-Mail: rahhal@process.cs.mcgill.ca

ISO 9000 Questionnaire

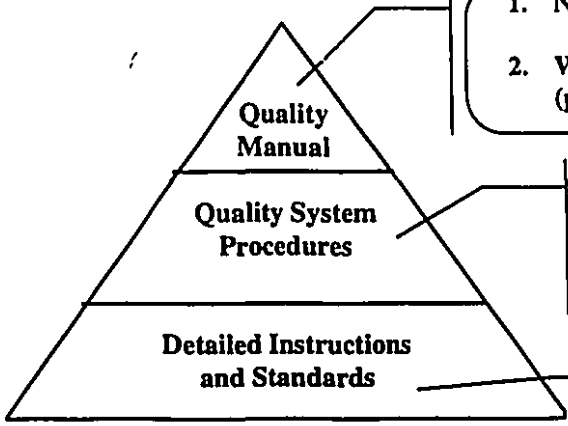
Please complete, then return pages 1 & 2 by February 24, 1995.
We appreciate your time...

ISO 9000 - Questionnaire

This Questionnaire should be filled out by the Quality Manager or the ISO Coordinator.
(If it has reached you in error, please forward it to the appropriate person. Thank you.)

1. Number of employees covered in the **ISO 9001** registration audit:	_______Employees
2. Number of sites covered in the **ISO 9001** registration audit: *Note: A Site is a separate facility/Plant or location.*	_______Sites
3. Number of elapsed (calendar) months from start to completion and successful **ISO 9001** registration audit.	_______Months
4. Total time spent (work-effort) in preparation for meeting the requirements of **ISO 9001**, from start to completion and successful registration audit. *Example: 4 employees full time for 7 months = 2 Years and 4 months.*	_______Years _______Months
5. Of the total time spent, approximately, what percentage was spent on each of the following activities: I. Development of all Procedures and Instructions, including Quality Manual. II. Staff training time and training material preparation .... III. Audit activities, Process review and Corrective/Preventive action .... IV. **ISO 9001** Project coordination and administration............................... V. Other, Specify _ ________________________ .	 ________% ________% ________% ________% ________% **Total 100 %**

6. Number of pages and the time spent (work-effort) on the ***development, review and approval*** of the following.



Quality Manual
1. **Number of pages ?** ____________
2. **Work-effort spent (person-months) ?** ____________

Quality System Procedures
1. **Number of pages ?** ____________
2. **Work-effort spent (person-months) ?** ____________

Detailed Instructions and Standards
1. **Number of pages ?** ____________
2. **Work-effort spent (person-months) ?** ____________

7. Estimate the percentage of the **ISO 9001** requirements your facility was meeting <u>before any preparation work was started</u>. *Example: If your facility had no quality management system in place and no defined and documented procedures and no quality records, then your facility's compliance was close to 0%. If your facility had every thing in place, and no preparation was necessary, just required auditing, then your compliance was close to 100%, otherwise, choose the appropriate level of compliance.*	________ %

Additional comments: ____________________

____

____

____

____

To receive the ISO 9000 survey results, please ensure that the following information is correct.

Company Name	: BIONAIRE INC.
Address	: 2000 - 32nd Avenue Lachine, PQ Canada, H8T 3H7
Registration	: ISO 9001
ISO 9000 coordinator	: George Laszlo
SIC	: 5064

Thank you very much for completing this questionnaire! All information will be kept in complete confidence, and reported at aggregate level only (SIC code).

For your convenience, please use the enclosed self addressed envelope ✉.

If you should have any questions, please contact:

Silas F. Rahhal
McGill University
School of Computer Science
3480 University Street
Montreal, Quebec, Canada H3A 2A7
email: rahhal@process.cs.mcgill.ca
☎ *(514) 398-5928*