# UNIVERSITY OF ABERTAY DUNDEE

# RISK BASED ANALOGY FOR E-BUSINESS ESTIMATION

# ELIAS PIMENIDIS

A thesis submitted in partial fulfilment of the requirements of the University of Abertay Dundee for the award of the degree of Doctor of Philosophy.

**July 2011**

# Acknowledgments

During the many years that I have toiled with this work I have been indebted to a large number of people and I would like to thank them each one of them for the time, support and useful materials that they have provided.

Especially, my gratitude goes to my director of studies Dr Colin J. Miller, who has managed the very difficult act of balancing the roles of mentor and friend extremely well. I could not have accomplished this without his patience, advice and support.

Many thanks are due to my supervisor Ms Claire MacEachen, who has provided timely and constructive feedback when it was needed most.

No work can be completed, if one does not have the right support at home. I dedicate this work to my wife Timokleia for her long enduring patience and the timely motivation when I almost felt like giving up and to my parents Irene and Dimitrios, who have always trusted my choices, supported my efforts and taken pride in everything that I have achieved.

# Contents

# List of Tables

# List of Figures

# Abstract

Over the past fifteen years, new technologies have enabled the evolution of e-business and as new trends on web based applications develop at a very fast pace, projects to implement new and adapt existing e-business systems are becoming more complex in their structure. At the same time, organizations involved as stakeholders are becoming more demanding in the delivery of fast, accurate and timely results. Where interconnected e-business systems, such as Value-Added-Communities, are concerned the key criteria of success, in addition to the quality of the technical work, are the timely completion of the project and an accurate estimate of the launch time. Given the nature of such e-business development projects, estimating their duration cannot be based on traditional estimation methods that target the estimation of human effort originally and then attempt to convert this to cost and duration. This research explores a new approach aiming to improve the accuracy of estimates for e-business development projects, obtained at the early stages of the project. The key objective is to modify the estimation by analogy approach by using risk as the key element that along with project size and complexity are used to identify analogues and determine analogies amongst projects.

The author established the behaviour of risk with the variation of certain project attributes that define a project's size and complexity at the very early stages of the project, before detailed requirements are identified. The variation of risk was subsequently employed, with a modification to the estimation by analogy method, to establish analogies on suitably identified analogues to the project to be estimated. The ISBSG project data repository has been used as source for potential analogue projects in testing the proposed method. Four project cases were considered. New estimates have been obtained using the modified estimation approach method and the results were compared to the original estimates for the projects, obtained using traditional methods.

Results obtained within the context of this research are encouraging and suggest that there is credibility to the proposed modification. The accuracy of the estimates is within acceptable tolerance levels and shows signs that it can improve on the quality of estimates obtained from traditional methods when applied to e-business development projects. More research is required though to expand and fully exploit the potential of this approach. In particular the author believes that the approach could be further fine tuned for improved accuracy of the estimated duration. This risk based estimation by analogy approach could then be applied to other types of projects which share characteristics with e-business projects such as, demands for short time development, complex interactions of stakeholders and accuracy in the time of completion. Types of projects where the proposed method could possibly be utilized are web services implementation, computer games development and cloud computing projects.

# INTRODUCTION

Conducting business electronically is not a recent development. Traditional business models were process-centred where emphasis was placed on improving processes throughout the organization, aiming to improve performance and to meet the organization's strategic objectives. Current business models that have evolved in the early years of the 2000's under the concept of creation of "Value-added communities" (McWiliam 2000, Means and Schneider 2000, pp.19-32) aim at synergy across the supply chain.

Value-added communities are groups of businesses that function at the various points of the supply chain and are connected electronically to enable maximum customer satisfaction. At the same time this electronic network should offer maximum return to the "community" as a whole. Through the electronic business facility of the trading organization (brand-owning company), information is processed, filtered and forwarded, through the relevant networks, to other computer systems such as MRP, MRPII, ERP, that each may support the function of one of the members of the "community". Thus planning and coordination of activities within the "community" can be performed according to market trends as they evolve and are continuously revised on a real time basis (Andal-Ancion et al 2003, Starr et al 2003). The creation and evolution of VACs will not be possible without the support of Web based technologies and other information systems, which will provide the necessary bond to tie the participating companies together. Thus Internet technologies can contribute to the coordination of operations for the members of a VAC, leading to minimization of wasteful activities and improved efficiencies. Participating companies need to focus on the development and adoption strategies regarding new technologies and e-business related systems if they are to create the conditions for improved markets for themselves (Venkatraman, 2000 and McAfee, 2005 and 2006).

A VAC is not a static compound, but a constantly evolving organism that responds to the challenges and opportunities offered by emerging technologies, i.e. web 2.0, web services and cloud computing, and the business environment. The technologies provide organisations with a plethora of options that they could embrace in enhancing the way they communicate and collaborate with partners and customers (Brynjolfsson and McAfee, 2007).

The information infrastructure required to support e-business, is therefore a collection of large complex systems. The fact that these are diverse systems processing the same type of input increases the difficulty of the development project as they often have to serve different requirements and need to update systems from different generations. As a consequence the complexity of managing and the degree of risk in estimating the project should change and increase (Wu and Cao 2009).

At the time of the first VACs in the early 2000s Research showed, that some 75% of new information systems development projects could fail to meet their objectives at first attempt (Davies, 2000, the Standish group 1995, Boehm and Basili, 2001). This figure could prove even worse when focused on e-business systems alone. This is because developers do not realize the complexity of the project and its multidimensional and cross-functional nature. E-business projects are of the nature of building not just a single web browser application, but also the whole communication infrastructure between the participating organizations in a "value-added community", as well as completing the integration of the various existing systems. This in itself has the added burden of dealing with organizations that, before the integration, might have different priorities (Rifkin 2001). These differences in priorities lead to different levels of commitment and response to the requirements of the project. Such differences can only increase the complexity and add new dimensions as to the difficulty of estimating the effort and time required to complete it. Failure, to estimate the project correctly, or within reasonable degree of accuracy, may lead to delayed launch of the e-

business facility, or poor service to the customers and the value-added community, with potentially disastrous financial results.

Since the publication of the 1995 Standish Group Report there has been considerable improvement in project successes as the Standish Group reports in its 2011 "Chaos Manifesto" with 32% of projects completing successfully, 44% considered challenged and only 24% failing altogether (The Standish Group 1995, The Standish Group, 2011). However good an improvement (up from 25% in the original report) this might sound though, there is still considerable room for progress in terms of successful completions rate. In particular, a lot of developers view e-business implementation projects as primarily web site development projects and the business community is led to adapt to this view (Reifer 2001, Mendes 2008). This increases the risk of project failure beyond the degree that has been identified for traditional information systems, as project managers might be led to ignore or severely underestimate the effort and time required to integrate a large diversity of systems, that apart from the technical challenges they present might be implicated into issues of conflicting inter or intra organizational priorities (Bygstad, 2004, Miranda and Abran 2008 and Wu and Cao, 2009)

## Aims of the Research

Since the mid 1990's and the early days of the "dot-com" era electronic commerce and electronic business have evolved considerably. The "gold rush" of early e-business adoption, spectacular successes and heavy failures has naturally slowed down nowadays and has given way to more mature and organized business models (Laseter, 2011). Internet based technologies though continue to evolve the world wide web is moving to a new more interactive "semantic" format and the requirements for more online transactions continue to increase. As a result information systems are continuing to be interlinked and interconnected

in business alliances and e-business communities called Value Added Communities. With the evolution of technologies employed in such projects and the ever increasing need to link old (legacy) systems to new to offer enhanced online facilities at a much reduced time compared to a few years ago, the need for more accurate estimation of the completion of a project is ever more pressing (Cusumano and Hopkins, 2011).

This research aims at addressing this specific need, by focusing on the role of risk, its effect on e-business projects and its potential relationship to the estimation of the duration of an e-business project. More specifically this research aims at:

- Reviewing existing estimation models and techniques and identifying gaps in terms of suitability to estimation of e-business projects

- Identifying a risk list that is relevant to e-business projects and establish it validity

- Explore and verify patterns of risk variation with an e-business project size and complexity

- Explore the relationship between risk variation and project analogies

- Propose a modified approach to estimation by analogy method to estimate the duration of e-business projects

- Validate the proposed approach using cases of completed projects with known actual duration.

- Explore the potential of extending this approach to estimation to other types of information systems development projects.

## Contribution to Knowledge

Estimation of information systems projects has traditionally focused on establishing models that could describe the size and / or technical complexity of the project based on historical

data. Subsequently those models could be applied to projects similar in nature for which certain "dimensions" of the project such as the number of lines of code or function points could be either predicted from expert knowledge or calculated once the requirements analysis phase of a project was at an advanced stage (Jorgensen and Boehm 2009). The estimation models of this nature are used to estimate the volume of the work required, subsequently translating it into human effort and hence calculating the cost and in conjunction with a detailed plan the duration of the project (Buglione and Ebert, 2011). The success rates of such models have improved over the past four decades that such work has been constantly evolving, but large numbers of projects still complete well outside the predicted estimates. Despite the great progress in estimation methods and models over the past forty years, there is still room for improvement especially as new types of projects with new challenges appear in the information systems arena and their integration with the business world is becoming stronger than that of a few years ago (Jorgensen and Boehm 2009, The Standish Group, 2011).

The case of e-business and in particular the building of value-added-communities is one characteristic example where current estimation methods are not adequate enough to accommodate the needs for estimating such projects. With most such projects the success or failure of the project lies on public perception, often related to a specific launch day upon which expectation for a new service or an enhanced one is built (Wilcocks et al 2001, Starr et al 2003, Pimenidis et al 2002, 2004). This is still valid under a different perspective nowadays; that a successful launch could lead to the organizations behind the VAC achieving all their original objectives, both individual and collectively as a community (Birkinshaw et al, 2011, and Buglione and Ebert, 2011). Getting the project duration right is such an important element in managing an e-business development project.

Two very recent examples where an unsuccessful launch could still prove critical, despite the services involved being monopolies, were the revamping of the tax return system in Greece in 2010-11 and the online tickets sales system for the 2012 Olympic Games in London. In the first case with the country in a major financial crisis and in desperate need for cash injection, the tax return system was being revamped to accommodate some changes to the system and the addition of links to other government systems with a target date of the 1$^{st}$ of March 2011. This failed dramatically and had resulted in the government issuing a series of rolling extensions to the deadline for submissions that led to delays in tax payments of up to three months. The overall impact was millions of Euro in unplanned taxes as the need for more short term borrowing increased (TA NEA, 2011a, 20011b).

In the second case due to the site failing to respond to surging demand, processing of payments had failed on a number of occasions, with the process being delayed by more than a week in the first round of sales and with potential buyers missing out on opportunities in the small scale second round of sales (SKY.com, 2011, Howard 2011 and The Telegraph 2011).

Most literature on estimation still focuses on "web-based" projects, by only looking at the technical (software development) attributes of the projects and not considering the impact of interrelationships across the partners (Mendes 2008, Glass et al 2008, Hill 2005, and Cataldo et al 2009).

This research demonstrates a new approach to estimating project duration with a different focal point. Instead of trying to estimate the project detailed size (lines of code, function points, etc), emphasis is placed on risk and its variation with more general project dimensions. The author shows that risk can reflect not just the structural complexity (partly expressed by project size in traditional approaches), but also that of the interaction between all stakeholders of the project. The need to estimate the duration of the project at a very early

stage does not allow for detailed requirements to be elicited and therefore makes traditional estimation methods difficult and possibly inaccurate to use.

Risk based analogy, will overcome the issue of lack of detailed requirements as it considers the project as a whole. With further work the applicability of this concept could be applied to other project types, with similar characteristics in terms of development conditions and requirements.


## Overview of the thesis

In line with the research objectives stated above, this thesis proceeds by exploring the background literature on estimation models and techniques, reviewing the most well established models for software effort and cost estimation, in chapter 1. The author further explores the potential of such models being utilized for estimating e-business projects under the perspective of integrating systems across different organizations, with potentially conflicting requirements which shifts the main focus away from software development. To this effect more recent variations of the traditional models explored in chapter 1, with emphasis on web-based projects are reviewed in chapter 2 and a gap in the literature relating to the estimation of e-business projects is identified.

Shifting the focus on e-business development from software to business interactions and interrelations amongst participating organizations, risk appears as an area common to both the technical and business aspects of the project. Literature has identified that risk is strongly related to the complexity of a project and tends to increase as the size and complexity increases. Chapter three addresses this issue and explores the nature of risk in information systems projects, reviews literature as to risk and e-business projects and discusses tolls and methods for monitoring and evaluating risk in project development. Furthermore the author

explores how risk has been used in estimation methods in other industries and how lessons learned in such approaches can be filtered into this research work.

The presentation of experimental work, its results and evaluation of them begins at chapter four. Here the work towards exploring the role of risks in e-business projects, from identifying a valid risk list to establishing a pattern of variation of risk with project size and complexity is discussed. Chapter five presents the next phase of the research work with the use of two established tools for estimating projects, namely the ISBSG software projects database and the ANGEL estimation tool. Their use along with the rate of risk variation with project size and complexity in e-business projects is discussed in relation to a modified use of the estimation by analogy method for estimating projects. To validate the approach discussed in chapter five, four project cases are considered in chapter six. The approach to estimation by analogy using risk is applied to obtain estimates for the duration of those four projects, with the results compared with the actual duration achieved and the results are discussed and evaluated. Finally chapter seven draws conclusions from the work overall and suggests further avenues in improving and expanding the outcome of this research, improving the accuracy of the results obtained and expanding the areas of application with other types of projects suggested.

# CHAPTER 1

# Project Estimation

Project estimation is core to managing projects and is considered as one of the key activities that could define the success or failure of any project. Traditional information systems estimation focuses on estimating the size of the project from the requirements specification stage and proceeds by estimating human effort required. This could be further translated to cost by calculating the cost of human effort hours and adding to that the projected cost of equipment and materials required to complete the project (Hall 1998, Schwalbe 2010).

This chapter provides an overview of the different approaches to estimating information development projects and reviews and discusses some of the most well established models and techniques used in industry by professional project managers in estimating their projects.

## 1.1 Project Estimation Approaches

There are two major software estimation approaches in common use in the information systems industry. These are classified as Macro and Micro Estimation. Both of them are based on traditional principles, developed by the engineering practices over centuries of experience, of using historical data or expert opinion to base their estimations.

There are many models that have been developed in the relatively short history of information systems development projects and these can be grouped in four types of techniques, which in turn are assigned to one of the two estimation approaches mentioned above. The types of techniques that are usually encountered within the two approaches are expert judgement, analogy, decomposition and statistical or parametric methods (Hill, 2010)

Expert judgment is based on the brainstorming of one or more experts who have experience with similar projects. The attributes of those projects are compared to those of the proposed

project, assessing the median project productivity, delivery rate and speed of delivery for each attribute using a consensus mechanism that produces the estimate. There are a few advantages in the use of this technique, in that it can be objective, repeatable, verifiable, efficient and if used correctly can yield a good guide on the effort required to complete the proposed project. Weaknesses include the possibility that expert knowledge used might no longer be valid. For this reason it is recommended that experts used are well familiar with the environment / organization the project is developed for (Schwalbe, 2010 and Buglione and Ebert, 2011).

Analogy estimation is based on being able to find a completed project that is a very good match to your proposed project allowing for comparison of previous, similar activities, and analysis of the most relevant project and service attributes. The analogue project that is identified is used as a guide to allow the estimator to gauge the new project's effort and cost values through estimator experience. As with expert judgment, this technique requires skilled people who can properly understand and see relationships and implicitly evaluate qualitative and quantitative figures to determine possible clusters of projects. In the absence of experience suitable repository mining tools can be used. It shares the same advantages of as estimation by expert judgement, but practitioners could find difficulty in applying it should the pool of projects from which the analogue is drawn prove not well aligned to the new project's environment (Hill, 2010 and Buglione and Ebert, 2011).

Statistical (parametric) models are a set of related mathematical equations that have been derived from the analysis of large volumes of historical project data. The parameters in the equations are changed to match the known facts (or predicted requirements / attributes) of the project to be estimated. The project manager / estimator can create alternative scenarios by changing the parameter values where there is uncertainty as to some of the requirements of the project. Project managers use such models or parametric estimation tools to get a useful

indicative, or "ball park" estimate of a project's duration, effort and cost. The key strength of such techniques is the depth of the historical data used to derive the equations employed. However, project managers should always bear in mind that such models can be too imprecise for accurate estimation (Hill, 2010).

The above three technique types form the group that is classified as Macro Estimation approach.

The last technique type, decomposition or work breakdown is a bottom-up estimation technique that tries to make a granular list of initially planned tasks. In this technique the effort and duration of each component of the project is estimated separately and the results are aggregated to produce an overall estimate for the project.

Buglione and Ebert (2011) believe that the more granular the tasks associated with a certain requirement in a work breakdown structure (WBS), the closer the estimated effort could be to its actual value. This would be reducing the mean relative error and possible slippage in project deliverables. Hill (2010) though warns that such a technique can be subjective and it could prove optimistic. He further cautions that to succeed in yielding accurate estimates this technique requires detailed knowledge of the project's structure and extensive knowledge of the organization and development environment; something that might not be possible in today's fast track development project and distributed development team era.

From the above it is evident that for proper estimation there are two important ingredients, people and historical data. These are interrelated when estimating a project and people often prove the more important of the two. Most organizations lack historical data and to compensate for this they use large project data repositories available from specialist organizations. As stated above though, this might prove a weakness of the technique applied should the data prove not relevant to the domain area of the project to be estimated. This is the reason that a number of organizations estimate primarily through experience. This might

work if the project managers involved have the right level of experience and periodically measure and put their estimates versus actual data into a historical database which will empower the organization with the potential of more accurate estimates.

## 1.2    The COCOMO Model

The original COnstructive COst MOdel (COCOMO) was first published by Dr. Barry Boehm in 1981 in his book Software Engineering Economics. It is an algorithmic software cost estimation model which uses a basic regression formula, with parameters that are derived from historical project data and current project characteristics. References to this model typically call it COCOMO 81.

Its publication followed a development and trial period during Dr. Boehm's tenure with TRW from 1976-1979. During this period, the number of source instructions (called the equivalent delivered source instructions or EDSI), total development time, and the total effort necessary for 40 aerospace industry software projects were studied. From this information, the estimation formulas for COCOMO were developed and have been used to calculate estimates for the time of implementation and the amount of human effort required to develop the software for which an estimate is required (University of Southern California, 1994).

Boehm originally defined three levels of application, basic, intermediate, and advanced, based on the phase of systems development during which the model is applied. The basic level is applied early in the lifecycle, while the intermediate and advanced levels, which require more accurate information on cost driver inputs, are applied later in the lifecycle.

Subsequently it was discovered that the schedule and effort are influenced by certain factors related to the difficulty of the project. The level of difficulty (or familiarity) is classified into 3 modes: organic, semi-detached and embedded.

The Organic mode is used to calculate the effort for a project where constraints during the implementation phase are mild. Furthermore, the project to be estimated has been pre-dated by a number of similar projects that could assist in defining the agenda of development.

The Semi-detached mode is used for projects where the constraints are greater than in the organic mode, but there still remains some flexibility; the project may only be pre-dated by a limited number of similar projects (University of Southern California, 1994)

Finally, the Embedded mode is used for a project that has very tightly defined constraints and therefore cannot rely upon previous projects completed.

All modes apply to all three levels of the original model (Smith et al, 2001).

For each mode of effort estimation, the effort result is given in units of Person-Month (PM). PM is the number of months one person would need to develop a given project. The schedule estimation is given in the actual number of months needed for development by a properly staffed full-time development team.

Evaluation of the three available levels (Basic, Intermediate, and Advanced) leads to the conclusion that the Intermediate model provides significantly better prediction than the Basic model, while the Advanced model is not materially better than the Intermediate model (Boehm, 1981 and Smith et al 2001).

The COnstructive COst MOdel version II (COCOMO II) is the revised version of the COCOMO 81. This model was published in the Software Cost Estimation with COCOMO II book, by a group of authors led by Barry Boehm (Boehm et al. 2000).

The main objectives behind the development of the COCOMO II model were:

"To develop a software cost and schedule estimation model tuned to the life cycle practices of the 21st century.

To develop software cost database and tool support capabilities for continuous model improvement.

To provide a quantitative analytic framework, and set of tools and techniques for evaluating the effects of software technology improvements on software life cycle costs and schedules" (Boehm et al., 2000 and University of Southern California, 2000).

The full COCOMO II model includes three stages where, stage 1 supports estimation of prototyping or applications composition efforts, stage 2 supports estimation in the Early Design stage of a project, when less is known about the project's cost drivers and stage 3 supports estimation in the Post-Architecture stage of a project (University of Southern California, 2000).

## 1.3 Function Point Analysis

Function points were first introduced by Albrecht in 1979. Function points are intended to measure the functionality of a software system as observed by the user, independent of the technology being used. The advantage over other measurements used in the estimation of information systems management projects is that they can be calculated fairly early in the software development process, using the requirement and design specifications.

Calculating function points is part of a larger estimation method called Function Point Analysis (FPA) (Smith et al 2001).

Function points measure five characteristics of a software system, namely

- User Inputs – Unique user data or control that enters the external boundary of the system

- User Outputs - Unique user data or control that leaves the external boundary of the system

- Internal Files – Each major logical grouping of data in the application

- External file Interfaces – Files passed or shared between applications

- User Inquiries – Unique input that causes and generates an immediate output

In performing FPA, the five characteristics are weighted based on their value to the software customer. The weights for each characteristic are based on a complexity estimate (simple, average, or complex). The sum of the weighted function point count for each characteristic is termed the unadjusted function point count (UFP) (Pandian, 2004, pp. 18-19). Albrecht's original method was criticised for lack of support of further complexity. As a result the complexity factor was enhanced by the consideration of a set of fourteen general application characteristics which are weighted as to their applicability / relevance to the project. The sum of these weights is used as a factor to adjust the UFP count.

The final count of function points is obtained from

Function points= information processing size X technical complexity adjustment (Smith et al, 2001).

A full example of calculating FP and applying them for estimating a project is given in appendix D

Information Systems development organizations are gradually moving away from Traditional software size measures, such as Source Lines of Code, which have been the subject of much criticism. Instead, as an alternative, Function Points have been gaining wide popularity for estimating application size.

The advantages of Function Point Analysis (FPA) are that it can be applied early in the Software Life Cycle and that the calculations are objective. Furthermore UFP are independent of the technology used to develop the application and therefore can withstand the rapid changes in that area of computing. Finally, function point counting and their application is supported by an active, worldwide user organization, the International Function Point Users Group (IFPUG) (Galorath and Evans, 2006).

However, one of the major criticisms of FPA is its inability to address complexity adequately. This can result in a disproportional measurement of size, which in turn will affect both effort

and duration estimation, limiting the objectivity of such estimates (Hastings and Sajeev, 2001). This is a considerable disadvantage in the cases of e-business projects and in particular those of value added community type as they are characterised by high levels of complexity.

Galorath and Evans (2006), discuss a further weakness of FPA which is also relevant to this research, that of 'semantic difficulties'. The function point standards have been drafted in the 1980's based on traditional information systems concepts and terminology. Most applications nowadays are much more complex and less rigidly structured to allow for an accurate calculation of UFP.

To overcome such difficulties, there have been several Function Point extensions proposed, particularly targeting the issue of complexity.  Of these, MKII Function Points and

Feature Points are the most tested and accepted alternatives.

Of those the most successful one is the MKII approach. This requires calibration that may prove difficult for specific application types when there is little or no history. Such a difficulty is something that could affect its applicability to modern types of applications that have no actual history of development but when successful tend to attract wide attention and the numbers of new projects increase considerably in a very short period of time (Hastings and Sajeev 2001).


## 1.4    Estimation by Analogy

Analogy based estimation is a technique for early life cycle macro-estimation.

In simple terms, when estimating by analogy the process involves finding one or more projects that are similar to the one to be estimated and then deriving the estimate from the values of these projects. Estimation by analogy can, for example, be performed as

• Pure expert estimation, where the "knowledge base'' comprising previous projects is in the expert's head.

• Expert estimation informally supported by a database containing information about finished projects, which can be consulted on occasion when the experts feel they might need to cross check their own memory / knowledge of previous projects,  or

• Estimation based on the use of a clustering algorithm to find similar projects, using a repository of projects that have been defined as cases (Jorgensen et al, 2003 and Scwable, 2010).

In this last case the key activities for estimating by analogy are the identification of a problem as a new case, the retrieval of similar cases from a repository, the reuse of knowledge derived from previous cases and the suggestion of a solution for the new case. This solution may be revised in the light of actual events during the project's life cycle and the eventual outcome could be used to update the repository of completed cases.

This is not as straightforward an approach as it might appear, as it poses two fundamental problems. First, how are the different projects cases characterised and classified?

Second, how will similar cases be identified and retrieved? In other words how will similarity be measured?

Characterising the cases depends on the pragmatic issue of what information is available?

Generally two types of variables are used. These are, continuous, i.e., interval, ratio or absolute scale measures or, categorical, i.e., nominal or ordinal measures.

When designing a new Case Based Reasoning system to support estimation by analogy, experts should be consulted to try to establish those features of a case that are believed to be significant in determining similarity, or otherwise, of cases available in the project repository (Shepperd and and Scofield 1997).

In the case that it is not possible to define enough criteria that would determine the similarity of the project, one would have to consider the hybrid approach to estimation by analogy

where the project repository would serve as supporting material to an expert human estimator.

In view of the above challenges in terms of establishing similarities, the following steps should be followed in estimating project effort by analogy:

1. Establishing the attributes of the planned (target) project, then measuring or estimating the values of those project attributes. Such attributes are:

Software size

Business area type

Maximum team size

Development type

Application type

User base – locations

User base – business units

User base – concurrent users

Primary programming language

Language type

Use of a methodology

Use of a CASE tool

2. Searching a repository of completed projects for a project that closely matches the target as a source analogue to compare against.

3. Using the known effort that was used in developing the source analogue as an initial estimate for the target project

4. Comparing the chosen attributes for the target and source projects

5. Establishing or adjusting the initial effort estimate in light of the differences between the target and source projects.

Advantages of estimating by analogy:

It is useful where the domain is difficult to model and it is quite dynamic in terms of requirements shifting with time or evolving business models. It is known that many factors influence the effort required to complete a software project; however, it is less known how these factors interact with each other, or how best to model the wealth of factors via software metrics. Estimation by analogy can be used successfully without having a clear model of how effort is related to other project factors. It relies primarily on selecting a past project that is a close match to the target project, rather than assuming a general relationship between effort and other project characteristics that applies to all projects.

It can be used with partial knowledge of the target project, an attribute that favours its use in estimating projects in their early stages.

It can avoid the inaccuracies of equation based model use, as models are often based on rather dated historical data.

It has the potential to mitigate problems with outliers, while it offers the chance to learn from past experience and to update knowledge by incorporating new project data into the repository, as soon as the project has been completed and its data validated.

Drawbacks of Estimation by analogy:

- The availability of an appropriate analogue, this might not always be possible depending on the repository, the types of projects data stored and whether this data is relevant and or current.

- The soundness of the strategy and the accuracy of the process for selecting the analogue project.

- The manner in which differences between the analogue and target are allowed for when deriving an estimate (Hill 2005, Hill 2010, and Gruschke and Jorgensen 2008).

## 1.5 An Overview of Estimation models

Over the past thirty years there has been a continuous debate amongst researchers and practitioners as to which is the most dominant project estimation technique, i.e. more frequently used and the one producing more accurate results; formal model-based or expert-judgment-based?

Formal effort estimation models, such as COCOMO and function-point-based models are based on a mechanical quantification element such as a formula. On the other hand, judgment-based estimation methods, such as work-breakdown structure-based methods, are based on a judgment-based quantification step, i.e. what the expert believes is the most likely use of effort to complete an activity. Furthermore, judgment- based estimation processes range from pure "gut-feelings" to structured, historical data and checklist-based estimation processes. The difference between models and expert judgment isn't always clear.

In a published debate Jorgensen and Boehm (2009) attempt to demystify the "myths" surrounding the different estimation techniques and analyse their differences, strengths and weaknesses.

They argue that the belief that models are more objective when compared to expert judgement is not true. This is due to the fact that in order to derive the models essential input is required on expert judgment that codifies and validates the historical data used.

A further "myth" is that judgment- based effort estimation is a "black-box process" which can degrade to a contest between experts, not necessarily adding value to the accuracy of the estimates obtained and making it very difficult to improve on those estimates. On the contrary, there are many ways to improve judgment-based effort estimation which are mostly specific to the type of project and the development environment used. Similar views have been previously documented by Keung et al (2008) and much earlier by Stamelos and Angelis (2001) which discuss the potential of improvements to estimation by analogy

estimates for specific cases and focusing on improving the sensitivity of matching suitable analogues to the project to be estimated.

A further argument that more advanced (complex) estimation models are more likely to lead to substantially more accurate effort estimates, is also refuted by the two renowned researchers. They question the belief that models such as the intermediate version of the COCOMO models or a neural-network-based estimation model will likely be more accurate than much simpler models. In their view in software engineering, as with most forecasting disciplines, the perhaps most stable result is that simple models typically perform just as well as more advanced models and that more sophisticated models are vulnerable to "over-fitting" to possibly dated or unrepresentative data sets. In such case, the improvement potential of the model side will consequently be lower when compared to that of judgement based estimation. In earlier work, Jorgensen (2004) had discussed the importance of specific domain knowledge (case – specific data) claiming that it is higher in software development projects than in most other studied human judgment domains. He then proceeded to propose a step-wise approach to improving the accuracy of estimations, aiming to address uncertainty in software development. The need for such approaches is also supported by Pfleeger (2008).

- "Evaluate estimation accuracy, but avoid high evaluation pressure.

- Avoid conflicting estimation goals.

- Ask the estimators to justify and criticize their estimates.

- Avoid irrelevant and unreliable estimation information.

- Use documented data from previous development tasks.

- Find estimation experts with highly relevant domain background and good estimation records."

In supporting expert estimation processes, Jorgensen further suggested a bidirectional approach to estimation, by estimating both top-down and bottom-up, independently of each

other. He also proposed the use of estimation checklists and the need to assess the uncertainty of the estimate. This last point was further pursued in by Laird (2006) and by Gruschke and Jorgensen (2008).

In concluding, the majority of the researchers agree that expert estimation is the dominant strategy when estimating the effort of information systems development projects. Furthermore there is no substantial evidence that supports the view of the superiority of model estimates over expert estimates. There are situations where expert estimates are likely to be more accurate. In these situations experts have important domain knowledge not included in the models or situations where simple estimation strategies can provide accurate estimates (Gruschke and Jorgensen 2008). This last point is also supported by the work of Mendes and her colleagues (2001a, 2001b, 2003, 2005, 2008).

Likewise, there are situations where the use of formal models may prove useful in reducing large situational or human biases that could easily distort the accuracy of an estimate by expert judgement (Jorgensen 2004).

# CHAPTER 2

# Web Based Projects

Web based applications and their relevant projects are quite distinct from traditional information systems development ones.

Hill (2005, p.66) claimed that most aspects of web development are not distinctly different to the development of a client-server development environment. He bases this claim on the fact that web based systems development involves significant effort directed towards database design, implementation of business rules and business objects and developing interfaces to other systems. If web based projects were to be limited to the above attributes only, it makes perfect sense to apply equation (parametric) based models to estimate those standard software development tasks and create a cumulative figure for the estimated effort.

Reifer (2000) though, had claimed that the then still evolving web development practice confronted project managers and the estimation practitioners with new challenges, precisely because it introduced new modes of working (shorter development times, extensive reuse, etc). He then raised the need for new metrics and new models for estimating web based projects, simply from the development effort perspective.

Web based projects though in their majority go beyond the scope of simple software development. This chapter presents an analysis of the challenges facing project managers in estimating web based development projects.

## 2.1    Value Added Communities

The evolution of the Internet and the World Wide Web have led to radical changes in the design and establishment of new businesses, as well as the reshaping of existing ones since the late 1990's. New models of doing business have emerged and new rules are defining the

way business is conducted, goods are bought and sold and most importantly define locations where work is to be shifted and carried out in the future (Brynjolfsson and McAfee 2007).

One particular type of business that is mostly suited to, and affected by, such new models is e-business. Venkatraman (2000) and Tapscott (2001) were amongst the first to discuss such trends and had predicted that they will continue to develop further. New technologies that have emerged since then have enhanced the capabilities of collaboration considerably and are contributing to redefining how organizations shape, develop and evolve through complex collaboration schemes. Web 2.0 technologies will have even more dramatic impact on how information is shared, aggregated and interpreted across online communities (Birkinshaw et al 2011).

A decade ago e-businesses were seen as varying considerably from other forms of business principally because they approach the business world in a holistic way that redefines the boundaries between organizations, information sharing and business partnerships (Tapscott, 2001).

Following the collapse of the "dot com" era, researchers were expecting business models to reach and maintain a state of advanced maturity, with research focusing more closely on the implementation of the e-business infrastructure (Christiaanse 2003, Papazoglou et al 2000).

To survive as a standalone player in the e-business marketplace is almost impossible and this has led to the evolution of Value-Added Communities (VACs) (McWilliam, 2000, Venkatraman, 2000, Andal-Ancion et al 2003). These are groups of companies, usually complementary within the supply chain, linked together through the use of computer networks. This electronic network offers maximum return for the community as a whole achieved through the establishment of communicating computer systems that support the key activities of each of the participating businesses across the supply chain. Customer demand is used as the empowering input for all the above systems. Through the electronic business

facility of the trading organization (brand-owning company), information is processed, filtered and forwarded through the relevant networks to other computer systems such as MRP, MRPII, ERP, that each may support the function of one of the members of the community. Thus planning and coordination of activities within the community can be performed according to evolving market trends and continuously revised on a real time basis. Internal cross-business coordination leads to the restructuring of supply networks with more supplier integration and a proliferation of product customization, business complexity and uniquely defined customer relationships (Kopczak and Johnson 2003, Dyer and Hatch 2004). Figure 1 below shows a block-diagram structure and the communication flows of a VAC.

Developing E-businesses and, in particular, under the above concept of VACs involves exposure to a considerable level of uncertainty and risk, arising from the need to integrate a number of different business functions, belonging to different organizations, which may be linked to diverse and conflicting objectives, or differing levels of commitment to the evolution and functioning of the VAC (Smolander, 2003). Complexity and uncertainty are bound to increase further as current research explains. E-business models and the way businesses continue to collaborate and interact are changing faster than before responding to the enhanced capabilities of new technologies. This means businesses must cultivate agility that is the ability to adapt quickly to, or even anticipate and lead, change. Such change would reflect to the whole online community in which a given business is actively collaborating (Cusumano and Hopkins 2011).

The concept of the Value-Added Communities might appear dated in the fast changing web driven world. On the contrary it is well in line with current business strategies and philosophy. Laseter (2011) claims that the only way businesses will survive and continue to be competitive is not by growing in scale, but by growing with focus on their unique

capabilities. This is the exact philosophy of VACs and this is the characteristic that dictates the need for collaboration.

Although, this research focuses on the systems elements of a VAC one must not ignore the business elements of the concept. It is these elements that differentiate the development of e-businesses and VACs in particular, but introducing additional complexity and uncertainty to the project.
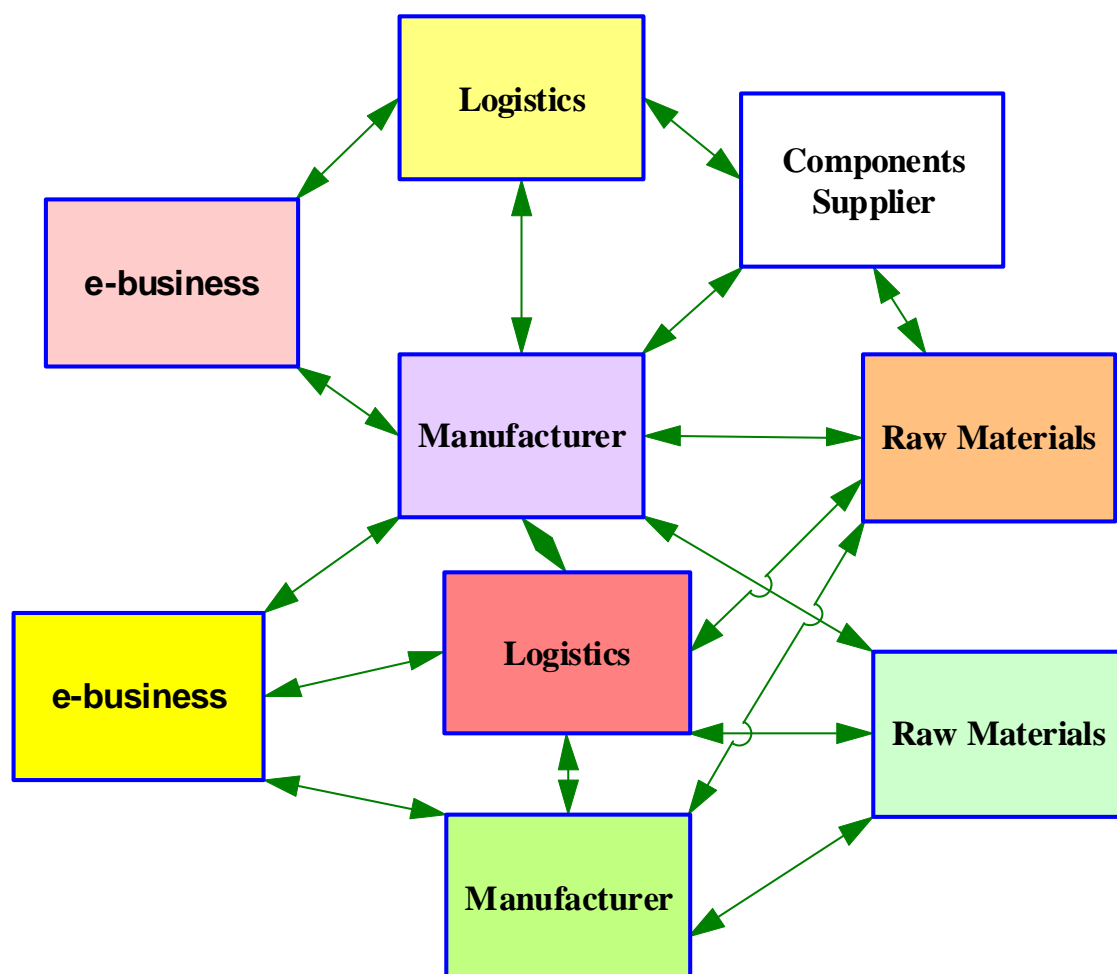


Figure 1. A schematic representation of a VAC

The different blocks in figure 1 above represent different companies (and their respective information systems). The green arrow-headed lines represent bidirectional communication in exchanging information. There is no particular colour code in terms of the colours of the

blocks in the diagram. Different colours are used to differentiate between companies. Where more than one block bears the same name, it is to show that a VAC can accommodate more than one company of the same type. These might be competitive to each other but in the confines of the VAC share the load of the business volume according to predetermined rules that govern the functioning of this partnership. These same rules drive the operation of the systems that plan the activities within the community.

## 2.2    Estimating Web Based Projects

Web software development requires the performance of tasks significantly beyond those performed in traditional projects like client-server development ones. In traditional systems development there is very little interaction between different business functions, like marketing and supply chain management with software consultants and developers, with a lot of projects loosing focus and drifting away from the original requirements. Web based projects though require the collaboration of all areas of an organization. They integrate supply chain management systems to customer relationship management tools and automate a lot of functionality that crosses in to the area of responsibility of more than one business functions. Furthermore, all this is usually driven by compressed deadlines dictated by marketing functions that focus on business objectives rather than the rational development of a complex system project (Hill 2005, pp. 66-67).

To address this shortfall in web based estimation projects various researchers have focused on adapting existing estimation techniques and metrics to measure software size and hence estimate development effort for web based applications. This is by no means an easy task as has been clearly documented by Reifer (2000). At the time Reifer had stated that the key problem was that existing methods had been developed by studying historical data over the previous twenty years, validating models and refining those on the basis of a large amount of

data with considerable stability. At the time web development projects had been outdating the then existing software methodologies at a rapid pace and thus creating a new breed of software (Boehm and Fairley 2000).

One consistent contributor to research outcomes relating to estimation of web based projects over the past ten years has been Mendes and her colleagues who have focused primarily in cost estimating web application projects.

One approach attempting to estimate the effort required to complete web development tasks was to use measurement principles to evaluate the quality and development of existing Web applications. Mendes et al (2001b) expected to be able to understand and potentially make improved predictions about web based software applications. It was hoped that obtaining early feedback from developers could assist the estimators to improve and correct their early estimates during the execution of a project.

To seek more accurate estimations at an early stage of a web based project Mendes et al (2002) employed Case Based Reasoning as an estimation technique, concluding that CBR "is a candidate technique to effort estimation and that with the aid of an automated environment it is a practical technique to apply to Web hypermedia development effort prediction".

While work has specifically focused on early project size measurement to enhance effort prediction CBR was further exposed to comparison with more traditional techniques. Amongst the conclusions is that the quality of estimates would vary with the quality of the dataset used for the estimation. The more successful estimates appeared to be those coming from an own company dataset, but then the size of such datasets might be limited and consequently its usefulness could be for that organisation alone.

In expanding their research Mendes and her colleagues pursued the accuracy of estimates obtained primarily with Case Based Reasoning in combination with estimation by analogy. This is compatible with other researchers who have suggested that a combination of

techniques could be applied to strengthen the estimate (Jørgensen and Shepperd 2007, Gruschke and Jorgensen 2008). Mendes et al (2005) concluded that estimates using model based on an organization's own historical data, are no better to those compared to estimates obtained from a cross-company model, or estimations based on a median effort.

Mittas and Angelis (2008) tested a Combination of Regression and Estimation by Analogy in a Semi-parametric Model for Software Cost Estimation, but despite seeing some improvement on initial estimates there was not enough stability to the results to confirm this method. There have been no further results published from this angle of research.

In a more recent effort Mendes and Mosley (2008) have explored the used of Bayesian Network models to estimate web development effort. The models were applied to given datasets and were benchmarked against simpler expert estimation techniques using the median effort. The results show that the more complex models based on Bayesian Networks did not improve the accuracy of results while their application was more resource intensive.

## 2.3   The Gap in Estimating VAC

As seen in the previous two sections of this chapter most of the methods used to manage an information system development project date back to the 1980s and are suitable for projects where the emphasis has been placed on the estimator's ability to establish a measure for the software size. As the problem domain of a system becomes less clearly specified, the degree of difficulty in managing the project increases. This is particularly true for software development in the 21$^{st}$ century as analysed by Boehm (2008) who explains that nowadays software developers and consequently project managers face serious challenges of having to simultaneously deal with rapid change, uncertainty and emergence, dependability, diversity, and interdependence.

This is particularly true for an e-business project where the conditions are much more volatile when compared to more conventional systems and action takes place in a much more dynamic field. Developers and contributors to the project might be dispersed in diverse geographical locations. Demands on the system change continuously along with the continuous evolution of the business model and the development of business-to-business relationships. In addition, development times have to be much shorter than those for traditional information systems if the business is to capture the market before its competitors (Cloyd 2001, Rajlich and Bennett 2000, Lee 2005, and Buglione and Ebert 2011). Failing to understand such differences, results in unrealistic cost estimates and estimates of project completion times offered by the developers, at a scale far greater than has been encountered with traditional systems development projects.

Boehm (2008) argues that in the case of complex systems where uncertainty is high due to conflicting requirements, it is important to manage stakeholders' expectations. This can only be achieved when software engineers and project managers are knowledgeable not only about software concepts and techniques but the concepts and techniques of the organizations using the software.

With e-business development all of the above is true and in particular with VACs. The way forward for such organizations is for each participant to focus on their own capabilities for growth (Laseter, 2011). In doing so individual stakeholders will increase the diversity of requirements and will induce more conflict in priorities. This will in turn impact on the integration phase of the development of the online community, with new systems linked to existing and legacy ones, increasing the level of uncertainty further. Lam and Shankararaman (2004) advise that in the case of an enterprise integration project (such as a VAC one) that a project manager should estimate resources based on the integration project's complexity, the

type of integration work, and the skills required, placing the emphasis on the complexity which should be the driver in estimation.

Collaboration in complex projects though is often hindered by compatibility, connectivity, and security issues across multiple organizations with different IT infrastructures. Communication is costly and often considered as "overhead" to be minimized. Thus to support collaboration among multiple stakeholders in a project and to control the impact of uncertainty Wu and Cao (2009) propose a three step approach; explore the past cases (experiences),find a similar case, and reuse the solution for the past case in the new problem situation.

E-business development projects definitely include some web based project elements. However, they are not solely web software development ones. A large amount of the effort and time required to complete such a project is consumed in managing stakeholder requirements and conflicts, integrating systems to a common communications and data processing infrastructure and tackling uncertainty due to the large scale and diversity of focus of the participants (Andal-Ancion et al 2003, Birkinshaw et al 2011).

The efforts in estimating web based projects reviewed in the previous section of this chapter point towards software size estimation as the first and core target of the researchers. This may prove valid and adequate for pure software projects involving the integration of a database to a website application, but are not capable of reflecting the effort required to manage and overcome the uncertainty and conflicts that characterise an e-business development project. As explained in the first section of this chapter and documented by Hill (2005, 2010), the key driver in an e-business project is the business sector. As such when estimating (in particular in the early stages) without much clarity about requirements the target is to identify a bulk estimate of the completion time. To do so the manager has to consider the project holistically as suggested by Mendes and Mosley (2008). Thus software sizing is not an appropriate

metric for estimating an e-business project. This author proposes to use risk as the key driver in estimating the duration of such projects.

# CHAPTER 3

# Risk and Projects

Risk is present in every aspect of life. Anything that has a specific objective assigned to it and the attainment of that objective could be disrupted by an unwelcome event is susceptible to risk. In simple terms, risk can be defined as the probability (or likelihood) of failing to achieve particular cost, project objectives, and the consequence of failing to achieve those objectives. Risk is inherent in information systems development projects. It needs to be managed to be controlled if a project is to execute with minimal disturbance (Aladwani 2002, Keil et al 2006). To be able to manage risk a project manager needs to be able to evaluate it at every stage of the project. To do so, project managers use ordinal risk assessment scales. These do not represent probability but "uncertainty" by evaluating separately the level of uncertainty in a project and the level of consequence of the impact (Conrow and Shishido 1997). Uncertainty can be defined as a combination of three elements, namely project size, project structure, and technical newness (Aladwani 2002).

## 3.1 Risk and Information Systems Projects

The role of risk in projects and in particular in information systems development projects is not a new concept. Its study and evolution is parallel to that of software engineering, with many renowned researchers in the field of software project management having pursued the concept of risk evaluation and risk management. Boehm (1991) had identified that project postmortems indicated that successful project managers had almost invariably used some project management approach in their way to a successful completion of their projects. He further proceeded by proposing a six step risk management procedure comprising, risk identification, risk analysis, risk prioritization, risk management planning, risk resolution and

risk monitoring. Boehm subsequently proceeded by proposing a risk taxonomy based on his assessment of risk uncertainty and risk impact using probabilities and thus calculating the cost of software failure. Since then his six step management process has been adopted by a large number of practitioners and researchers, but his calculations of risk cost on the basis of probabilities has often been replaced by the use of ordinal scales as proposed in Conrow and Shishido (1997).

At present, many information systems developers have adopted detailed and heavily process-oriented methodologies in an effort to control the effects of risk and so reduce delays and the number of failures in the projects they undertake.

Invariably, the first step is that of identifying those elements of a project or those events (external or internal) that could possibly cause the project to be late or fail. Following Boehm's 1991 taxonomy, many others have proposed and refined risk taxonomies that reflect systems development as it has evolved in the present and these can be used as guides to project teams in the identification stage of risk management. These taxonomies are lists of areas or activities within a project that could potentially yield a risk consolidated from project histories (Bandyopadhyay et al. 1999).

Most such lists or taxonomies however are limited as to their use since they primarily focus on internal factors, ignoring many external elements of influence such as politics, changing business requirements, development platform deficiencies and so on. In the risk identification phase teams must also be aware that risks change with time, something that early researchers like Boehm (1991) and Conrow and Shishido (1997) have identified but has been consistently ignored where spectacular failures are encountered (Keil et al 2006). Furthermore new risks arise that the team has not planned for (Murthi 2002).

Conrow and Shishido (1997) present a list of "software risk issues" grouped in six areas namely project level, project attributes, management, engineering, work environment, and a

general one termed as "other". It is this notion of "other" the unknown or constantly varying risk area that is often the most critical in terms of identifying and managing as emphasis shifts with time and new conditions arise (Letens et al, 2008). Keil et al (1998) had acknowledged the fact that in order to develop meaningful risk management strategies the relative importance of the various risks needs to be established. This would be ideal if risk were a static concept. They also suggested that a full risk identification process should also focus on explaining why certain risks are perceived to be more important than others. With risk being a very dynamic concept such classification has to be revised on a continuous basis, under the risk monitoring step proposed by Boehm. Keil et al (1998, 2006) proceed by identifying a different taxonomy of risk factors which again is limited to internal factors within a software development project. The same limitations are experienced in the model for accurate status reporting proposed by Snow and Keil (2001). Here, the difference in risk exposure of various projects is considered in order to assess project management reporting bias.

## 3.2    Risk Assessment in information systems projects

According to Addison and Vallabh (2002) the failure rate of software projects at the time has been proven to be very high, and the incidence of failure was showing an increasing trend as more companies venture into software development. Risk management as a process was consequently defined as the use of methods aimed at minimising or reducing the effects of project failure.

Similarly Zafra-Cabeza et al (2008) define risk management in a project environment as the systematic process of identifying, analysing and responding to uncertainty as project-related events or conditions which are not definitely known and which have the potential of adverse consequences on a project objective. In any risk management process, risk assessment is the

key element that can potentially protect the project manager and the project from considerable delays and decaying effects.

Tiwana and Keil (2006) in investigating the role of functionality risk review literature on studies of risk in software. They address the issue of information integration in forming risk perceptions. Risk is assessed through individual manager's judgement used to establish the perceived level of risk. Such judgements are based on heuristics and the systematic integration of information about a variety of characteristics of a given project. These individuals use cognitive models through which they weigh each attribute that can influence risk perceptions. All this information is integrated into an overall assessment of the project's risk perception. The same authors argue that, the lack of required knowledge in the project personnel along with the introduction of new technology are major risk factors. Prior related experience in similar projects helps develop a cognitive representation of a largely intangible investment (such as software) which reduces the level of perceived risk associated with developing it. This leads into the conclusion that the higher the related technical knowledge in the domain of the project, the lower is the managerial perception of overall project risk. The emphasis is placed on the project domain itself and not the specific technical areas of the project as managers would consider the risk to the project holistically. The above agrees with the way risk is considered in this work and supports the way it is utilized in the estimation approach proposed here.

Risk is usually assessed by the combination of probability and impact. The mathematical expression of risk in the "risk formula": Risk=probability x damage seems to indicate that risk is an objective entity (Hall 1998). However, both aspects of the risk formula are not objective entities but rather unstable social constructs. Probabilities depend on the observer and that person's past experiences or his / her ability to project historical data into the future which does not induce any objectivity in this attribute of risk.

Damages associated with risk are not objective either. The idea of objective damages directly implies that they should be measurable something that is not always possible, even if in some cases there are methods and formulae available to monetarise them (calculate the amount of money required to rectify the damage) (Stahl 2007). This is true in the case of VACs where the damage cannot be calculated in monetary terms as it is often related with loss of reputation or loss of an opportunity in capturing a slice of a niche market.

Stahl further comments that the identification of risks is usually done by developing a list of risk factors through expert knowledge. In this approach risk is implied as being objective and the approaches focus on the concept that managers identify risks that are there independent of them.

The list of identified risks should be comprehensive as unidentified risks can become a major threat to the organization or result in significant opportunities being missed.

In an attempt to address this concern, many organizations either expand the standardized checklists of possible risks that very often are a formal part of standards or guidelines, or refer to separate risk lists that exist in literature. Literature review shows that these classifications of risks, the so-called "risk taxonomies," are structured in several different ways. Some risk taxonomies list risks according to the sources of the risk, e.g., making the distinction between environmental risks, political risks, and economic risks. The various taxonomies illustrate that risk lists, whether as a separate tool or as a part of a standard or guideline, very often seem to stress particular risks more than others, or even omit certain types of risks (Tesch et al, 2007). Consequently, some risks may not be identified in early critical stages of a project therefore, may be excluded from further risk analysis and risk management and may affect the planning and estimation phases causing considerable disturbance in later stages. Thus Letens et al (2008) argue that project managers should

consider expanding those project lists based on their own experiences and their own fact gathering techniques for the benefit of future projects.

While various risk checklists (e.g., the ''top-10'' list of risk factors described by Boehm(1991) and frameworks (e.g.Keil et al 1998) have been proposed, the underlying dimensions of the software project risk construct and their influence on a project remain largely unexplored. A better understanding of the dimensions of software project risk and the trends or patterns that they are likely to follow in different types of projects could help project managers formulate more specific risk management strategies by allowing them to focus on areas that are at potentially high risk. Wallace et al (2004) claim that, earlier efforts did not attempt to examine the ways in which the dimensions of risk vary across different types of projects. While the specification of risk and measures may allow managers to audit risk levels, it does not provide them with information to help formulate a tailored strategy for countering the risks on a specific project.

To this effect the above authors propose exploring the differences between low, medium, and high risk projects as a means of focusing managerial attention on recurring patterns. This could in turn be used to offer insight into the relative trends in risk dimensions that could in turn enhance managerial understanding of the nature of vulnerability of a specific project.

The authors maintain that project characteristics also impact the risk level, and proceed to investigate three of them: project scope, the degree to which the project is of strategic nature, and whether it is outsourced. These three were chosen because they have been previously proposed in the literature as factors that may affect project risk levels. They study them as to the effect they might have on the following six dimension of risk, namely team risk, organizational environment risk, requirements risk, planning and control risk, user risk, and complexity risk . They conclude that the most prominent risks associated with high risk projects differ from those classed for medium and low risk ones. For high risk projects,

requirements risk, planning and control risk, and organizational risk are the most prominent risks, whereas for low risk projects complexity is the most prominent. This is somehow in contrast to Aladwani's (2002) findings which consider complexity a major element of uncertainty and therefore a key risk contributor in all projects and not just the "low risk" ones.

Similar to Wallace et al (2004), Boehm and Turner (2003) had considered five elements of agility and have studied how the assessment of risk factors relating to the above elements could dictate a suitable project structure to achieve the objectives of an agile driven project.

Risk is an ever present feature of projects and in times of economic crisis, as the one experienced on a worldwide scale over the past few years, emphasis is placed on improving on risk assessment and potentially minimizing the impact and losses due to information systems development risk. On this note Kulk et al (2009) propose to apply statistical methods for quantifying the impact of factors that influence the quality of the estimation of costs for IT-enabled business projects. These factors are termed as risk drivers since they influence the risk of the misestimating of project costs. The above authors claim that their method can be effortlessly transposed for usage on calculating the effect on other important information system key performance indicators (KPIs), such as schedule estimation or functionality delivery. They claim that to do so project managers have to have at their disposal "decent quantifications". They transpose methods used in perinatal epidemiology to the world of information systems to address the issue of misestimating project costs. They proceed to demonstrate how with quantification of risk, it may be possible to quantify the expected return of a portfolio of projects. The expectation in this case is that with an ordering of projects by increasing risk, investments in the need of management attention will surface and that audit attention will optimally be allocated to address those needs.

They further believe that by quantifying the information systems risks for an entire project portfolio it becomes possible to quantify the aggregate expected return of this information systems projects portfolio, and thus it becomes more straightforward to assess whether it is of value to the organization investing in this portfolio. This is a rather isolated view of information systems projects ignoring the impact of interactions with the external environment as this has been considered by most of the previously quoted researchers in this work.

## 3.3 The Role of Risk in Estimation

The consideration of the role of risk in the estimation of information systems projects is not new to research literature. A long list of distinguished researchers consider risk as being inherent in projects and they have shown risk to have some role in managing projects, with some venturing into the consideration of risk impact on estimation. E-business project and the development of VACs in particular are not an exception to this. This author believes that risk can be used as a definitive element in estimating the duration and consequently the completion of a VAC development project.

One of the first to highlight the need to link risk to project estimation was Kansala (1997) where he stated that "Risk assessment estimates the needed contingency due to the impact of anticipated risky events. Armed with this information, the project manager can confidently commit to the aggregated result of cost estimation and risk assessment".

More recently though, Donaldson and Siegel (2007) argued that, the traditional planning approach in information systems projects often lacks an explicit allocation of resources to reduce risk. To improve on this they proposed a five step approach to risk assessment, comprised of; deciding on the number of risk levels, defining risk criteria for each risk level, defining the number of matches required to assign the project to for each risk level, define the

default risk level (in case the matches are insufficient to assign one), and decide on recommended resource allocation percentages for management, development, and product assurance for each level. The authors claim that projects with more risk will demand more risk reduction resources. They propose an approach to estimation whereby percentages of the budget are assigned to product assurance and management. These percentages can be increased as the project risk increases. Such an approach although valiant does not directly contribute to the main aim of this work which is to estimate the project at the outset, as the above approach aims at introducing corrections to the estimate at various points of the project and such corrections may have considerable impact on the original estimate.

Earlier, Laird (2006) had identified one of the shortcomings of estimation methods as not accommodating risk and its assessment in estimating information systems projects and discussed how this might benefit project managers and other practitioners.

In recent works researchers have focused more sharply on risk and its role in estimating projects (Glinz 2008, Gruschke and Jorgensen 2008). Boehm and Bhuta (2008) warn project managers that risk assessment should be integral to all phases of any project, including that of planning and estimation, or they could face considerable rework and overruns. Furthermore, Jorgensen and Boehm (2009) accept that not all models are objective and conclude that research should focus more on improving and refining judgment-based methods of which risk assessment is a considerable element. The above agree with the earlier work by Laird (2006) who, in discussing the limitations of estimation, highlights the fact that risk and its management are often ignored in planning and estimating and concludes that risk consideration in estimating could lead into considerable improvements and potentially minimize inaccuracies in estimating systems development projects.

Approaching the problem of estimation in a different industry, Zafra-Cabeza et al (2008) use risk metrics to forecast the final duration of a project and its cost estimate at completion.

Their method selects a set of risk mitigation actions to be undertaken in order to reduce risk exposure. They use the mitigation actions as the manipulated variables while cost and time are the controlled variables. Modeling techniques are used to demonstrate how decisions as to when risk mitigation actions should be taken, to reduce the impacts of the possible risks that could affect the project. These authors propose that their method can be used for rescheduling the project depending on the time at which mitigation actions are required. The work is applied to semiconductor manufacturing projects, but the author of this work believes that the principle is also relevant to VAC projects. Semiconductor development projects are very structured, process driven projects, but highly prone to risk and especially changing risks throughout the duration of the project. Thus the risk element is significant in both types of projects. At the same time though, one would have to consider the need for accuracy of the initial estimate for VAC projects. This is essential to support other non software development activities integral to the project such as marketing plans and actions, physical resources reallocation and relocation, etc. Therefore any approach that would consider revised estimates through reviewed risk levels during the duration of the project is not going to fulfill the requirements of estimating an e-business project.

Finally Kluk et al (2009) in discussing their work of risk quantification claim that their approach could benefit project managers in that it could be possible to identify risk drivers and manage new projects based on the right values of the risk drivers with a positive influence on the correct estimation of information systems KPIs. They aim at identifying the risk drivers that lead to cost misestimating. Knowing the reason for project misestimating attributed to the identified risk drivers, better estimates can be obtained in the future and this in turn will support effective investment decisions that could potentially lead to more stable projects.

# CHAPTER 4

# Exploring the Role of Risk in e-Business Projects

## 4.1    Experimental Method

Value Added Communities are primarily information systems development and integration projects and as such the risks encountered during a VAC project would not differ in nature to those encountered in large and complex information development projects. What could prove different though are how such risks and their overall effect on a VAC may vary as the size and complexity of the VAC vary. It was therefore natural to start the research into risks relating to VACs from established risk lists that have been developed studying large and complex systems projects. Those, the author expected to be able to modify by adding elements relating to e-business and more specifically VAC development projects.

In pursuing the above, the following three-phase approach was followed (Addison and Vallabh 2002, Benaroch and Goldstein 2009):

•       Literature sources were used to identify relevant risk lists and taxonomies. From those, a preliminary list of elements that included technical issues but mostly related to business or integration issues was compiled. The preliminary list contains some entries that might appear as duplicate risks, but came from different sources and were maintained in this list to allow for more clarity in the discussions with the experts. This is shown in table 1a, in appendix A.

•       This preliminary list of risks was then refined by incorporating the views of four experts which provided a first level of verification.

•       Finally a questionnaire was used to experimentally test the proposed list. The data collected was used as means of validating the proposed risks through a wider expert panel and at the same time rank those risks as to order of significance.

In the case of this research, the experts involved came from a variety of fields across the business and local authority sectors.

Local authorities were considered as valuable inclusion as they offer a particular fortuitous advantage in that they are all facing e-business pressures in their continuous quest for rationalization of resources and cost cutting and therefore have agendas similar to those of commercially oriented VACs. These parallel developments offer greater statistical stability and perhaps greater predictive power (Pimenidis et al, 2004).

## 4.2    Establishing a Risk List

The author's first step towards proposing a risk list relevant to VAC projects was based on Aladwani (2002) who considers project uncertainty focusing on the three variables namely, size, complexity and diversity of the implementation project. Thus risks relating to VAC development were sought along those three axes. This approach was combined with the proposal by Addison and Vallabh (2002) whereby risk factors have been categorized according to project manager experience, to add a further variable. Finally the work by Reifer (2002) which compares traditional to Internet/Intranet development risks was also consulted. This proved useful in confirming that traditional risks such as those primarily relating to the technical issues of an Internet related project are still relevant. Personnel shortfalls, volatile requirements and new methods and unstable tools are the key risks identified by Reifer. The addition by Reifer is compatible with risks identified for large and complex projects in earlier work by Boehm (1991) and confirmed by the work by Costa et al (2007).

Murthi's (2002) list of risks extending beyond traditional sources helped identifying issues relating to the integration of systems within the bigger spectrum of the value-added community. The work by Keil et al. (1998) that discusses the areas of potential conflict and has been of considerable help in particularly when combined with that of Wallace et al (2004)

who focus more closely on risks relating to systems development in the Internet era and the effect these could have on the successful delivery of business systems. Finally Lippert and Zullighoven (2002) provided useful suggestions as to differences in approaches, tools and methodologies used in e-business development as opposed to the implementation of traditional information systems, thus identifying a further area of risk relevant to VAC projects.

This above preliminary list was further complemented, confirmed and reviewed through interviews with four project managers with considerable experience on e-business development projects. The four experts between them have experience of implementing VAC projects in 7 different countries in Europe and the U.S.A. From those one came from the banking sector, one from the manufacturing industry, and the other two from retail. Two of the interviews were contacted on a face-to-face communication and the other two over the telephone. The results of the interviews were used to reconsider the draft list obtained through literature review and the final group of project risk factors was reduced to the following seven, listed in alphabetical order:

1.     Difference in readiness of partners to function on the e-business model.

2.     Different priorities in terms of launch time

3.     Inexperience of developing team.

4.     Legacy systems not compatible with modern technology.

5.     Loss of expert resources (members of the development team).

6.     Low commitment of individual partners.

7.     New systems to be integrated but not previously tested.


## 4.3     Validating the Risk list

Following the compilation of the above list, a questionnaire was constructed and distributed to sixty nine project managers or sponsors of VAC projects. The purpose of this questionnaire was to collect data that would validate the proposed risk list. The respondents were either people well known amongst practitioners for their experience in having managed projects relevant to this research, or were identified from research in the field. The latter were contacted prior to them receiving the questionnaire to verify their suitability as respondents.

The questionnaire was distributed electronically, as an attachment to emails sent to potential respondents. This proved to be an efficient method of data collection resulting in a 50% completion rate. A considerable number of the eventual respondents were initially contacted by telephone and some via email in order to explain the purpose of the questionnaire and possibly assess the suitability of the potential respondent. This was considered as an essential step in ensuring that data was obtained from sources with relevant expertise, as these are still quite scarce.

The target in devising the questionnaire was to classify the risks identified according to their impact and to assess their probability of occurrence, thus providing a means of ranking the risk factors identified. The works by Fehlman (2002), Willis (2002), Sherer (2004) and Keil et al (2006), providing the business perspective in terms of risk factors in e-commerce were particularly useful in this part of the work. Furthermore the work by Letens et al (2008) discussing practices of risk identification and assessment in engineering practice provided additional insight and helped in confirming the soundness of the approach.

The questionnaire was initially distributed to 44 businesses and 25 councils. Of those contacted 21 businesses and 14 councils responded by completing the questionnaire, yielding a return rate of just over 50%.

A preliminary analysis of the results has been attempted on a qualitative basis. The reason for this is dual:

1.      The sample is not large enough to allow solid conclusions as the data in terms of experience levels is quite scattered and

2.      It might induce bias in the next stages of the research where more data is sought out to validate / endorse these preliminary results.

The second point raised above is confirmed by Heemstra et al (2003) who argue that individual views or perceptions of risk and their effects are highly prone to bias. The above two reasons are confirmed by Persson and Mathiassen (2010) when discussing the development of risk mitigation plans in distributed teams, something that is very much part of the reality of e-business development projects.

To provide a more meaningful representation of the respondents' perception of the above list of risks the results were quantified by assigning a frequency and severity weighting of 1 to 4. This follows the approach in quantifying the impact as a combination of frequency and severity to accentuate the significant observations as suggested by Willis (2002), Kan (2002) and Pandian (2004).

The Frequency weighting of 1-4 is assigned in the order of Extremely Unlikely – Frequent, while the Severity weighting of 1-4 is assigned in the order of Negligible to Catastrophic. To produce a first feel for an overall rating of each of the risks in this list the weighted sum of each was calculated according to the following formula:

Rating $= [ \sum_{i=1..4} [(i*nfi)]/Nf ] * [ \sum_{i=1..4} [(i*nsi)]/Ns ]$

where

nfi = number of frequency i values

Nf = total number of occurrences for frequency

nsi = number of severity i values

Ns = total number of occurrences for severity

The results obtained are shown in tabular form in table 1 below which has been sorted according to the overall rating achieved for each of the proposed risks. The overall ratings obtained suggest that the two most critical risks amongst this list are: "New systems to be integrated but not previously tested" and "Inexperience of development team", while all others follow close behind. Although simple, this first set of results confirms that all risks identified in the original list are seen as significant and so qualified to be included.

Table 1. Risk Ratings – First Phase

| RISK | BUSINESS | COUNCILS | OVERALL |
|---|---|---|---|
| New systems to be integrated but not previously tested. | 9.39 | 7.29 | 16.68 |
| Inexperience of development team | 8.42 | 9.18 | 17.50 |
| Loss of expert resources (members of the development team). | 9.04 | 5.90 | 14.94 |
| Different priorities in terms of launch time | 6.49 | 6.40 | 12.89 |
| Legacy systems not compatible with modern technology. | 4.39 | 6.59 | 10.98 |
| Difference in readiness of partners to function on the e-business model. | 7.05 | 4.44 | 11.49 |
| Low commitment of individual partners. | 4.74 | 7.71 | 12.45 |

The individual ratings achieved from the business and local authority sectors are shown separately above. The data from the individual councils were broadly similar one to another

and so differences between the two groups may be related purely to context, e.g. the financial position of local councils compared with the commercial sector.

A larger survey during the next stage of this research was intended to allow the confirmation or realignment of this list and hopefully lead to a proposed taxonomy of risks for VAC projects that will complement existing taxonomies (Cho and Park 2002, Addison and Vallabh 2002).

The first five risks listed above are mostly generic to large-scale and complex projects (Houston et al 2001, De Oliveira et al 2004), however, the author was seeking to clarify whether these risks might have enhanced significance in the context of VAC development projects.

The results obtained have been normalised and are shown in table 2 below. This shows the risks in the order they appear in table 1 with the normalised rating being that for the overall value as shown in table 1.

Table 2. Normalized Risk ratings (wi=i)

| RISK | Normalised Rating |
|---|---|
| New systems to be integrated but not previously tested. (A) | 0.52 |
| Inexperience of development team (B) | 0.55 |
| Loss of expert resources (members of the development team). (C) | 0.47 |
| Different priorities in terms of launch time. (D) | 0.40 |
| Legacy systems not compatible with modern technology. (E) | 0.34 |
| Difference in readiness of partners to function on the e-business model. (F) | 0.36 |
| Low commitment of individual partners. (G) | 0.39 |

To accentuate the significance of extreme ratings (e.g. "catastrophic") the ratings calculations were repeated with wi = $i^2$. Figure 2, shows the revised results which have also been segregated to distinguish the business and local government sectors. The order of the risks on the x-axis is arbitrary and has been selected to differentiate the patterns obtained by results from the business sector compared to those from councils.
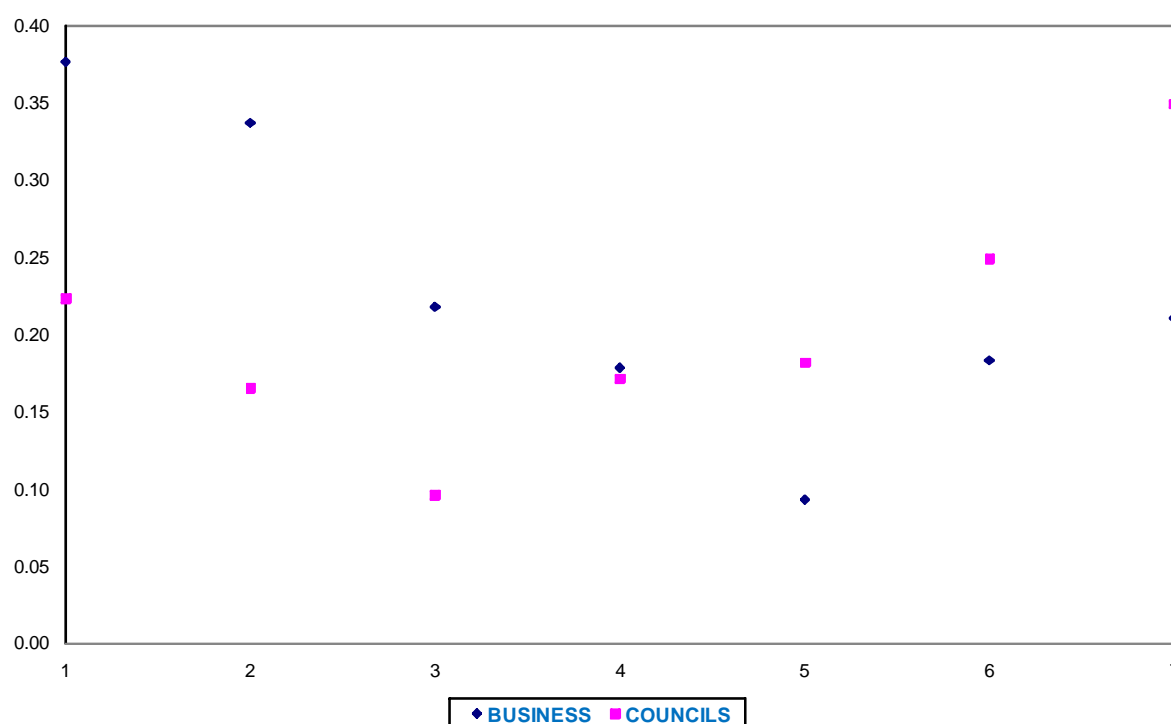


Figure 2. Graph of Normalized Risk Ratings ($i^2$) – First Phase

Points 1-7 on the x-axis correspond to the risks shown in table 2 in the following order B, D, G, F, E, C and A.

The different values obtained for the two sectors demonstrate the differences in perception. In the business sector the risk of having an inexperienced team is considered as the top rated one. The perception in the councils sector is much different with the risk of having new systems integrated into the VAC without any prior testing achieving the highest rating. Such differences should not cause a great surprise as the needs and priorities are quite distinct within the two sectors. For the business sector the emphasis is on the successful launch on the

publicised date and any diversion from that could be catastrophic. The involvement of an inexperienced team is likely to induce delays and these must be avoided at all costs (Starr et al 2003, and Tesch et al 2007). Councils on the other hand normally have less financial resources available per project in comparison to business led VACs. Usually the funding for such projects at local authority level comes from special projects which are closely scrutinized re their costs against objectives and any major rework at a later stage of the project might have a severe impact on the cost and the success of the project overall (Ferro and Molinary 2010). Thus they perceive the risk of integrating previously untested new systems as the most dangerous one; being closer to the main constraint.

There is one instance where the two values for risk F are very similar and the interesting point of this coincidence of perception is that it involves the risk that is least likely to be affected by the size or the complexity of the project. Differences in the readiness of partners to embrace the e-business model are primarily a business issue that is perceived at the same level of risk for both groups of organizations. This is the reason the rating is almost identical for the two sectors, regardless of the differences in perception as to the rest of the risks. Legacy systems not compatible with new technology appear to be perceived as less disrupting to the business environments, while the local authorities appear to be less worried about the commitment of individual partners.

Once confidence in the initial list of risks was instilled, the author sought to extend the survey in an attempt to draw solid conclusions as to the order of significance of the risks and their behaviour with varying properties of the project. A second phase of the questionnaire survey yielded a much higher sample size (65 respondents in total). The results obtained were very similar to the ones obtained from the first phase of the experiment as is demonstrated in Figure 3 below which shows the revised risk versus ratings graph that includes the two sets of data combined.
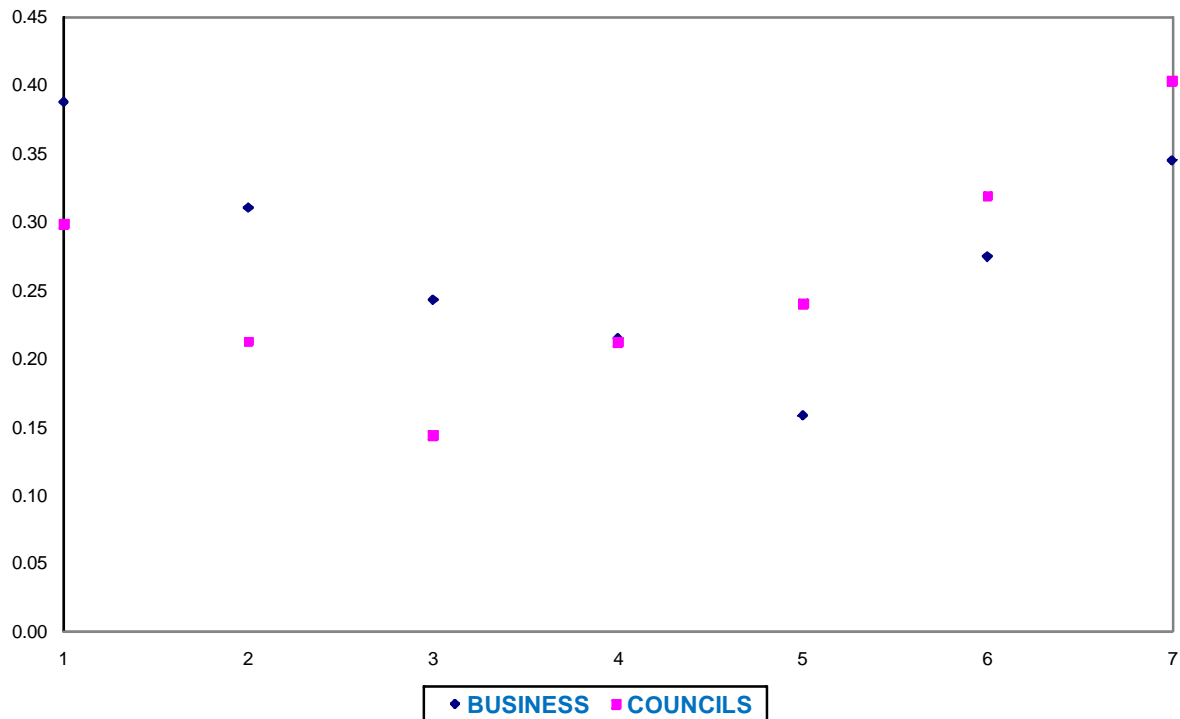
Figure 3. Graph of Normalized Risk Ratings ($i^2$) – Combined data from the two phases

The two graphs shown in figures 2 and 3 are very similar, although the individual values differ with the graphs in figure three occupying positions higher as to the vertical axis. The actual differences in values obtained are of little significance at this point of the research when compared to the similarity of behaviour, a fact that strengthens the confidence in the significance of the experiment and the ranking of the risks identified. As the source of the data is individual perception of risks affecting VAC projects it could be inappropriate to apply any analytic statistical processing method. Jørgensen and Moløkken-Østvold (2004) suggest that there are good reasons to believe that individual responses to qualitative questionnaires could be affected by the position / role of the responder and the way that data collection was completed. Although their work focuses on assessing software estimation errors the principle on data collection applies here too and the differences in the actual scale over the two graphs can be explained. With the selection process applied to potential

respondents the author had tried to ensure that respondents to the survey would come from a similar background in terms of experience and level of position To demonstrate the similarity a further graph figure B1 is shown in appendix B this shows the curves obtained the results from plotting the results of the two phases of the above research, as well as the result of combining the two sets of data over a single graph. Despite some small differences in the inclination of the curves the similarity of their behaviour can be further verified there.

The resulting ratings as obtained from individual observations were plotted in two graphs (figures 4 & 5). The expectation was that if individual data behaves in a manner similar to that demonstrated for the weighted average values (figures 2 &3) this would further strengthen the validity of the experiment and the confidence in the emerging ranking of risks specific to VAC projects. The two graphs in figures 4 and 5 show the top three individual risk ratings for each of the two sectors. These are the most significant and the ones that have attracted most observations of the two sectors investigated here. The figures of 1, 0.56 and 0.32 correspond to the highest three normalised ratings that can be achieved with the chosen scale for $i^2$. Thus the top rating of 256 $(4 \times 4)^2$ yields a normalised value of 1, the second highest rating $(4 \times 3)^2$ yields a normalised value of 0.56 and the third highest $(3 \times 3)^2$ a normalised value of 0.32. The figures on the vertical axis of the graphs show the number of occurrences each of the three values has been recorded in the responses to the questionnaires, for each individual risk.
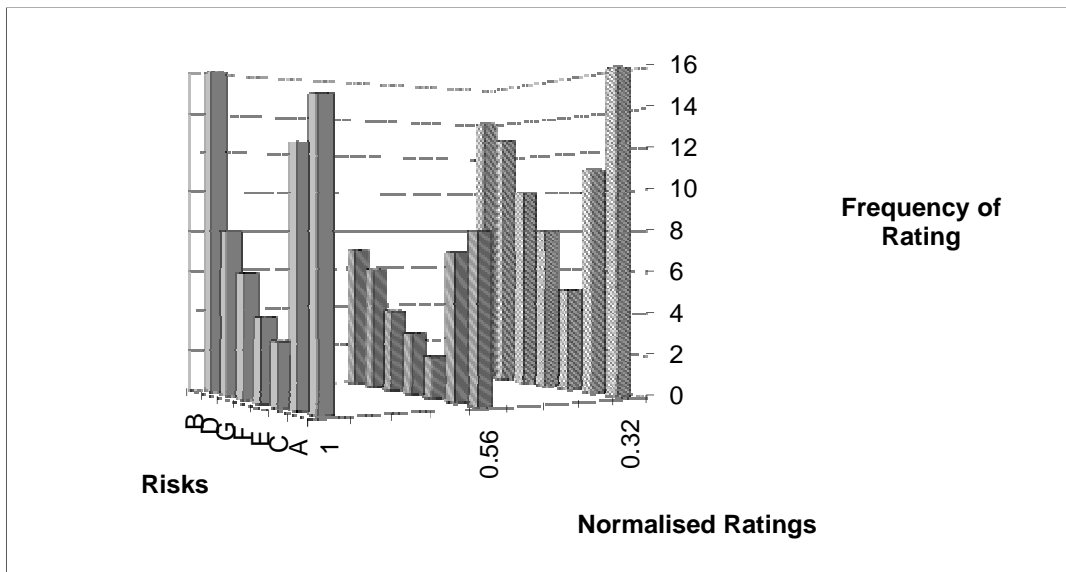
Figure 4. Frequency of Normalized Individual Ratings per Risk (Business Data)
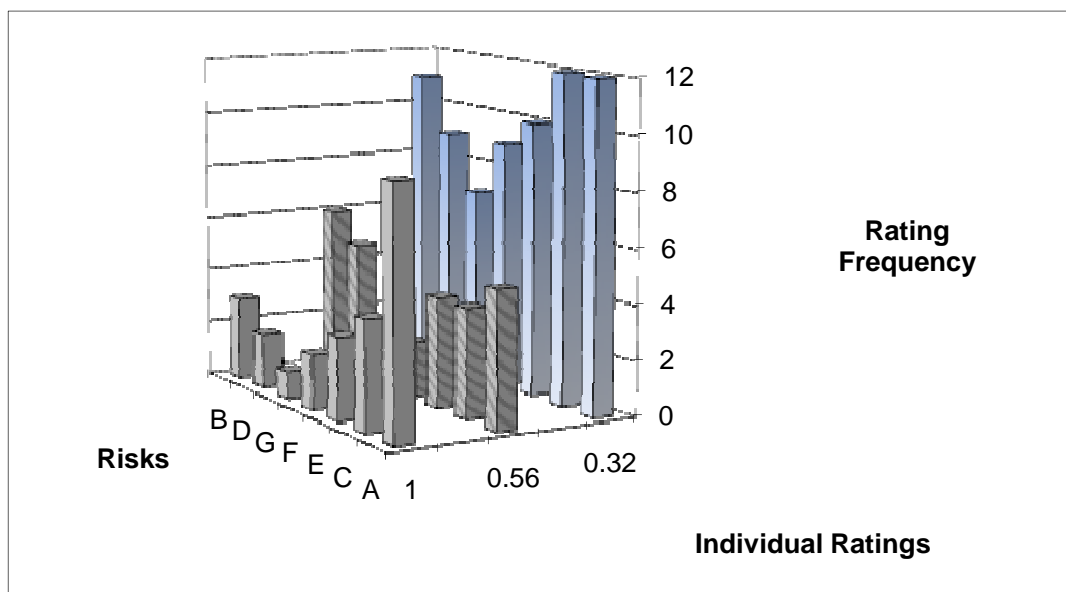


Figure 5. Frequency of Normalized Individual Ratings per Risk (Councils Data)

Figure 4 shows that there is a strong resemblance between the pattern of change of column height for each of the three rows for rating values 1, 0.56 and 0.32 as viewed across the seven risks and those of the curves for business values shown in figures 2 and 3. Similarly figure 5, demonstrates a strong resemblance to the patterns of figures 2 and 3 for the councils data. This is more evident for the front two rows (the higher ratings) while the pattern is somehow

different (with much larger number of observations) for the third rating. However, the variation of column heights shows a low point for risk G as all others and follows the general trend identified in all the others. The higher concentration of observations for this rating can justify the generally lower values of ratings for the left hand side of the graphs for councils in figures 2 and 3.

Therefore the individual observations largely conform to the patterns of variation across risks identified by the weighted average ratings, thus establishing the stability of the data and confidence in the results.


## 4.4    Establishing a Risk Variation Pattern

In previous work Risk ratings were plotted against dimensions of VAC projects to assess how these ratings varied as project dimensions increased (Pimenidis and Miller 2005). At the time the risk ratings had been plotted against the number of different types of systems and against the total number of systems integrated in a VAC project. Those early results showed that risk increased as those two elements of the project increased in numbers, thus confirming the relationship between risk and project characteristics such as size and complexity. At the same time those results had indicated that there was some difference in the behaviour of risks versus size or complexity, in the two sectors (Business – Councils). At the time the author felt that the sample size was not sufficient to explore and firmly support such conclusions. The second phase of the experiment had resulted in a greatly expanded sample and with the stability of the data confirmed this hypothesis was revisited and a further sample of graphs was published in Pimenidis and Miller (2006) to corroborate the earlier results. The full set of graphs is shown in figures B2-B9 (Appendix B). The data gathered from the first two phases of the experimental work has been processed to give weighted risk ratings for the various

risks shown in table 1 above. In the case of the graphs shown in appendix B, more detail describing the size and complexity of the projects is presented. The variation of risk rating as captured from the responses to the questionnaire has been plotted against the number of different systems integrated per project, number of development teams involved, number of types of systems integrated and number of organizations involved as participants in the VAC. The above four project attributes have been identified as key during the interviews with the experts at the start of the research and have been included in the questionnaire. Table 6 presented in appendix B shows average values of these four 'size' and 'complexity' attributes for the projects which the respondents to the questionnaire have been involved.    In earlier phases of the work the author had found that these risks vary linearly with each one of the four project attributes mentioned above (Pimenidis and Miller 2006). Figures B2 to B9 present the plotting of the individual risk rating for each of the four project attributes and with the data separated into Business projects and Councils (local government) projects, confirming the validity of the earlier indication of linear variation. The usefulness of the variation of individual risks is rather limited and a further set of graphs was produced showing the average rating of risk as it varies with increasing numbers of the four project attributes discussed above. Figures 6 and 7 show the variation of average risk ratings with increasing number of systems, development teams, types of systems and organizations involved in a VAC project. In each case a linear trend line has been applied to each plot and the resulting linear equation is shown for each cluster on each of the two graphs.

These linear equations will later be used to calculate the analogy amongst projects used to estimate the duration of a of VAC development projects, applying a modified estimation by analogy method as proposed by the author.
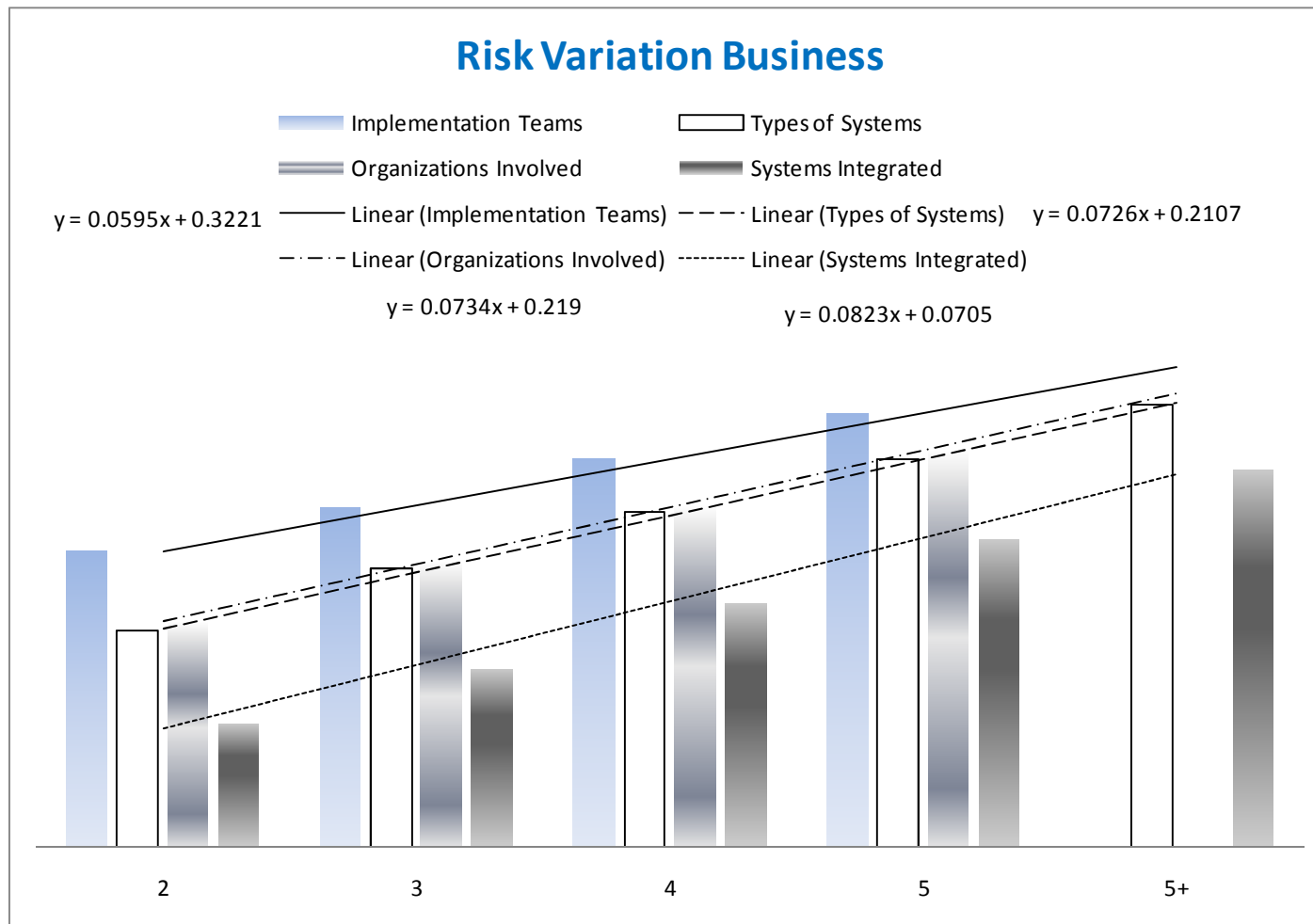
Figure 6. Risk Rating Variation with four project attributes (business data)
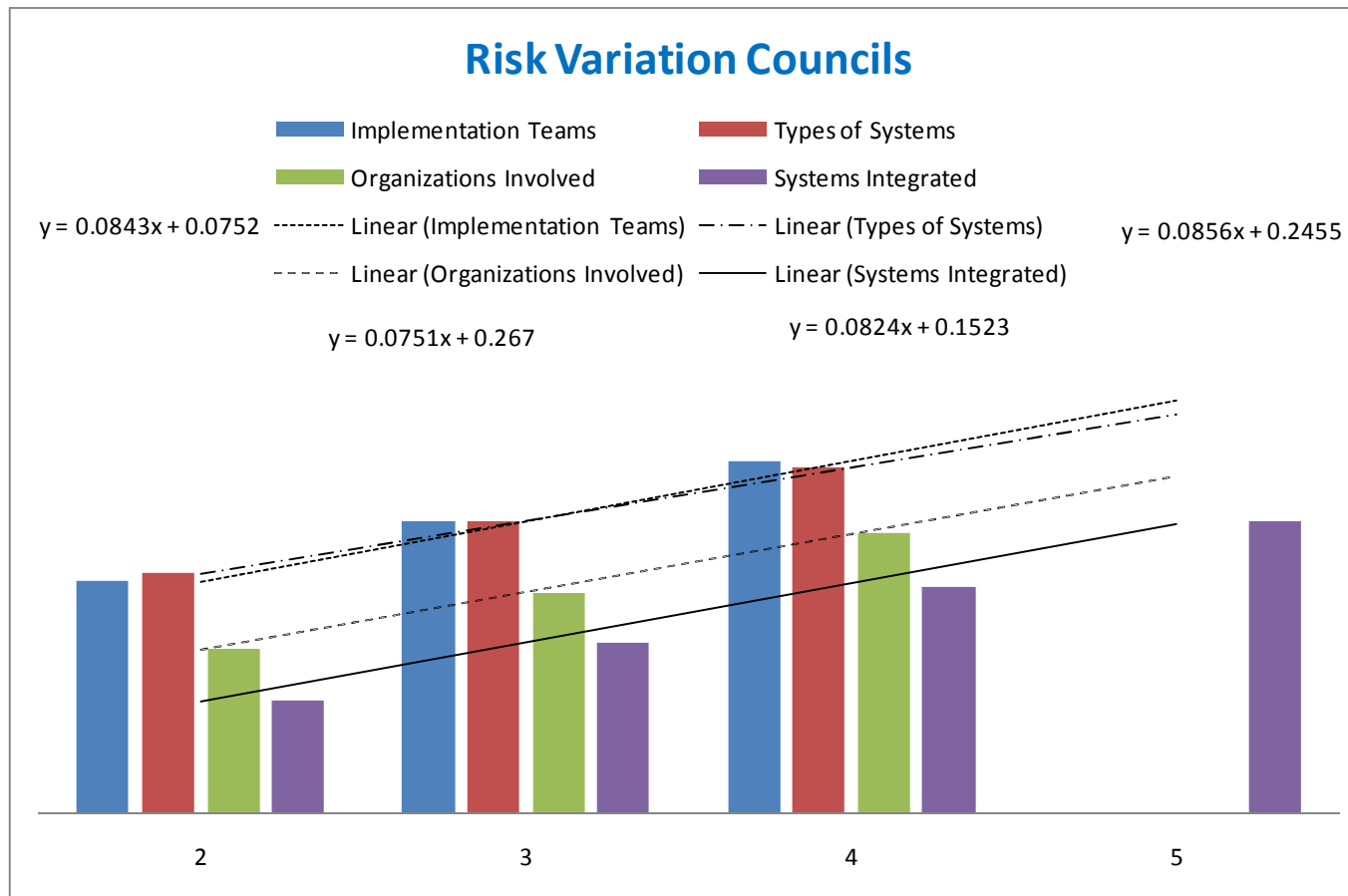
Figure 7. Risk Rating Variation with four project attributes (councils data)

The author believes that this is in line with the general trend of projects being in the public sector tend to move at a slower pace as opposed to those in the business sector (De Saulis, 2011). Also the public sector tends to take more cautious steps attempting to integrate subsystems first and slowly building up to the full VAC community adding partners and systems in incremental steps (Aagesen and Krogstie, 2010 and Castelnovo, 2011). This in itself leads to smaller projects as shown in figure 7. The incremental approach is partly dictated by the need to be cautious with public money spending, the lack of major pressure to complete within a short period of time dictated by market forces and the relatively smaller cash flow capability of local authorities. As these results confirm the initial results of the first phase of the experiment, the author believes that projects in the two sectors should be considered separately. This agrees with Lee D. H-D. (2005) who argues that e-government initiatives (VAC projects in the councils sector) have a distinct set of business value sources different from those of e-business projects. The difference in the behaviour of risks for the two sectors is particularly important in the next step of this work which uses these results in estimating VAC projects.

# CHAPTER 5

# Using Risk to Establish Analogies

## 5.1 The ISBSG Repository of Project Data

The International Software Benchmarking Standards Group (ISBSG) has established, grows and maintains a database of software project data that can be used by software project managers, IT managers, CIOs and IT customer business managers.

According to Buglione and Ebert (2011) a very important requirement for an estimation tool or method is to have the opportunity to run benchmarks against best-in-class projects and browse within such data. This is the reason all major tools have recently used the International Software Benchmarking Standards Group (ISBSG) history database (www.isbsg.org), one of the most renowned public repositories of information systems project data that is continuously maintained and updated.

The ISBSG database (Release 9), under license from ISBSG for research purposes, was used for this experiment. The database contains data on more than 3000 completed projects from different parts of the world. These are software development projects and are of different types of systems. The ISBSG data comes from projects submitted by organizations from nineteen different countries, with more than 65% of the projects having been completed in 2000 or later. The data is rich in detail, but not all projects contain the same level of detail. The quality of data has been checked by the ISBSG before the entry being allowed into the repository made available and therefore all data is useful.

Data details available vary from project to project and to avoid any inconsistencies in the validation experiment that followed in the research work, the author chose to limit the extent of the repository to be used only to those projects for which data was available for all the factors that were included in the questionnaire. Furthermore, the author selected those

projects for which original estimates, actual delivery dates and project duration were available, as all these elements are directly relevant to this research work. Hill (2010) confirms that there are 222 projects which will satisfy the above criteria and all those were included in the original selection. To fine tune the experiment and to minimise any effects from mismatches due to differences in the type of project, a further refinement yielded a final total of 116 selected projects that were used for this validation experiment. The refined list of projects was produced by excluding all those projects which despite being web based projects, could not be clearly classified as VAC ones. The excluded projects were lacking the level of complexity inherent in VAC projects and involved either just one organization or just one development team.  The above process follows the suggestions by Jørgensen and Grimstad (2008), who argue that a key source for inaccuracies in estimation is irrelevant and misleading information. They further explain that such information can impact effort estimates because estimation is based partly on unconscious cognitive processes where estimators frequently do not control their use of information when estimating effort or providing subjective input to formal effort estimation models.  A similar issue was also discussed by Zhang and Sheth (2006) who consider incomplete information received by managers as one of the key reasons for project mismanagement and estimation is at the core of project management activities. They see this incomplete information processing as one of the main reasons why projects fail and advise that project managers should actively seek to verify the completeness of the information upon which they make decisions and act.

The projects were divided in two groups those from the business sector and those from the government sector to allow for more representative projects to be identified as analogues for each of the two categories of projects to be estimated. Although the database includes various details that might relate to the size of the project in terms of function points or lines of code,

the selection of data that was used for the group of projects utilized was focused on the following four attributes that are relevant to this research.

• Number of systems integrated,

• Number of type of systems integrated,

• Number of organizations involved,

• Number of development teams involved.

Further attributes added were those of the duration of the project, categorical attributes defining whether the project is a new development, and enhancement or a re-development of an existing system, and finally the type of industry (in the case of business projects) that each project was developed for. These further attributes would contribute to a closer matching of the sought after analogue to the target project which this method aims at estimating. The two data models comprised 74 records for the Business data and 42 records for the Government projects data. Details of the construction of these data models can be found in appendix C.

## 5.2    The Angel Software Tool

The ANGEL software tool was developed at Bournemouth University in the late 1990's as a project estimation tool. It aims to allow the creation of one or more databases (known as data models) as collections of software development data such as effort, size, duration etc. Data is stored on a project basis (although it is feasible that units smaller than projects, such as phases, might be collected and stored). The collection of historical project data allows ANGEL to create estimates of effort, for new projects, by a technique known as estimation by analogy.  ANGEL operates upon the principle that predictions of effort should be based upon the most similar completed projects for which effort is already known.  This approach has the advantage that the system will automatically adapt to the local environment of the

user without the need for calibration. In addition, the reasoning behind any estimate can be easily understood since it will be based upon a list of most similar projects (Shepperd and Schofield 1997, Shepperd et al 1996).

It is the capability of the ANGEL tool which allows the user to define own data models / databases that the author is exploiting in this work, utilizing the selection of projects from the ISBSG database as the source for this work. The tool is used to select the best match to the project to be estimated from the data model used. This is called an analogue for the project sought to be estimated. The selection is based on the attributes of the project listed in section 5.1 and relating to the project attributes upon which the survey questionnaire had prompted respondents to rate risks relating to e-business projects.

Although the ANGEL tool is primarily used to estimate effort based on project dimensions such as function points, it is well suited to working with other attributes of project size such as the ones discussed in section 5.1 above. The system has been used in various experiments by renowned researchers, including (Shepperd and Schofield 1997, Jorgensen et 2003). Version 2.02 of the tool has been used for this experiment (available to download from the University of Bournemouth web site). Screen shots of the setting up of data models and loading of the selected project data from the ISBSG database to create data models can be found in Appendix C.

## 5.3 Using Risk with the ISBSG Data

Following the creation of the data models, populated with the ISBSG data, ANGEL was used to identify analogues for four projects for which data were collected through interviews with the project managers. The projects and the project managers involved in those four cases are unrelated to the ISBSG database and have also not been involved in the earlier phases of this research work. Two of the projects involved commercial applications in Greece while the

data for the other two projects originate from local government projects in the UK. In each case, to identify project analogues the attributes identical for the two projects (the one to be estimated and the selected analogue) are used to verify the matching. Once the analogue is established these identical attributes play no further role in the estimation process. The non matching attributes were used to establish the analogy (the measure of the difference) between the two projects, expressed as a ratio. Once this was established, it was simply multiplied by the actual duration of the analogue project to provide the estimated duration of the project under observation. To establish the analogy, the risk ratings for each of the differing attributes were calculated using the equations obtained from the corresponding graphs in figures 6 and 7 in chapter 4 and shown in table 3 below.

Table 3. Equations used for the calculation of risk based analogy

| Project Attribute | Business Projects Equations | Councils Projects Equations |
|---|---|---|
| Number of systems integrated | $y = 0.0823x + 0.0705$ | $y = 0.0824x + 0.1523$ |
| Number of type of systems integrated | $y = 0.0726x + 0.2107$ | $y = 0.0856x + 0.2455$ |
| Number of organizations involved | $y = 0.0734x + 0.219$ | $y = 0.0751x + 0.267$ |
| Number of implementation teams involved | $y = 0.0595x + 0.3221$ | $y = 0.0843x + 0.0752$ |

Where the projects differed by just one of the four attributes discussed earlier, the analogy was established as simply the ratio of the two calculated risk ratings for that attribute, i.e. that of the project to be estimated over the analogue. Where more than one attribute differed, the ratio of the mean values of the calculated risk ratings for each of the two projects was taken as the analogy. The results of the estimation obtained both by the method discussed here and

by using the ANGEL tool were compared to the actual duration of each of the four project cases and to the original estimate at the time of starting that project. A summarised view of the results of this experiment is given in table 4 and a graphical representation is provided in figure 8 of chapter 6.

A further data set containing just twenty one projects gathered through various sources in Greece, Spain and Serbia was compiled. These were not integrated with the ISBSG data as they have not undergone a similar exposure to researchers and scrutiny as to the integrity of the data. The data for this third data set was collected by the author between 2008 and 2010 and the data entries were not as rich as those of ISBSG, limiting the fields to those factors that were considered under this research. The purpose of this smaller more selective set was to test how a more limited (or controlled) dataset can affect the accuracy of the identification of a suitable analogue. A separate data model for use with the ANGEL tool using this small group of local government projects data was compiled. This was then used to estimate the two local government projects that were earlier estimated based on the ISBSG data and the results are presented in table 5 and figure 9, shown in chapter 6 of this work.

# CHAPTER 6

# Estimating Using Risk Based Analogues

This chapter presents, discusses and evaluates the case of the four projects that were used for validating the proposed approach to project duration estimation using the estimation by analogy method. The estimates obtained by the risk based method for calculating the analogy amongst the project to be estimated and the analogue identified, are benchmarked against estimates obtained by using the ANGEL tool as an estimation tool.

## 6.1    Estimating using the ISBSG data

The four projects to be estimated were obtained from environments independent to each other and incorporating systems as distinct in nature as possible within the context of e-business development and VACs.

The London Borough of Havering is one of the 32 Boroughs of the Greater London Authority and the City of London with a population of about 170,000 people. The Borough has over the past ten years invested heavily in online systems aiming to improve the quality of local government services in London. It provides a range of services online to casual and registered users linking to further supporting third party organizations and companies in an effort to improve their services and provide a full e-service experience to users. The projects discussed here involve the integration of existing database systems, third party payment systems, Enterprise Resource Planning system and a Customer Relationship Management system.

"Protoporia" is a Greek bookstore chain with presence in the four biggest cities in Greece and the most successful online book retailer in the country. Their portal serves customers for placing purchase orders and provides a follow-up facility of their orders. It is also the

gateway to a small but efficient VAC. Data from online sales along with data captured from online checkout registers at the stores is sent to four of the major publishing houses in Greece. Their systems process the data and organize shipments of books to the bookstores shops around the country and its warehousing centre through a logistics company which facilitates the transport of orders. The publishers' systems update those of the logistics company and those of the bookstore and also update online customers as to the processing of their order and the expected date of delivery.

ELGEKA S.A. is the largest commercial company in the Greek food sector. Offering its collaborators, customers and suppliers, an integrated system of commercial services comprising Sales, Marketing, Trade Marketing, and Logistics, ELGEKA constitutes an integral link in the sector's supply chain. The company has developed a B-2-B portal for supporting its customers (thousands of retail outlets in Greece) in placing orders and checking the status of their orders online. The portal links the company's offices to the warehousing facilities and a number of logistics support companies (including their own subsidiary) in planning order deliveries. The VAC, based on this portal, also serves in collating data for automated order placing with their suppliers both abroad (the vast majority) and in Greece. Table 4 below presents a summary of the outcome of the estimation experiment involving the above four projects. Furthermore, figure 8 provides a diagrammatic comparison of the results obtained for each project using the risk based approach to estimation developed by the author, the estimate provided by the ANGEL tool and the original estimate at the time of the start of the project.

Table 4. Summary of Estimation Results using the ISBSG data for analogues

| Company | Project Title | Data Category | Duration (Days) | Variation from Actual (%) |
|---|---|---|---|---|
| LB Havering | Planning Online & Content Management | Actual | 205 | |
| | | ANGEL Estimate | 198 | -3.41 |
| | | Risk Based Estimate | 214 | 4.39 |
| | | Original Estimate | 195 | -5.13 |
| | Online Bookings & Payments | Actual Data | 426 | |
| | | ANGEL Estimate | 353 | -17.14 |
| | | Risk Based Estimate | 458 | 7.51 |
| | | Original Estimate | 390 | -8.45 |
| Protoporia | Order Placing and Coordination System | Actual Data | 426 | |
| | | ANGEL Estimate | 383 | -11.10 |
| | | Risk Based Estimate | 437 | 2.58 |
| | | Original Estimate | 385 | -9.62 |
| ELGEKA | Customer B-2-B & Logistics Portal | Actual Data | 262 | |
| | | ANGEL Estimate | 290 | 10.68 |
| | | Risk Based Estimate | 246 | -6.10 |
| | | Original Estimate | 240 | -8.40 |
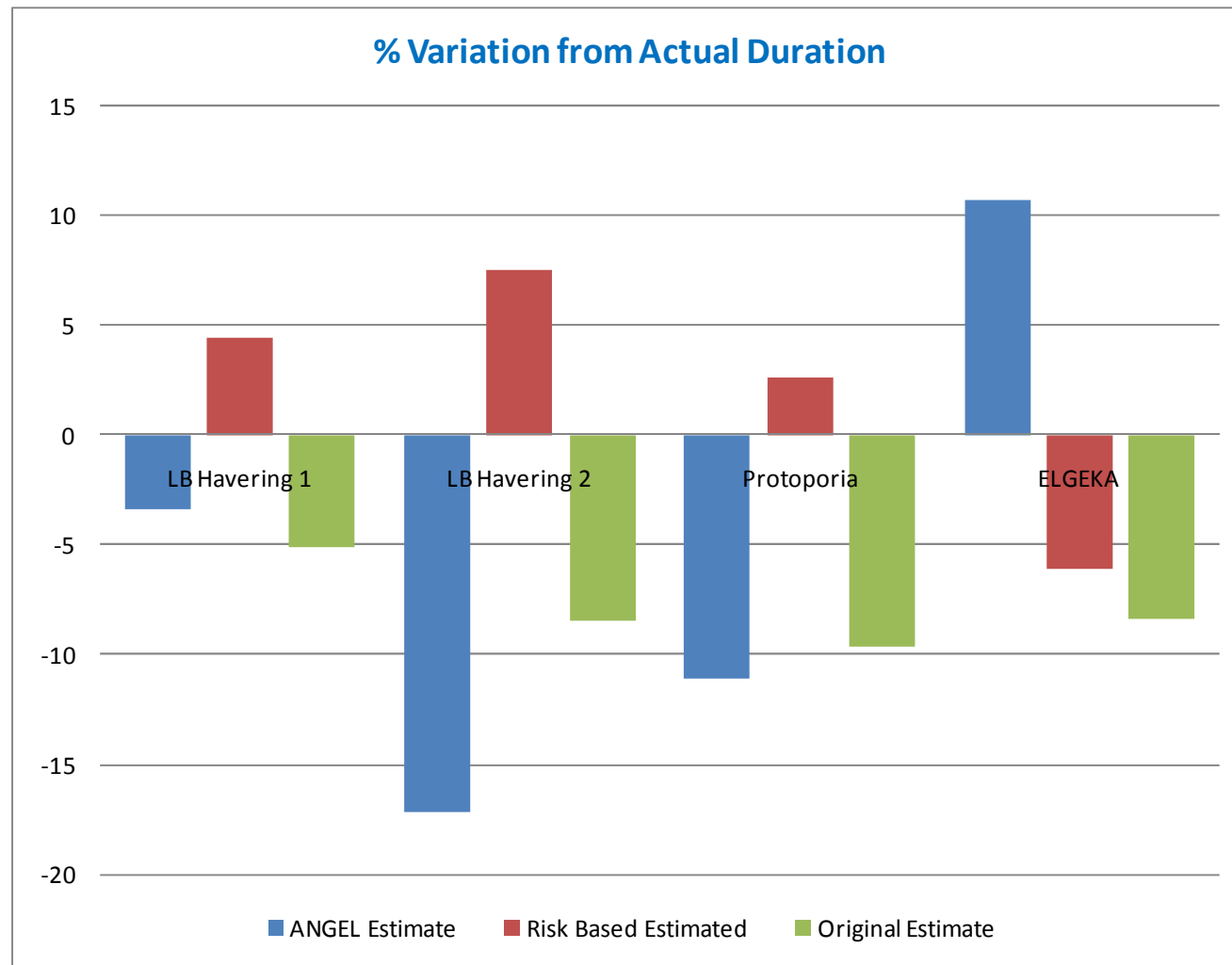
Figure 8. Variation of estimates for the four project cases, using ISBSG data

## 6.2    Evaluation of the results

In contrast to other approaches to estimating projects, the method discussed uses analogy in attempting to estimate the duration of the project instead of the effort required to complete the project or its individual components. The reason for doing so is simple and is explained in the context of the study of risk behaviour for VAC development project that precedes this last phase of the work. The duration of VAC projects is not only affected by the development effort put into completing each individual task and phase of the projects but equally by the interactions between the different partners and the different teams and subtasks operating across the projects. For such reasons it is not possible to apply the oversimplified suggestion that to estimate a web based project simply divide the project into clearly defined subtasks and estimate each one individually; adding up the individual components would yield the sum of efforts and this would in turn be converted into time (Hill 2010, Mendes et al 2005, Mendes 2008). Such approach might work when estimating a simpler web based project (like developing an online database) as the researchers above have demonstrated, but it is not possible to work with the conditions affecting VAC projects.

A further characteristic of VAC projects is that due to the interactions mentioned above planning is often revised and hence the balance between tasks run in sequence or in parallel changes frequently, rendering the above simple summation of effort approach rather impractical. This is a further reason that this work does not concentrate on estimating effort but simply on the overall duration of the project aiming to address the most critical success factor of the project, the launch day of the VAC.

The results shown in table 4 and figure 8 demonstrate that there is good reason to want to further develop and experiment with the use of this approach to estimating projects such as VAC development ones. In all cases the estimated project duration was better than the original estimates obtained. The source containing the data on projects comprising the

database from which analogues were extracted were not directly related to the environment of the projects used as test cases for the estimation experiment. This demonstrates the potential applicability of the process to a wide range of projects that share the complexity issues surrounding VAC projects. Also the origin of the projects and in particular those from Greece demonstrates that the approach is not sensitive to any cultural or local elements that might affect the running and delivery of a project to be estimated, since the ISBSG database does not contain any data on projects from Greece or any other countries in the geographical vicinity.

In scrutinizing the results the reader can see that in most cases the proposed approach to estimation by analogy has worked quite well and has achieved variations from the actual project durations that would be considered within an acceptable tolerance level (-6.10 % to +7.51 %). In all cases the estimates derived using the author's risk based approach to estimating VAC projects is better than the original estimates obtained for those projects. Regarding the actual figures the reader might be surprised that the above range is considered acceptable. However, according to Hill (2010) in an analysis of those projects for which both estimates and actual duration data had been submitted about half had been completed outside their duration estimates, with the majority of those completing late. The average of overrun was 100% for small projects. About 10% of those completing outside the estimate were early completions, but in most cases this was achieved at a considerable expense of extra effort and resources. The above appears to be in agreement with the work originally performed by Moløkken and Jørgensen (2003) and subsequently continued by Gruschke and Jorgensen (2008). They argue that surveys of effort estimates for software development projects report that such estimates are frequently inaccurate, with the average effort overrun of software projects shown at a level of 30–40%. What is more worrying is that the accuracy of estimates has not improved much over the last ten to twenty years, something that is broadly confirmed

by a survey of effort estimation by El Eman and Koru (2008). Gruske and Jorgensen (2008) consider important reasons for the lack of accurate effort estimates that information system development is a complex process that is often poorly understood by its practitioners. They further emphasise that most of the activities within the process are still primarily human rather than mechanical, and thus prone to all the subjective factors that may lead to considerable deviation in the performance of estimators. There are various types of uncertainties in estimation and statistical variance is always going to be there with the only potential improvement being the ability of the project manager to control its size. The Project Management Institute (PMI) recommends maintaining a 10-percent management contingency buffer in the absence of any other information. Jorgensen (2004) suggested that it is important in order to effectively communicate one's estimates, instead of giving just a number, to give a range suggesting that the actual results will be less than or equal to the estimate at a given percentage of the time (e.g. 90% of the time. This view is also shared in subsequent work by Koch and Mitlöhner (2009) and Kulk et al (2009). In addition, Laird (2006) states that "Estimates are typically the 50-percent view, meaning the probability for being under or over budget is the same. (Unfortunately, Parkinson's Law, which states that work expands to meet the time available, holds for software projects. So, this 50-percent view says we will be on budget 50 percent of the time and over budget the other 50 percent.)"

Within the above cloud of uncertainty re estimation results and considering that estimation accuracy increases as the work progresses and with further details of the project being identified the estimates obtained using the risk based method have to be seen as a good first set of results (Laird 2006). This is further supported by the fact that these were only early phase estimates with only an outline of the project details known.

Looking at figure 8 which shows the variation from the actual duration, the reader can see a more impressive outcome. In the three of the four cases not only is the variation from the

actual duration better than that of the original estimate, but the estimate obtained using the author's method is much better from a qualitative point of view. The three estimates, namely those for the two LB of Havering projects and that for the "Protoporia" project, are over the actual duration, while the original estimates for those projects obtained at the time of implementation are under the actual duration. This fact makes the estimates obtained using the risk based method much more useful within the context of estimating e-business development. One of the key success factors has been identified as that of a timely launch of the e-business and providing an over estimation minimises the risk of running short of time and defaulting on the original plans for launch. On the contrary an under estimation of the project's duration at the early stages can lead to considerable revisions of plans and the hiring and deployment of expensive additional resources at later stages of the project. This would be the result of the project manager trying to meet a very optimistic deadline, for launching the e-business, which has been based on the earlier underestimation of the duration of the project. A consequence of this could be an overburdened project budget that might affect the future of the business.

To revert to an earlier argument on the issue of cultural differences and different work environments and given the plethora of projects submitted to the ISBSG database, one cannot be quite certain as to how the actual working day was measured for each one and whether slight differences in measurement and recording methods can affect the result of the estimates obtained. Therefore, although one of the most consistent and widely trusted databases has been used for this work, a definitive answer as to the quality of the outcome of the estimates obtained cannot be provided. This is by no means a weakness of the method, but possibly the issue of lack of generally acceptable standards, as there is no particular level of tolerance that can be applied cross industry.

Overall, the variations from the actual durations for the four project cases have been significantly less than those yielded by the ANGEL tool, although both methods used the same analogues. This is not a criticism of the ANGEL tool which has been developed to estimate on the basis of function points and not the more general project dimensions that were used as attributes in this work. It demonstrates though that the proposed approach to estimation by analogy can deliver results of possibly acceptable accuracy. There was one case though where both methods produced considerable differences from the actual (7.51% for the proposed approach) and this could raise some concerns for the proposed approach. The author will not consider this as a serious problem though as the nature of the data has to be considered carefully before any conclusions are drawn. A close inspection of the data set used in the experiment demonstrates a significant difference in one of the attributes used that of the duration of the analogue, which appears considerably different (lower) from other projects of similar size. The quality of the historical data upon which analogues are drawn can influence the outcome of estimations by analogy and it is the author' view that this is the case here (Gruschke and Jorgensen, 2008, Chapman and Ward 2003, Hall 1998). Many researchers have claimed that as project managers build their own repositories of project data they will be able to understand the strength and weaknesses of their data and hence to trust the quality and accuracies of their estimations (Fenton and Pfleeger 1997, Mendes 2008, Li et al 2007, Pfleeger 2008, Stamelos  and Angelis 2001). This, the author believes, is the case with the proposed estimation approach. To strengthen this belief, the second experiment with the smaller set of data collected by the author from a very specific application domain that matches in principle the local government projects estimated here, yielded improved results when used with the risk based estimation by analogy method. Further building of domain specific repositories of project data could possibly enhance the accuracy of the proposed estimation approach. Although there is an encouraging indication in this second experiment,

the results are limited as to their extent of application and further tests would be required with possibly larger data sets. Furthermore, to test this hypothesis fully, datasets from industry (business) based projects should be built to explore this other application domain area. Finally, one should not discount the fact that the ISBSG historical database has been growing for more than a decade now and some of the project data might be dated. At the same time this particular database has had its data audited by a wide range of industry and research base users and one should be inclined to feel reasonably confident in its use. The building and validating of proprietary large datasets to further test the validity of this research work is time consuming, resource demanding and beyond the scope of this programme of research.

## 6.3 A Case of a Controlled and Restricted Estimation Experiment

Following suggestions from the literature with experiences from other authors that estimation based on a combination of expert opinion and historical project data often yields better results when one is using a pool of data that closely matches the type of project to be estimated (Jørgensen and Grimstad 2008, and Smith et al 2001), the author has attempted a smaller controlled experiment to test this hypothesis in the case of risk based estimation by analogy. Over the space of two years (2008-2010) data for twenty one relatively small projects from local authorities were collected from three different countries, quite distinct to each other both geographically and politically. The projects originated from Spain, Greece and Serbia and are of duration between six and fourteen months each. They are for systems that were developed and integrated in local government authorities at various levels of administration, serving primarily internal admin functions and integrating systems across different government functions and in some cases external agencies and suppliers. A full listing of project data is shown in Appendix C.

This new collection of data was used as a new data model (dataset) with the ANGEL tool instead of the ISBSG data and was used to estimate the two LB of Havering projects considered in section 6.2 above. The results obtained can be seen in tabular and graphical format in comparison with those for the ISBSG data in table 5 and figure 9 respectively. They demonstrate a better estimate, but only marginally and still do not match the actual duration of the project. They do maintain the trend shown by the previous set of results, where the estimates obtained are over the actual duration, and combined with the fact that the accuracy of these two estimates is better than the ones obtained previously raise the issue of exploring further the concept of own data models to be used with the ANGEL tool.

Table 5. Comparison of estimates of the two council projects, Independent data

| | % Variation from Actual Duration | | |
| --- | --- | --- | --- |
| PROJECT | RISK BASED ISBSG | RISK BASED INDEPENDENT | ORIGINAL ESTIMATE |
| LB Havering 1 | 4.39 | 2.83 | -5.13 |
| LB Havering 2 | 7.51 | 4.79 | -8.45 |
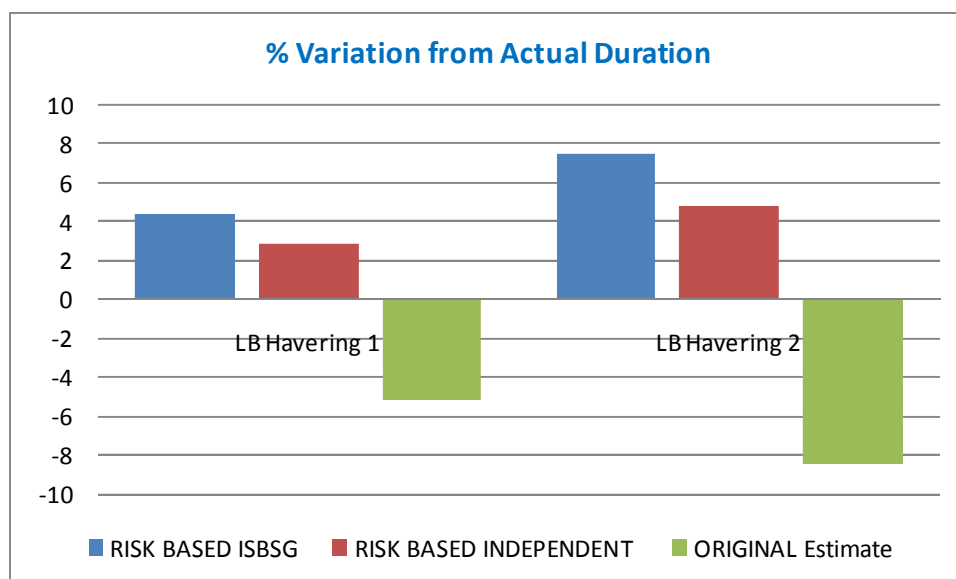


Figure 9. Variation of estimates for the council's projects using independent data

# CHAPTER 7

# Conclusions and Recommendations for Further Research

This work aimed at seeking a means of improving the accuracy of estimates for the duration of e-business development projects, while at the early stages of a project and before the detailed requirements of the project are made known. The author has developed a variation to estimation by analogy where analogies are established by means of the variation of risk with project size and complexity.

## 7.1    Conclusions

The results obtained demonstrate a good level of accuracy that is within an acceptable range when considering the accuracy of established estimation methods and techniques. All estimates obtained using the risk based method are better than the original estimates, obtained at the time of the start of each of the projects that were used as cases for the validation of the method.

More encouraging is the fact that in three of the four cases the author's method yielded overestimations of the duration of the projects instead of the underestimations of the original estimates. This is a strong indication that this method is suitable for estimating the duration of e-business projects where the accuracy of the launch date of the business is critical and an underestimation of that might prove catastrophic. The strength of this aspect of the results was reconfirmed when a second experiment with a small dataset used for the two councils projects yielded similar results as to the overestimation, only with better accuracy as to the actual duration of the projects.

To develop the method, the author has applied a risk evaluation approach that has allowed for the identification of a risk list relevant to e-business projects which was validated by field experts. The variation of the validated list of risks with the variation of project attributes that

determine the size and complexity of an e-business project was established and confirmed. The resulting equations describing the variation of risks were utilised in estimating the duration of four value-added community development projects. The risk variation equations provided the means by which the actual duration of suitable project analogues where adjusted to reflect an estimate for the duration of the projects used for testing the method.

One of the most renowned, accurate and verified projects database (ISBSG) was used as the source from which analogues were identified.

To allow for full confidence in the method and to test its potential use in other project domains further work is required which is beyond the scope of this work. Suggestions for such work follow this conclusions section.

The author acknowledges that there are limitations to this work, but believes that these do not undermine the quality of the results obtained. Such limitations are the size of the repositories used and the specific nature of the projects the method was tested on. A reader could see a limitation in the nature of the data collected as this comprises risk perception rather than actual facts. The consistency of the results obtained though instils confidence in both the quality of the data and the accuracy of the method presented and discussed in this work.

## 7.2    Recommendations for Further Research

The work for this research was confined to specific types of e-business development projects, testing the hypothesis that risk and its variation with project characteristics can be a means of establishing suitable analogues for estimating the duration of projects at the early stages of implementation when detailed project requirements might not be clear.

The work performed showed that this hypothesis is valid, but to instil further confidence and to possibly broaden the use of such an approach more work is required. Further research should seek out whether the accuracy of the results obtained is sensitive to the type of project

data collected. To this effect access to other project repositories could be obtained or the building of further domain specific project databases could be attempted. Both could be time consuming as there are only few large project databases available on a worldwide basis that are available to researchers and that contain quality data as the one used in this work (Mendes and Mosley, 2008). Furthermore, developing a domain specific database of comparable size would be very hard to achieve and the quality of the data contained would have to be validated thoroughly before the results obtained through its use can be considered with any confidence.

Another element of further work is that of broadening the application horizon of the results of this work. Further research with project data from other domains could be conducted and the method could be applied to a wider range of project types. This would allow for further validation of its applicability and would demonstrate the full value of the work and results discussed here.

The author believes that the work and the estimation method presented here could apply to various types of projects in the wider computing area. These projects will be of such types that justify the approach discussed here. They'll be projects that involve many different functional teams, different stakeholders from diverse backgrounds, with the detailed requirements not always known during the earlier stages and engaging new and emerging technologies that might have not had that much exposure to testing in real world systems. Such projects by definition are characterised by high degree of uncertainty (Boehm, 2008 and Persson and Mathiassen, 2010). Given the hypothesis behind this work the fundamental principles governing the outcome of this research could be extended to such types of projects. Indicatively the author would suggest web service development projects as a potential category of projects. These are characterised by high uncertainty, embrace continuously evolving technologies in order to stay at the high end of service delivery and accommodate or

even define new business innovations (Featherman and Wells 2004, Shi 2007, and Birkinshaw et al, 2011).

A further area of potential application of this estimation method is that of computer games development projects. Similarly to web services ones, new technologies to offer cutting edge graphics and state of the art programming capabilities are utilized. Depending on the project size a large number of partners and stakeholders might engage and the development work could be equally distributed to a number of diverse teams with varying experience and geographically dispersed. All these project characteristics are ingredients for high levels of uncertainty and that on its own would make such projects reasonable candidates to test the expandability of the application of the estimation method discussed in this thesis.

An emerging area of applicability is that of cloud computing. Currently most cloud applications are proprietary and the exposure to uncertainties due to uncontrolled or loosely controlled project hierarchies is rather limited. They do embrace state of the art technologies, they can involve diverse project teams and as dissemination of technologies and platforms increases more independent applications are expected. Then projects will increase in uncertainty as the competitive nature of different developers and service providers will yield a state of ever increasing conflicting requirements. Both computer games development and cloud computing projects operate in a state of ever shorter demand times imposed by external influences and realistic duration estimates are the only sensible requirement there.

As a last recommendation for further work the author would raise the need to investigate the effect on the accuracy of the method discussed here that the actual contents of the database from which analogues are identified could have. An issue that might arise is that more than one project might share the same project attributes, but have different project durations. This might reflect cultural differences or different work practices that differentiate in the way that a work day is defined and accounted for. This issue did not surface in the experimental work

here, but it could be something to expect in the case of large datasets, in particularly when project data is coming from diverse backgrounds.

# References

Addison T., and Vallabh S. (2002) 'Controlling Software Project Risks – an Empirical Study of Methods used by Experienced Project Managers', *the 2002 Annual Conference of the South African Institute of Computer Scientists & Information Technologists (SAICSIT 2002)*, pp.128-140.

Aagesen, G. and Krogstie, J. (2010) 'Investigating Requirements for Transformational Government Information Infrastructures: The Case of the Approval Process for Building Applications', *the 10th European Conference on eGovernment*, National Centre for Taxation Studies and University of Limerick, Ireland, 17-18 June 2010, pp. 1-9.

Aladwani A. M. (2002) 'IT project uncertainty, planning and success: An empirical investigation from Kuwait', *Information Technology & People*, Vol. 15 No. 3, 2002, pp. 210-226.

Andal-Ancion, A., Cartwright, P.A. and Yip, G.S. (2003) 'The Digital Transformation of Traditional Businesses', *Sloan Management Review*, Vol. 44, No. 4, pp. 34-41.

Arora, A., Hall D., Ariel Pinto, C., Ramsey, D., and Telang, R. (2004), 'Measuring the Risk-Based Value of IT Security Solutions', *IT Professional*, Vol. 6, No. 6, pp. 35-42.

Bandyopadhyay, K., Mykytyn, P. P. and Mykytyn, K. (1999) 'A framework for integrated risk management in information technology', *Management Decision*, Vol. 37, No.5, pp.437-444.

Benaroch, M. and Goldstein J. (2009) 'An Integrative Economic Optimization Approach to Systems Development Risk Management', *IEEE Transactions on Software Engineering,* Vol. 35, No. 5, pp. 638-653.

Birkinshaw, J., Bouquet, C. and Barsoux, J.-L. (2011) 'The 5 Myths of Innovation', *Sloan Management Review,* Vol.52, No.2, pp. 43-50.

Boehm, B. (2008) 'Making a difference in the software century', *IEEE Computer,* Vol. 41, No. 3, pp. 32-38.

Boehm, B. (1991) 'Software Risk Management: Principles and Practices', *IEEE Software*, Vol. 8, No. 1, pp. 32-41.

Boehm, B. and Bhuta, J. (2008) 'Balancing Opportunities and Risks in Component-Based Software Development', *IEEE Software*, Vol. 25, No. 6, pp. 56-63.

Boehm, B. and Turner, R. (2003) 'Using Risk to Balance Agile and Plan-Driven Methods', *IEEE Computer*, Vol. 20, No. 3, pp. 57-66.

Boehm, B. and Basili, V. R. (2001) 'Software Defect Reduction Top 10 List', *IEEE Computer*, Vol. 34, No. 1, pp. 135 – 137.

Boehm, B., Abts, C., Winsor Brown, A., Chulani, S., Clark B.K., Horowitz, E., Madachy, R., Reifer, D.J. and Steece, B. (2000) *Software Cost Estimation with Cocomo II,* Prentice Hall

Boehm, B. and Fairley, R.E. (2000) 'Software Estimation Perspectives', *IEEE Software,* Vol. 17, No. 6, pp. 22 – 26.

Boehm, B. and DeMarco, T. (1997) 'Software Risk Management', *IEEE Software*, Vol. 14, No. 3, pp. 17-19.

Bourke, P., Oligny, S., Abran, A. and Fournier, B. (2007) 'Developing Project Duration Models in Software Engineering', *Journal of Computer Science and Technology*, Vol. 22, issue 3, pp. 348-357.

Brynjolfsson, E. and McAfee A.P. (2007) 'Beyond Enterprise 2.0', *Sloan Management Review*, Vol. 48, No. 3, pp. 50-55.

Buglione, L., and Ebert, C. (2011) 'Estimation Tools and Techniques', *IEEE Software*, Vol. 28, No. 3, pp. 91-94.

Bygstad B. (2004) 'Controlling Iterative Software Development Projects: The Challenge of

Stakeholder and Technical Integration', *the 37th Hawaii International Conference on System Sciences (HICSS-37 2004)*, Big Island, HI, USA, 5-8 January 2004

Castelnovo, W. (2011) 'Risk Management in a Cooperation Context', *the 11th European Conference on eGovernment*, Faculty of Administration, University of Ljubljana, Ljubljana, Slovenia, 16-17 June 2011, pp. 132-139

Cataldo, M., Mockus, A., Roberts, J.A. and Herbsleb, J.D. (2009) 'Software Dependencies, Work Dependencies, and Their Impact on Failures', *IEEE Transactions on Software Engineering*, Vol. 35, No. 6, pp. 864-878.

Chapman, C. and Ward, S. (2003) *Project Risk Management*, 2nd edition, Wiley

Cho S.E., Park K. (2002), 'Empirical taxonomy of services and service products in electronic commerce', *Electronic Commerce Research and Applications*, Vol. 1, No. 3-4, pp. 339-350.

Christiaanse, E. and Lynne Markus M. (2003) 'Participation in Collaboration Electronic Marketplaces': *the 36th Hawaii International Conference on System Sciences (HICSS-36 2003)*, Big Island, HI, USA, January 6-9, 2003.

Cloyd, M.H. (2001) 'Designing User-Centered Web Applications in Web Time', *IEEE Software,* Vol. 18, No. 1, pp. 62-69.

Conrow, E.H. and Shishido, P.S. (1997) 'Implementing Risk Management on Software Intensive Projects', *IEEE Software,* Vol. 14, No. 2, pp. 83-89.

Costa, H.R., Marcio de O. Barros, M., and Travassos, G.H. (2007) 'Evaluating software project portfolio risks', *The Journal of Systems and Software,* 80 (2007) 16–31

Cusumano, M. and Hopkins M.S. (2011) 'How to Innovate When Platforms Won't Stop Moving' *Sloan Management Review,* Vol. 52, No. 4, pp. 55-60.

Damian, D.,  and Chisan J. (2006) 'An Empirical Study of the Complex Relationships between Requirements Engineering Processes and Other Processes that Lead to Payoffs in

Productivity, Quality, and Risk Management', *IEEE Transactions on Software Engineering,* Vol. 32, No. 7, pp. 433-453.

Davies P. (2000) 'Survival of the fittest', *The Computer Bulletin*, The British Computer Society, September 2000

De Oliveira Barros, M., Werner, C.M.L. and Travassos, G. H. (2004) 'Supporting risks in software project management', *Journal of Systems and Software*, Vol. 70, No. 1-2, pp. 21-35.

De Saulis, M. (2011) 'Social Media and Local Government in England: Who is Doing What?', *the 11th European Conference on eGovernment*, Faculty of Administration, University of Ljubljana, Ljubljana, Slovenia, 16-17 June 2011, pp. 187-192.

Donaldson, S.E. and Siegel, S.G. (2007) 'Enriching Your Project Planning: Tying Risk Assessment to Resource Estimation', *IT Professional,* Vol. 9, No. 5, pp. 20-27.

Dyer, J.H. and Hatch, N.W. (2004) 'Using Supplier Networks to Learn Faster', *Sloan Management Review*, Vol. 45, No. 3, pp. 57–63

Dzbor, M., Cimiano, P., Uren, V., and Chapman, S. (2005) 'Study on Integrating Semantic Applications', *the ISWC 2005 Workshop on End User Semantic Web Interaction.* Galway, Ireland, November 7, 2005

El Emam, K., and Koru, A.G. (2008), 'A Replicated Survey of IT Software Project Failures', *IEEE Software*, Vol. 25, No. 5, pp. 84-90

Eveleens, J. L., and Verhoef, C. (2010) 'The Rise and Fall of the Chaos Report Figures', *IEEE Software*, Vol. 27, No. 1, pp. 30-36

Fairley R.E. (2005) Software Engineering Glossary: Software Risk Management, *IEEE Software,* Vol. 22, No.3, p. 101.

Featherman, M.S., and Wells, J.D. (2004) 'The Intangibility of E-Services: Effects on Artificiality, Perceived Risk, and Adoption', *the 37th Hawaii International Conference on System Sciences (HICSS-37 2004)*, Big Island, HI, USA, 5-8 January 2004

Fehlmann, T. (2002) 'Business-Oriented Testing in E-Commerce' in Meyerhoff, D., Laibarra, B., van Der Pouw Kraan, R. and Wallet, A., *Software Quality and Software Testing in Internet Times*, Berlin; Springer-Verlag, pp. 117-137.

Fenton, E.N., and Pfleeger, S.L. (1997) *Software Metrics – A Rigorous & Practical Approach*, 2nd edition, PWS Publishing Company

Ferro, E. and Molinari, F. (2010) 'Framing Web 2.0 in the Process of Public Sector Innovation: Going Down the Participation Ladder', *European Journal of ePractice*, Vol. 9, [online]. Available at: http://www.epractice.eu/en/journal/volume/9 (accessed 17 June 2011)

Galorath, D.D., and Evans, M.W. (2006) *Software Sizing, Estimation, and Risk Management*, Boca Raton, FL: Auerbach Publications

Glass, R.L., Rost, J and Matook, M.S. (2008) 'Lying on Software Projects', *IEEE Software*, Vol. 25, No. 6, pp. 90-95

Gleghorn, R. (2005) 'Enterprise Application Integration: A Manager's Perspective', *IT Professional*, Vol.7, No. 6, pp. 17-23.

Glinz, M. (2008) 'A Risk-Based, Value-Oriented Approach to Quality Requirements', *IEEE Software*, Vol. 25, No. 2, pp. 35-41.

Gruschke, T.M., and Jorgensen, M., (2008) The Role of Outcome Feedback in Improving the Uncertainty Assessment of Software Development Effort Estimates, *ACM Transactions on Software Engineering and Methodology*, Vol. 17, No. 4, Article 20.

Hastings, T.E. and Sajeev A.S.M. (2001) 'A Vector-Based Approach to Software Size Measurement and Effort Estimation', *IEEE Transactions on Software Engineering*, Vol. 27, No. 4, pp. 337 – 350.

Hall, E. M. (1998) *Managing Risk Methods for Software Systems Development*, Reading, MA, USA: Addison Wesley.

Hill, P. (ed.) (2010) *Practical Software Project Estimation: A Toolkit for Estimating Software Development Effort & Duration,* New York, USA: McGraw-Hill Osborne

Hill, P. (ed.) (2005) *Practical Project Estimation*, 2nd edition, Melbourne, Australia: International Software Benchmarking Standards Group (ISBSG)

Houston, D. X., Mackulak, G. T. and Collofello, J. S. (2001) 'Stochastic simulation of risk factor potential effects for software development risk management', *Journal of Systems and Software*, Vol. 59, No. 3, pp. 247-257.

Howard, B. (2011) 'London 2012: Ticket payment gaffe by Barclaycard' [Online]. Available at http://www.bbc.co.uk/news/business-13576179 (Accessed 30 May 2011)

Jørgensen, M. (2004) 'A review of studies on expert estimation of software development effort', *The Journal of Systems and Software* 70 (2004) 37–60

Jorgensen, M., and Boehm, B. (2009) 'Software Development Effort Estimation: Formal Models or Expert Judgment?', *IEEE Software*, Vol. 26, No. 2, pp. 14-19.

Jørgensen, M. and Grimstad, S. (2008) 'Avoiding Irrelevant and Misleading Information When Estimating Development Effort', *IEEE Software*, Vol. 25, No. 3, pp. 78-83.

Jorgensen, M., and Shepperd, M. (2007) 'A Systematic Review of Software Development Cost Estimation Studies', *IEEE Transactions on Software Engineering*, Vol. 33, No. 1, pp. 33-53.

Jørgensen, M.and Moløkken-Østvold, K. (2004) 'Reasons for Software Effort Estimation Error: Impact of Respondent Role, Information Collection Approach, and Data Analysis Method', *IEEE Transactions on Software Engineering*, Vol. 30, No. 12, pp. 993-1007.

Jørgensen, M., Indahl, U., and Sjøberg, D. (2003) 'Software effort estimation by analogy and ''regression toward the mean''', *The Journal of Systems and Software,* 68 (2003) 253–262

Kan, S.H. (2002) Metrics and Models in Software Quality Engineering. 2nd edn. Boston: Addison-Wesley

Kansala, K. (1997) 'Integrating Risk Assessment with Cost Estimation', *IEEE Software,* Vol. 14, No. 3, pp. 61-67

Keil, M., Li, L., Mathiassen, L. and Zheng, G. (2006) 'The Influence of Checklists and Roles on Software Practitioner Risk Perception and Decision-Making': *the 39th Hawaii International Conference on Systems Science (HICSS-39 2006)*, Kauai, HI, USA, 4-7 January 2006.

Keil, M., Cule, P. E., Lyytinen, K. and Schmidt R. C. (1998) 'A Framework for Identifying Software Project Risk', *Communications of the ACM*, Vol. 41, No. 11, pp. 76-83.

Keung, J. W., Kitchenham, B. A. and Jeffery, D. R. (2008) 'Analogy-X: Providing Statistical Inference to Analogy-Based Software Cost Estimation', *IEEE Transactions on Software Engineering*, Vol. 34, No. 4, pp. 471-484.

Koch, S. and Mitlöhner, J. (2009) 'Software project effort estimation with voting rules', *Decision Support Systems* 46 (2009) 895–901

Kopczak, L.R. and Johnson M.E. (2003) 'The Supply-Chain Management Effect', *Sloan Management Review*, Vol. 44, No. 3, pp. 27–34

Kulk, G.P., Peters, R.J. and Verhoef, C. (2009) 'Quantifying IT estimation risks', *Science of Computer Programming,* 74 (2009) 900-933

Laird, L.M. (2006) 'The Limitations of Estimation', *IT Professional*, Vol. 8, issue 6, pp. 40-45.

Lam W. and Shankararaman V. (2004) 'An Enterprise Integration Methodology', *IT Professional*, Vol. 6, No. 2, pp. 40-48.

Laseter, T. (2011) 'Focus and Scale on the Internet', *Strategy and Business,* Issue 63, Summer 2011.

Lee, D.H-D. (2005) 'Contextual IT Business Value and Barriers: an E-Government and E-Business Perspective', *the 38th Hawaii International Conference on System Sciences (HICSS-38 2005),* Big Island, HI, USA, 3-6 January 2005

Letens, G., Van Nuffel, L., Heene, A. and Leysen, J. (2008) 'Towards A Balanced Approach in Risk Identification', *Engineering Management Journal*, Vol. 20, No. 3, pp. 3-9.

Li, J., Rube, G., Al-Emran, A. and Richter M.M. (2007) 'A flexible method for software effort estimation by analogy', *Empirical Software Engineering*, Vol.12, pp. 65-106.

Lippert, M. and Zullinghoven, H. (2002) 'Using Extreme Programming to Manage High-Risk Projects Successfully', in Meyerhoff, D., Laibarra, B., van Der Pouw Kraan, R. and Wallet, A., *Software Quality and Software Testing in Internet Times*, Berlin; Springer-Verlag, pp. 85-100.

McAfee, A.P. (2006) 'Enterprise 2.0, the Dawn of Emergent Collaboration' *Sloan Management Review*, Vol. 47, No. 3, pp. 21-28.

McAfee, A.P. (2005) 'Will Web Services Really Transform Collaboration?', *Sloan Management Review*, Vol. 46, No. 2, pp. 78-84.

McWilliam G., (2000) Building Stronger Brands Through Online Communities, *Sloan Management Review*, Vol. 41, No. 3, pp. 43 – 54.

Means, G. and Schneider, D. (2000) *Meta-Capitalism*, New York, John Wiley & Sons.

Mendes, E. (2008) *Cost Estimation Techniques for Web Projects*, Hershey, PA, USA: IGI Publishing

Mendes, E. and Mosley, N. (2008) 'Bayesian Network Models for Web Effort Prediction: A Comparative Study', *IEEE Transactions on Software Engineering,* Vol. 34, No. 6, pp. 723 – 737.

Mendes, E., Lokan, C., Harrison, R. and Triggs, C. (2005) 'A Replicated Comparison of Cross-company and Within-company Effort Estimation Models using the ISBSG Database',

*the 11th IEEE International Symposium on Software Metrics (METRICS 2005)*, Como, Italy,19-22 September 2005.

Mendes, E., Mosley, N. and Counsell, S. (2003) 'Early Web Size Measures and Effort Prediction for Web Costimation' *the Ninth International Software Metrics Symposium (METRICS'03)*, Sydney, Australia, 3-5 September 2003

Mendes, E., Mosley, N. And Watson, I. (2002) 'A Comparison of Case-Based Reasoning Approaches to Web Hypermedia Project Cost Estimation': *The Eleventh International World Wide Web Conference (WWW2002)*, Honolulu, Hawaii, USA, May 7-11 2002.

Mendes, E., Mosley N. and Counsell, S. (2001a) 'Using an Engineering Approach to Understanding and Predicting Web Authoring and Design': *the 34th Annual Hawaii International Conference on System Sciences (HICSS-34),* Maui, Hawaii, USA, January 3-6, 2001

Mendes, E., Mosley, N. and Counsell, S. (2001b) 'Web Metrics – Estimating Design and Authoring Effort', *IEEE Multimedia*, Vol. 8, No. 1, pp 50-57.

Menzies, T., Chen, Z., Hihn, J., and Lum, K. (2006) 'Selecting Best Practices for Effort Estimation' *IEEE Transactions on Software Engineering,* Vol. 32, No. 11, pp. 883 – 895.

Miranda, E., and Abran, A. (2008) 'Protecting Software Development Projects Against Underestimation', *Project Management Journal*, Vol. 39, No. 3, pp. 75-85.

Mittas, N. and Angelis, L. (2008) 'Combining Regression and Estimation by Analogy in a Semi-parametric Model for Software Cost Estimation': the ESEM'08, October 9–10, 2008, Kaiserslautern, Germany.

Murthi, S. (2002) 'Preventive Risk Management for Software Projects', *IT Professional*, Vol. 4, No. 5, pp. 9-15.

Pandian C.R. (2004) *Software Metrics – A Guide to Planning, Analysis and Application*, Boca Raton, FL, USA: Auerbach Publications

Papazoglou, M. P., Ribbers, P. and Tsaglatidou, A. (2000), 'Integrated Value Chains and their implications from a business and technology standpoint', *Decision Support Systems*, 29, pp. 323-342.

Persson, J.S. and Mathiassen, L. (2010) 'A Process for Managing Risks in Distributed Teams', *IEEE Software*, Vol. 27, No. 1, pp. 20-29.

Pfleeger, S.L. (2008) 'Software Metrics: Progress after 25 Years?', *IEEE Software,* Vol. 25 No. 6, pp. 32-34

Pimenidis, E. and Miller, C.J. (2006) 'Evaluating Risk in E-Business Development Projects', *the IADIS International Conference E-Commerce 2006.* Barcelona, Spain, 9-11 December 2006, pp. 81-88.

Pimenidis, E. and Miller, C.J. (2005) 'Risk and E-Business Development Projects: in Search of a New Metric', *the IADIS International Conference E-Commerce 2005.* Porto, Portugal, 15-17 December 2005, pp. 303-311.

Pimenidis E., Miller C. J. and MacEachen, C.F. (2004) 'Identifying Risk for e-Business Development': *the European Conference on Information Technology Evaluation (ECITE2004)*, Amsterdam, The Netherlands, 11-12 November 2004pp. 349-358.

Pimenidis E., Miller C. J. and MacEachen C.F. (2002) 'A New Look at Project Estimation for E-Business': *the International Conference on Project Management "Breakthrough with Project Management – In the Era of Global Revolution by IT" (ProMAC2002),* Singapore, 31 July – 2 August 2002, pp. 353-358

Rajlich, V.T. and Bennett, K.H. (2000) 'A Staged Model for the Software Life Cycle', *IEEE Computer*, Vol. 33, No. 7, pp 66 - 71.

Reifer, D.J. (2000) 'Web Development: Estimating Quick-to-Market Software', *IEEE Software*, Vol.17, No. 6, pp. 57 – 64.

Rifkin, S. (2001) 'What Makes Measuring Software So Hard?', *IEEE Software*, Vol. 18, No. 2, pp. 41 – 45.

Shi, X. (2007) 'Semantic Web Services: An Unfulfilled Promise', *IT Professional,* Vol. 9, No. 4, pp. 42-45.

SKY.com (2011) 'Olympics Ticket Chaos Frustrates Fans Again' [Online]. Available at http://news.sky.com/skynews/Home/UK-News/Olympics-Fans-Left-Frustrated-Again-At-Ticket-Sales-As-Website-Struggles-To-Cope-With-Huge-Demand/Article/201106416017911?DCMP=Web-search-sslc&lid=ARTICLE_16017911_OlympicsFansLeftFrustratedAgainAtTicketSalesAsWebsite StrugglesToCopeWithHugeDemand&lpos=searchresults  (Accessed 6 July 2011)

Shepperd, M. and Schofield, C. (1997) 'Estimating Software Project Effort Using Analogies', *IEEE Transactions on Software Engineering*, Vol. 23, No. 12, pp 736-743.

Shepperd, M., Schofield, C., and Kitchenham, B. (1996) 'Effort Estimation Using Analogy', *International Conference on Software Engineering (ICSE 1996):* March 25-30, 1996, Berlin, Germany, pp. 170 – 178.

Sherer, S. A. (2004) Managing Risk beyond the Control of IS Managers: The Role of Business Management, Proceedings of the 37th Hawaii International Conference on System Sciences (HICSS'04), IEEE.

Smith, R.K., Hale, J.E. and Parrish, A.S. (2001) 'An Empirical Study Using Task Assignment Patterns to Improve the Accuracy of Software Effort Estimation', *IEEE Transactions on Software Engineering*, Vol. 27, No. 3, pp. 264 – 271

Smolander, K. (2003) 'The Birth of An E-Business System Architecture: Conflicts, Compromises, and Gaps in Methods': *the 36th Hawaii International Conference on System Sciences (HICSS-36 2003)*, Big Island, HI, USA, January 6-9, 2003.

Snow, A. P. and Keil M. (2001) 'The Challenge of Accurate Software Project Status Reporting: A two Stage Model Incorporating Status Errors and Reporting Bias': *the 34th Annual Hawaii International Conference on System Sciences (HICSS-34)*, 2001, Maui, Hawaii, USA, January 3-6 2001.

Stahl, B.C. (2007) 'Reflective responsibility for risk: a critical view of software and information systems development risk management', *International Journal of Risk Assessment and Management*, Vol. 7, No 3, pp. 312-325.

Starr, R., Newfrock, J., and Delurey, M. (2003) 'Enterprise Resilience: Managing Risk in the Networked Economy', *Strategy and Business*, Issue 30.

Schwalbe, K. (2010) *Managing Information Technology Projects.* 6th edn. Andover: CENGAGE Learning

Stamelos, I. and Angelis, L. (2001) Managing Uncertainty in Project Portfolio Cost Estimation, Information and Software Technology, 43 (2001) 759-768.

TA NEA (2011a) 'ΚΑΤΩΤΕΡΟΙ ΤΩΝ ΠΕΡΙΣΤΑΣΕΩΝ' [Online]. Available at http://www.tanea.gr/kyrioarthro/?aid=4629262 (Accessed 30 May 2011)

TA NEA (2011b) 'Στην ουρά για δήλωση στο TAXIS' [Online]. Available at http://www.tanea.gr/oikonomia/article/?aid=4629298

Tapscott, D. (2001) 'Rethinking Strategy in a Networked World (or Why Michael Porter is Wrong about the Internet)', *Strategy and Business,* Issue 24.

The Standish Group (1995) Chaos Report, [Online]. Available at http://www.projectsmart.co.uk/docs/chaos-report.pdf (Accessed 15 June 2011)

Tesch, D., Kloppenborg T.J. and Frolick, M.N. (2007) IT Project Risk Factors: The Project Management Professionals Perspective, *The Journal of Computer Information Systems,* Vol. 47, No. 4, pp.61-69

The Standish Group (2011) CHAOS Manifesto, [Online]. Available at http://standishgroup.com/newsroom/chaos_manifesto_2011.php (Accessed 15 June 2011)

The Telegraph (2011) 'London 2012 Olympics: Ticketmaster apologise to 700 fans double-charged in games ticket rush' [Online]. Available at http://www.telegraph.co.uk/sport/othersports/olympics/8633564/London-2012-Olympics-Ticketmaster-apologise-to-700-fans-double-charged-in-games-ticket-rush.html (Accessed 16 July 2011)

Tiwana, A. and Keil, M., (2006) 'Functionality Risk in Information Systems Development: An Empirical Investigation', *IEEE Transaction on Engineering Management*, Vol. 53, No. 3, pp. 412-425.

University of Southern California (1994) USC COCOMO Reference Manual,[Online]. Available at http://sunset.usc.edu/research/COCOMOII/cocomo81docs/cocomo.pdf (Accessed: 25 May 2011)

University of Southern California (2000) USC COCOMO II 2000 Software Reference Manual,[Online]. Available at http://csse.usc.edu/csse/research/COCOMOII/cocomo2000.0/CII_manual2000.0.pdf (Accessed: 25 May 2011)

Venkatraman, N. (2000) Five steps to a Dot-Com Strategy: How to find your footing on the web, *Sloan Management Review*, Vol. 41, No. 3, pp. 15-28.

Wallace, L., Keilb, M., and Raic, A. (2004) 'Understanding software project risk: a cluster analysis' *Information & Management* 42 (2004) 115–125

Willcocks, L.P. and Plant R. (2001) 'Pathways to E-Business Leadership: Getting from Bricks to Clicks', *Sloan Management Review*, Vol. 42, No. 3, pp. 50 – 59.

Willis, S. (2002) 'Using QA for Risk Management in Web Projects', in Meyerhoff, D., Laibarra, B., van Der Pouw Kraan, R. and Wallet, A., *Software Quality and Software Testing in Internet Times*, Berlin; Springer-Verlag, pp. 31-48.

Wu, H. and Cao, L. (2009) 'Community Collaboration for ERP Implementation', *IEEE Software,* Vol. 26, No. 6, pp. 48-55.

Zafra-Cabeza, A., Ridao, M.A. and Camacho, E.F. (2008) 'Using a risk-based approach to project scheduling: A case illustration from the semiconductor manufacturing', *European Journal of Operational Research*, 190(2008), pp. 708-723.

Zhang, Y. and Sheth, D. (2006) 'Mining Software Repositories for Model-Driven Development', *IEEE Software*, Vol. 23, No. 1, pp. 82-90.

# Bibliography

Abran, A. and Robillard, P.N. (1996) 'Function Point Analysis: An Empirical Study of its Measurement Processes', *IEEE Transactions on Software Engineering*, Vol. 22, No.12, pp. 895 – 910.

Azzeh, M., Neagu, D. and Cowling, P. (2008) 'Improving Analogy Software Effort Estimation using Fuzzy Feature Subset Selection Algorithm': *the Fourth International Workshop on Predictor Models in Software Engineering (PROMISE'08)*, Leipzig, Germany, May 12–13, 2008

Baskerville, R., Levine, L., Pries-Heje, J., Ramesh B. and Slaughter, S. (2001) 'How Internet Software Companies Negotiate Quality', *IEEE Computer*, Vol. 34, No. 5, pp. 51 – 57.

Bielak, J. (2000) 'Improving Size Estimates Using Historical Data', *IEEE Software*, Vol.17, No. 6, pp. 27 – 35.

Boehm, B. (2000) 'Project Termination Doesn't Equal Project Failure', *IEEE Computer*, Vol. 33, No. 9, pp. 94 – 96.

Coates, R. (2000) 'How e is your Company?', *IT Strategy:* The British Computer Society, Autumn 2000, pp. 21-23.

Drobik, A. (2000) 'E-business is not easy business', *The Computer Bulletin*: The British Computer Society, January 2000.

Fenny, D. (2001) 'Making Business Sense of the E-Opportunity', *Sloan Management Review*, Vol. 42, No. 2, pp. 41-51.

Isakowitz, T., Stohr, E.A. and Balasubramanian P. (1995) 'RMM: a methodology for structured hypermedia design', *Communications ACM*, Vol. 38, No. 8, pp. 34-44.

Johnson, P.M., Moore, C.A., Dane, J.A. and Brewer, R.S. (2000) 'Empirically Guided Software Effort Guesstimation', *IEEE Software*, Vol.17, No. 6, pp. 50 – 56.

Guadrado-Gallego, J.J., Rodriguez, D., Sicilia, M.A., Rubio, M.G. and Crespo, A.G. (2007) 'Software Project Effort Estimation Based on Multiple Parametric Models Generated Through Data Clustering', *Journal of Computer Science and Technology*, Vol. 22, Issue 3, pp. 371-378.

Han, W-M. and Huang, S-J.(2007) 'An empirical analysis of risk components and performance on software projects', *The Journal of Systems and Software*, No. 80, pp. 42–50.

Hinz, D.J., (2005) 'High Severity Information Technology Risks in Finance', *the 38th Hawaii International Conference on System Sciences (HICSS-38 2005)*, Big Island, HI, USA, 3-6 January 2005

Hughes, B. (2000) *Practical Software Measurement*, New York, USA: McGraw-Hill

Kitchenham, B., Pfleeger S.L. and Fenton N. (1995) 'Towards a Framework for Software Measurement Validation', *IEEE Transactions on Software Engineering*, Vol. 21, No. 12, pp. 929 – 944.

Kontio, J., Jokinen, J-P. and Rosendahl, E. (2004) 'Visualizing and Formalizing Risk Information: An Experiment': *the 10th IEEE International Software Metrics Symposium (METRICS 2004)*, Chicago, IL, USA, 11-17 September 2004.

Laseter, T., Houston, P., Chung, A., Byrne, S., Turner. M. and Devendran, A. (2000) 'The Last Mile to Nowhere', *Strategy and Business*, Issue 20.

Lientz B.P. and Larson L. (2006) Risk Management for IT Projects, Butterworth-Heinemann,

Li, J. and Ruhe, G. (2008) 'Multi-criteria Decision Analysis for Customization of Estimation by Analogy Method AQUA': *the Fourth International Workshop on Predictor Models in Software Engineering (PROMISE'08)*, Leipzig, Germany, May 12–13, 2008.

Li, J. and Ruhe, G. (2007) 'Decision Support Analysis for Software Effort Estimation by Analogy': *the Third International Workshop on Predictor Models in Software Engineering (PROMISE'07)*, Minneapolis, MN, USA, 20-26 May 2007

Li, J., Conradi, R., Slyngstad, O.P.N., Torchiano, M. and Morisio M. (2008) 'A State-of-the-Practice Survey of Risk Management in Development with Off-the-Shelf Software Components', *IEEE Transactions on Software Engineering*, Vol. 34, No. 2, pp. 271-286

Liu, K. and Ong, T. (1999) 'A Modelling Approach for Handling Business Rules and Exceptions', *The Computer Journal*: The British Computer Society, Vol. 42, No. 3, pp. 221-231. doi:10.1093/comjnl/42.3.221

MacCormack, A. (2001) 'Product Development Practices That Work: How Internet Companies Build Software', *Sloan Management Review*, Vol. 42, No. 2, pp. 75 – 84.

McConnell S. (2006) Software Estimation – Demystifying the Black Art, Microsoft Press

Mendes, E. (2000) 'Investigating Metrics for a Development Effort Prediction Model of Web Applications': *the 12th Australian Software Engineering Conference (ASWEC 2000)*, Canberra, Australia, 28-30 April 2000.

Mendes, E. and Counsell, S. (2000) 'Web Development Effort Estimation Using Analogy': *the 12th Australian Software Engineering Conference (ASWEC 2000)*, Canberra, Australia, 28-30 April 2000.

Munson J.C. (2003) Software Engineering Measurement, Auerbach Publications

Myrtveit, I. and Stensrud, E. (1999) 'A Controlled Experiment to Access the Benefit of Estimating with Analogy and Regression Models', *IEEE Transactions on Software Engineering*, Vol. 25, No. 4, pp. 510 – 525.

Peters, R.J., and Verhoef, C. (2008) 'Quantifying the yield of risk-bearing IT-portfolios', *Science of Computer Programming* 71 (2008) 17–56

Poulin L. (2005) Reducing Risk with Software Process Improvement, Auerbach Publications

Prechelt, L. (2011) 'Plat_Forms: A Web Development Platform Comparison by an Exploratory Experiment Searching for Emergent Platform Properties' *IEEE Transactions on Software Engineering,* Vol. 37, No. 1, pp. 95-108.

Sadiq, M., Rahman, A., Ahmad, S., Asim, M. and Ahmad, J. (2010) 'esrcTool: A Tool to Estimate the Software Risk and Cost': *the Second International Conference on Computer Research and Development*, 7-10 May 2010, Kuala Lumpur, Malaysia, pp. 886-890.

Seaman, C.B. (1999) 'Qualitative Methods in Empirical Studies of Software Engineering', *IEEE Transactions on Software Engineering*, Vol. 25, No. 4, pp. 557 – 572.

Shim, S.S.Y., Pendyala, V.S., Sundaram. M., and Gao, J. Z. (2000) 'Business-to-Business E-Commerce Frameworks', *IEEE Computer*, Vol. 33, No.10, pp. 40 - 47.

Takagi, Y., Mizuno, O. and Kikuno, T. (2005) 'An Empirical Approach to Characterizing Risky Software Projects Based on Logistic Regression Analysis', *Empirical Software Engineering*, Vol. 10, pp. 495-515.

Urban, G.L, Sultan, F. and Qualls, W.J. (2000) 'Placing Trust at the Center of Your Internet Strategy', *Sloan Management Review*, Vol. 42, No. 1, pp. 39-48.

Verdon, D. and McGraw, G. (2004) 'Risk Analysis in Software Design', *IEEE Security and Privacy*, Vol. 2, No. 4, pp. 79-84.

Verner, J.M. and Evanco, W.M., (2005) 'In-House Software Development: What Project Management Practices Lead to Success?', *IEEE Software*, Vol. 22, No. 1, pp. 86-93.

Wang, W., Hidvegi, Z., Bailey Jr, A.D. and Whinston, A.B. (2000) 'E-Process Design and Assurance using Model Checking', *IEEE Computer*, Vol. 33, No. 10, pp. 48 - 53.

Woolridge, R.W., McManus, D.J., and Hale, J.E., (2007) 'Stakeholder Risk Assessment: An Outcome-Based Approach', *IEEE Software*, Vol. 24, No. 3, pp. 36-45.

Zhu, H., Siegel, M. D. and Madnick, S. E. (2001) 'Information Aggregation -A Value-added E-Service', *the International Conference on Technology, Policy, and Innovation: Critical Infrastructures*, The Netherlands, June 26-29, 2001.

# APPENDIX A

# Survey Questionnaire

Table 1a. Preliminary Risk List

1. Project scope and objectives are inappropriately defined
2. Complex and unclear relationships between partners, customers and suppliers
3. Disagreement between involved partners
4. Lack of previous experience by the customer
5. Lack of clear definition of development methodologies or/and technological infrastructures
6. Lack of planning for replacement of current systems or/and interfacing with current systems
7. Lack of backup plan for delays or/and under-performance of new system
8. Significant need for re-engineering of current business processes
9. Inappropriate business plan and IS vision
10. Lack of senior management support or/and internal political resistance
11. Inefficient communication between all involved parties
12. Inexperienced team members in core business or technology project components
13. Emerging or unproven technologies
14. Unfamiliar technologies to the design and development team
15. Emerging or unproven programming and debugging technologies
16. Unfamiliar development environment to the project team
17. Project involves use of technology that has not been used in prior projects
18. Large number of links to other systems required
19. High level of technical complexity
20. Project involved the use of new technology
21. Immature technology
22. Highly complex task being automated
23. Conflict between users
24. Users with negative attitudes toward the project
25. Personnel shortfalls

Survey Questionnaire

1. What type of projects have you been involved with?

    a. ☐      web-site development
    b. ☐      e-business development with integrated systems
    c. ☐      e-learning
    d. ☐      other (please state):

2. If integrated systems how many systems were involved?

    a. ☐      2
    b. ☐      3
    c. ☐      4
    d. ☐      5
    e. ☐      5+

3. How many organizations were involved in the project?

    a. ☐      1
    b. ☐      2
    c. ☐      3
    d. ☐      4
    e. ☐      5
    f. ☐      5+

4. How many different types of systems were involved?

    a. ☐      1
    b. ☐      2
    c. ☐      3
    d. ☐      4
    e. ☐      5
    f. ☐      5+

5. What types of systems did the integration involve (select all that apply)?

    a. ☐      e-procurement
    b. ☐      e-payment
    c. ☐      e-learning
    d. ☐      Enterprise Resource Planning (ERP)
    e. ☐      Management Information Systems (MIS)

f. ☐      Decision Support Systems (DSS)

g. ☐      Other (please state):

6. How many different implementation teams did the project involve?

a. ☐      1
b. ☐      2
c. ☐      3
d. ☐      4
e. ☐      5
f. ☐      5+

From the table below, complete those rows as deemed appropriate, identifying sources of risk for the project failing to complete on time or exceeding the budget. Please grade entries against:

- Frequency: (How often has any of them been encountered).
- Severity: (What would the effect of such an event occurring be).

| SOURCE | FREQUENCY (select one only) | | | | SEVERITY (select one only) | | | |
|---|---|---|---|---|---|---|---|---|
| | FREQUENT | REASONABLY PROBABLE | OCCASIONAL | EXTREMELY UNLIKELY | CATASTROPHIC | CRITICAL | MARGINAL | NEGLIGIBLE |
| Inexperience of developing team. | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| New systems to be integrated but not previously tested. | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| Legacy systems not compatible with modern technology. | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| Loss of expert resources (members of the | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| development team). | | | | | | | | |
| Low commitment of individual partners. | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| Different priorities in terms of launch time | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| Difference in readiness of partners to function on the e-business model. | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| Other (please fill in - a): | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| Other (please fill in - b): | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |

The following information would be useful to streamline the results of this research and although not essential would be highly appreciated:

Industry type :

Job Title :

Would you like a complimentary copy of the relevant article to be sent to you when published? ☐

Please return the completed questionnaire by email.

# Appendix B

# Risk Variation Graphs
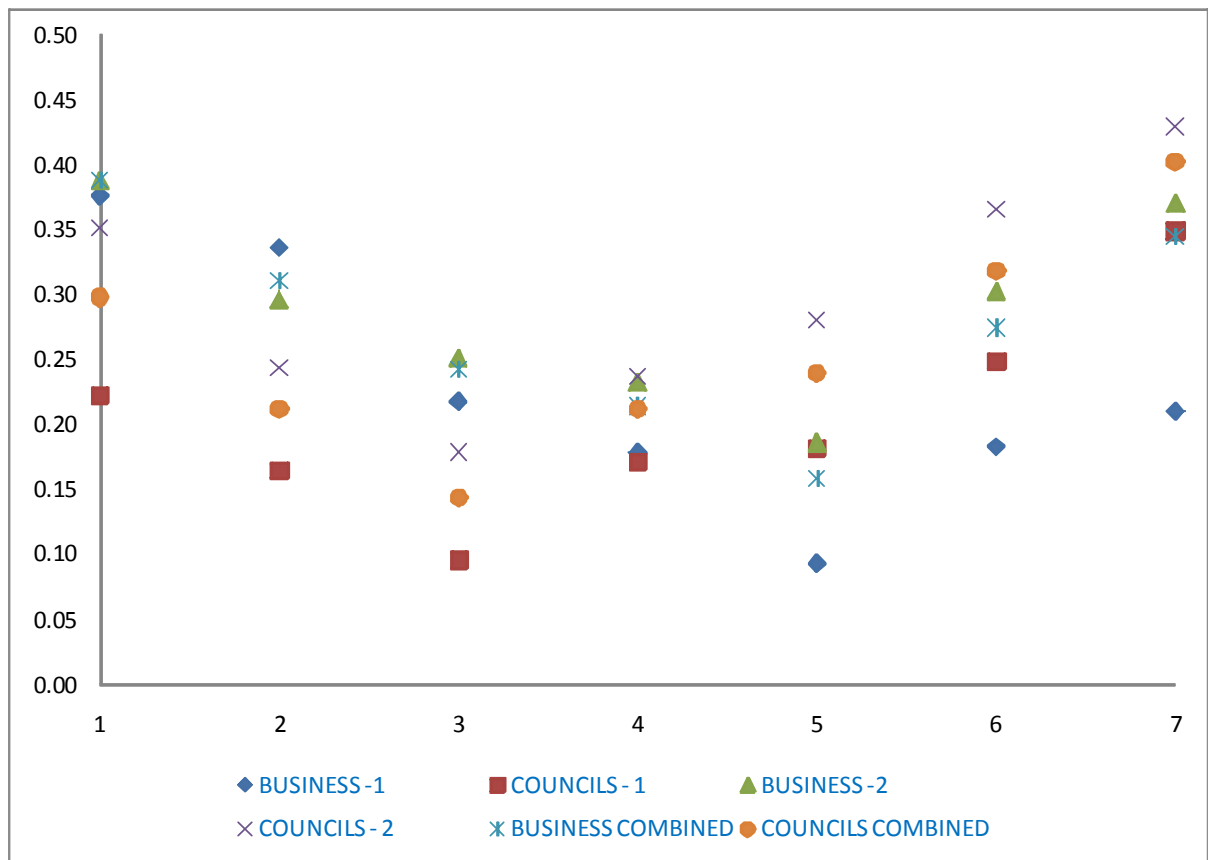
Figure B1. Graph of Normalized Risk Ratings ($i^2$)
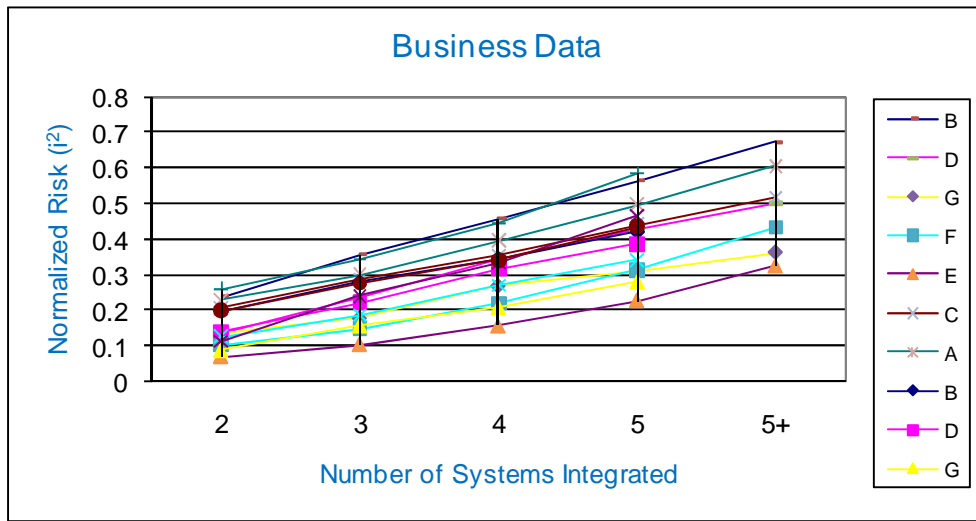
Figure B2. Variation of Individual Risks with Number of Systems Integrated (business data)
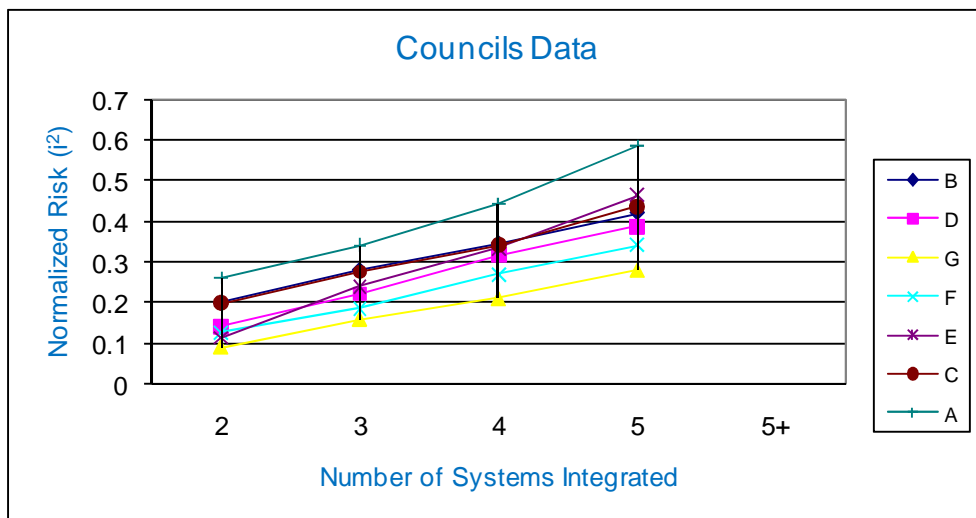


Figure B3. Variation of Individual Risks with Number of Systems Integrated (Councils data)
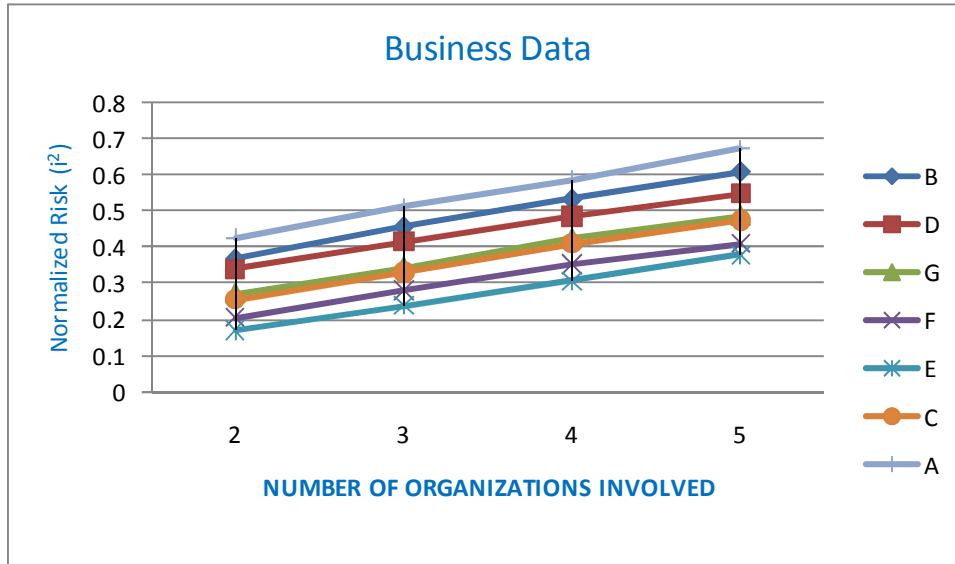
Figure B4. Variation of Individual Risks with Number of Organizations Involved

(Business data)
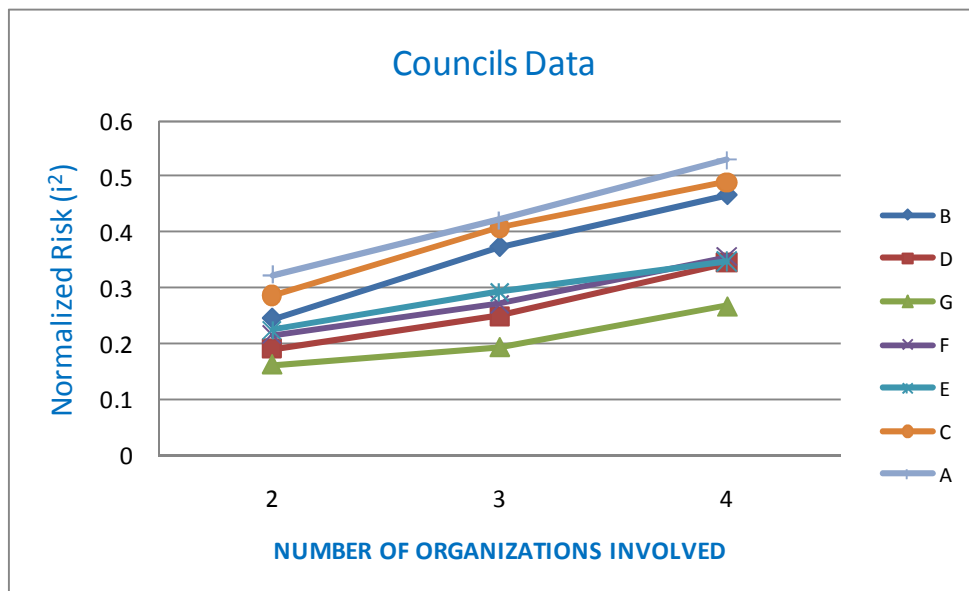


Figure B5. Variation of Individual Risks with Number of Organizations Involved

(Councils data)
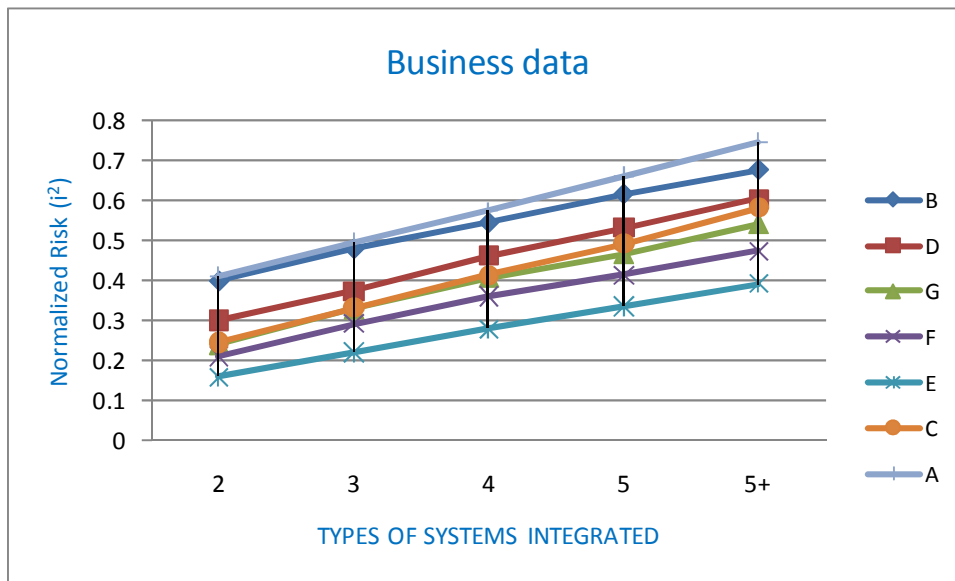
Figure B6. Variation of Individual Risks with Number of Types of Systems Integrated
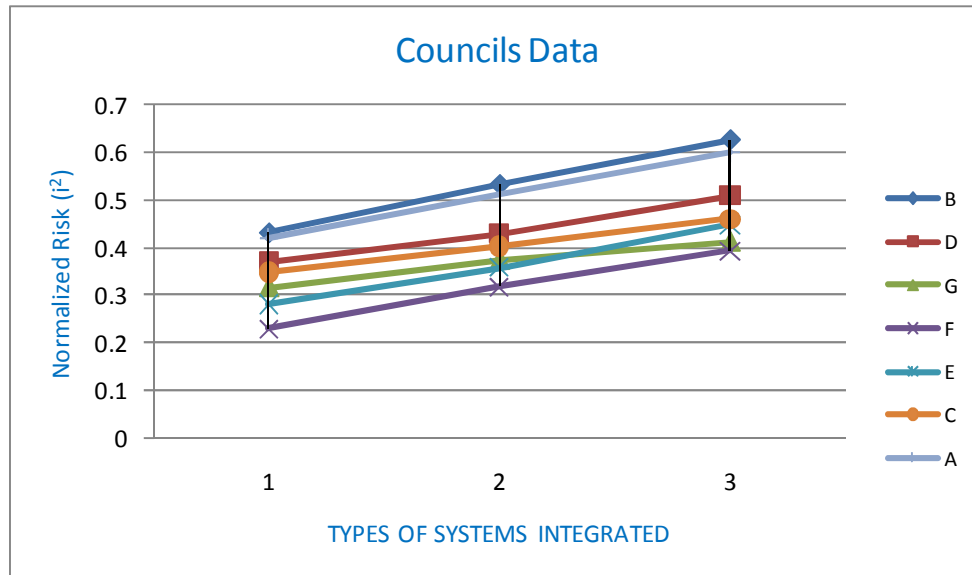
(Business data)



Figure B7. Variation of Individual Risks with Number of Types of Systems Integrated
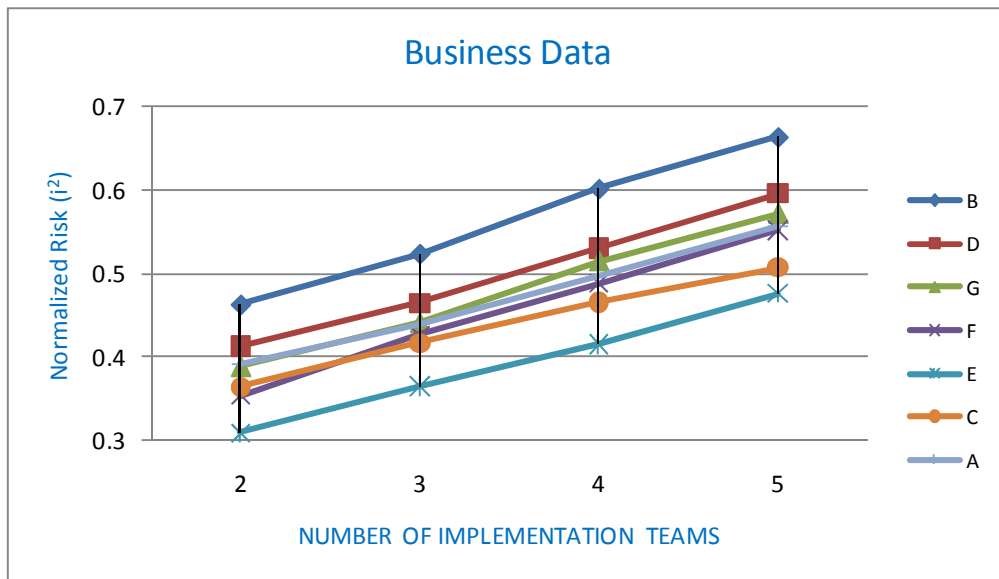
(Councils data)

Figure B8. Variation of Individual Risks with Number of Implementation Teams

(Business data)
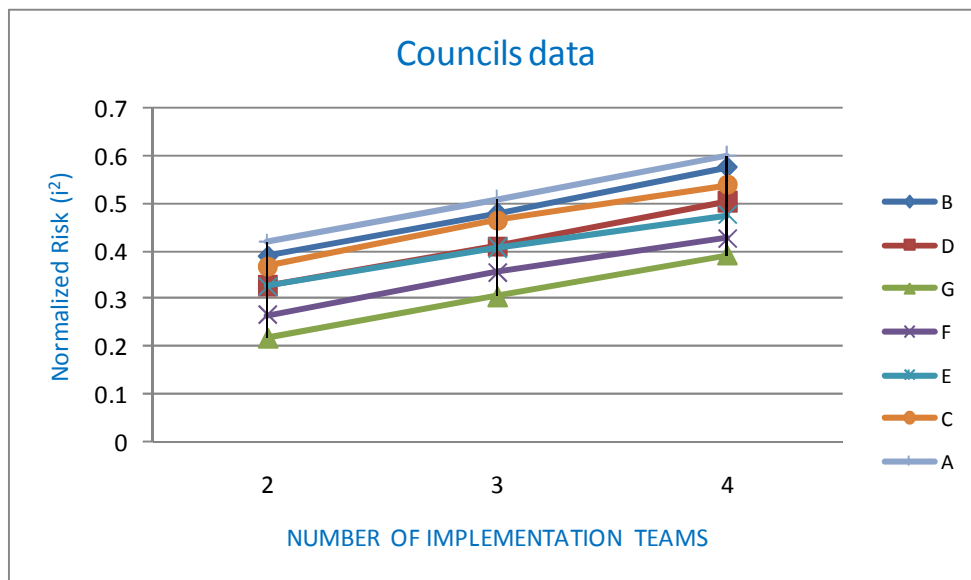


Figure B9. Variation of Individual Risks with Number of Implementation Teams

(Councils data)

Table 6. Average Numbers of project attributes

| PROJECT ATTRIBUTES | AVERAGE NUMBERS | |
| --- | --- | --- |
| | BUSINESS | COUNCILS |
| Systems Integrated | 4.05 | 2.89 |
| Organisations Involved | 3.00 | 2.38 |
| Types of Systems | 3.57 | 2.00 |
| Implementation Teams | 2.87 | 2.29 |

# APPENDIX C

# Calculations

Table 7. Project Attributes (of the four project cases)

| PROJECT | NUMBER OF SYSTEMS | NUMBER OF TYPES OF SYSTEMS | NUMBER OF ORGANIZATIONS INVOLVED | NUMBER OF IMPLEMENTATION TEAMS | ACTUAL DURATION |
|---|---|---|---|---|---|
| LB HAVERING 1 | 4 | 3 | 3 | 4 | 205 |
| LB HAVERING 2 | 5 | 5 | 4 | 5 | 426 |
| PROTOPORIA | 6 | 4 | 6 | 2 | 426 |
| ELGEKA | 5 | 4 | 4 | 5 | 262 |
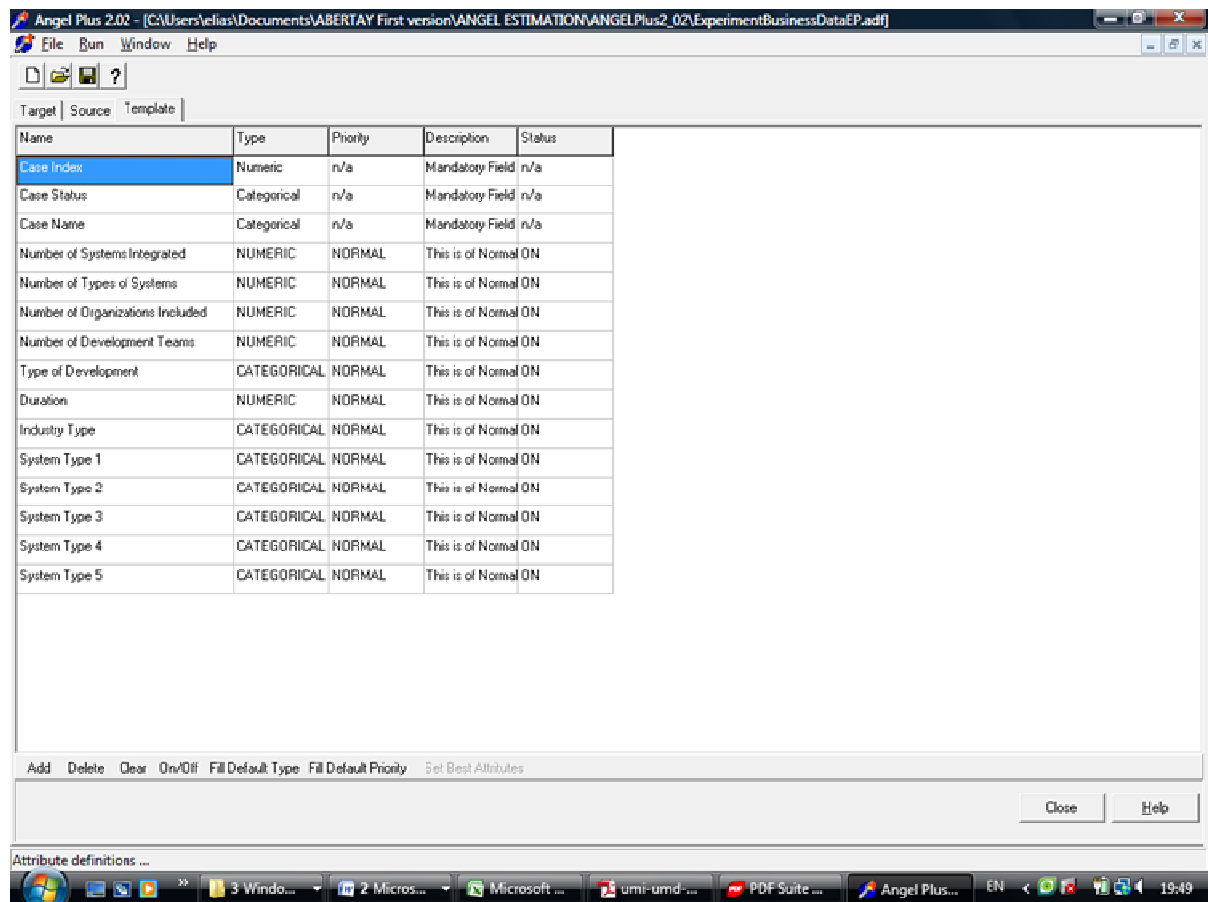


Figure C1. Defining a template for the ISBSG Business data
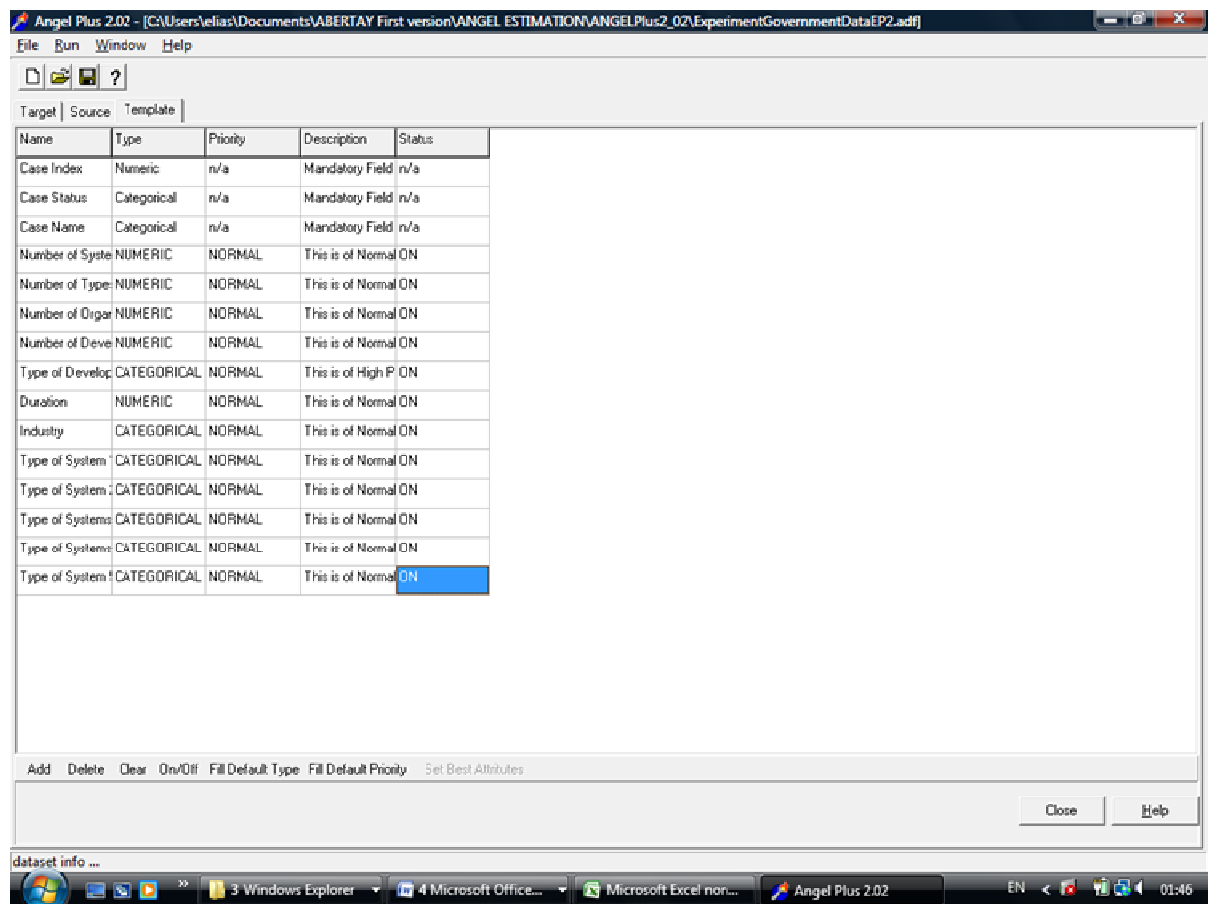
Figure C2. Defining a template for the ISBSG Councils data
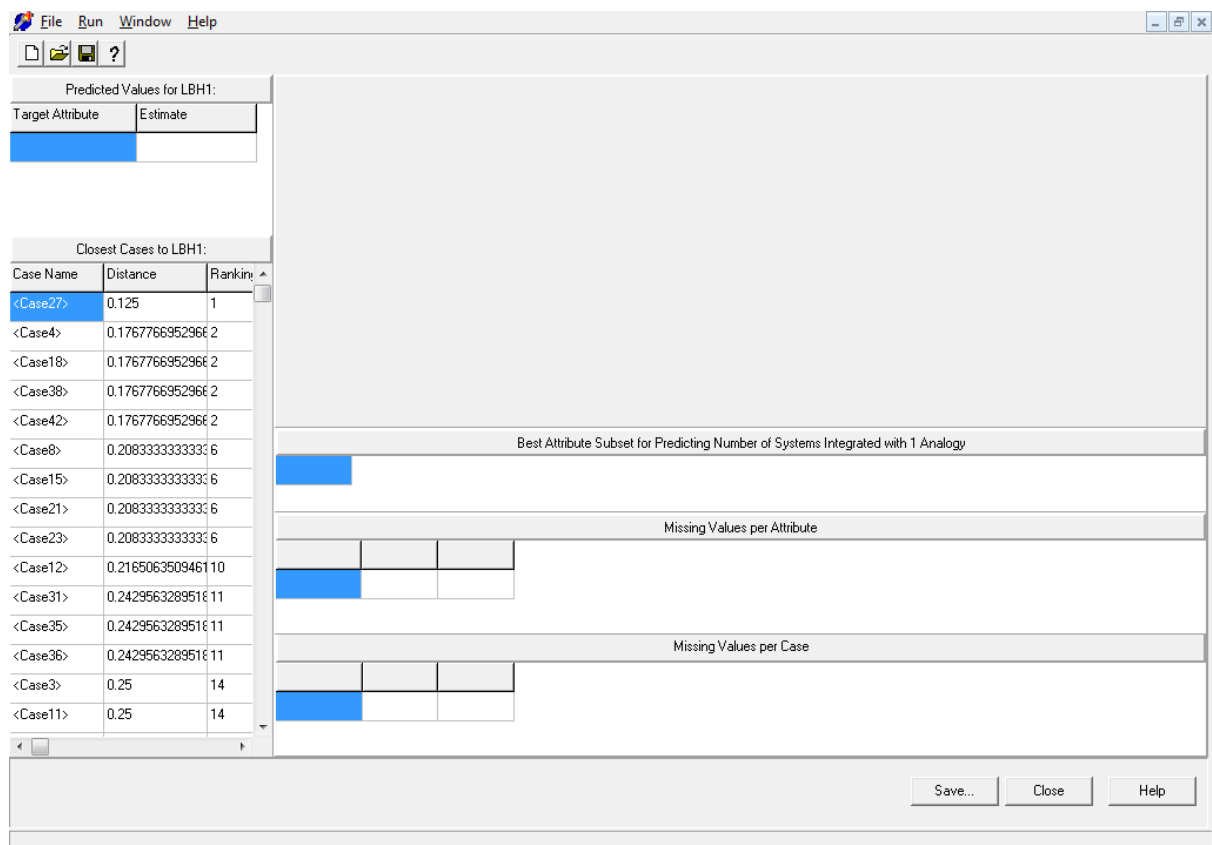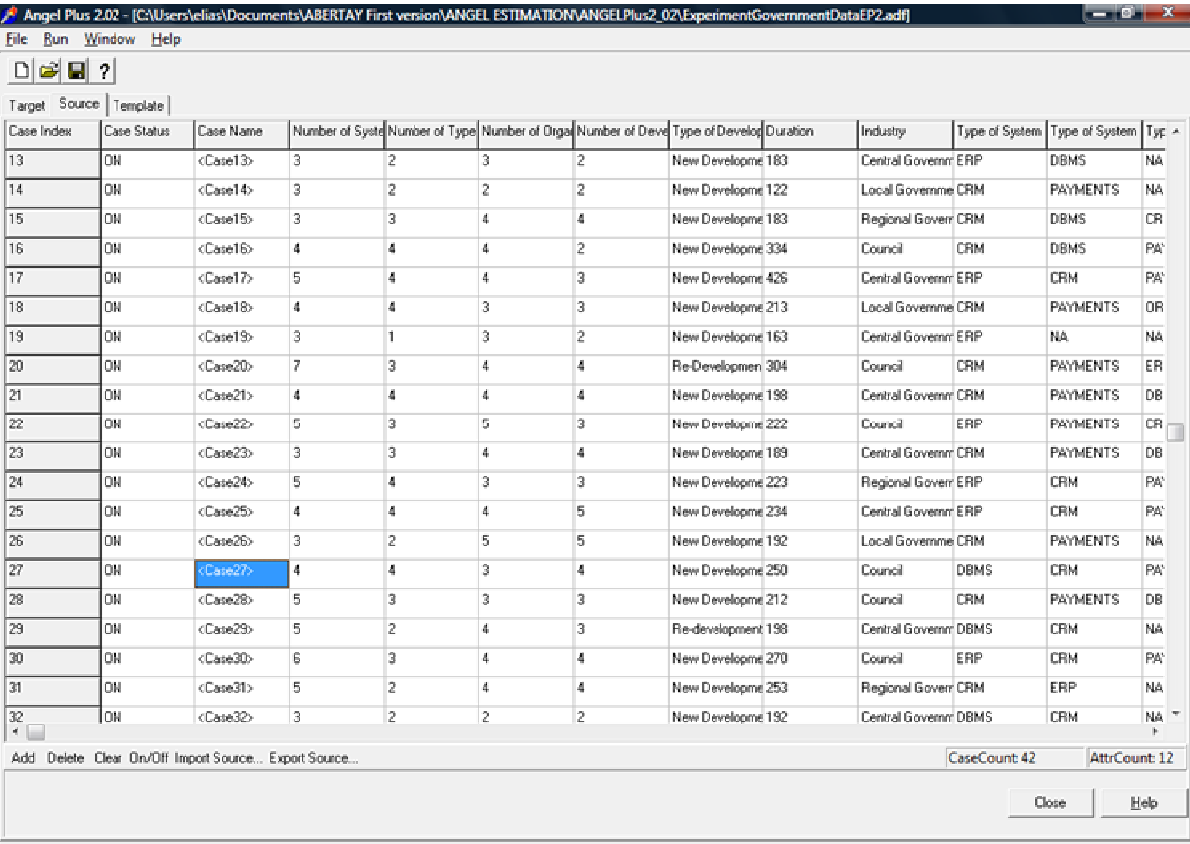
Figure C3 – Using the ANGEL tool to identify a matching case (project analogue) for the LB of Havering project 1.

# Estimating the 1<sup>st</sup> LB of Havering project



Figure C4. Details of the identified project analogue for the LB of Havering project 1

## Calculating the estimated duration of project 1 of the LB of Havering

The non matching attribute for Project 1 of the LB of Havering and Case 27 of the Councils

data model is that of number of types of systems integrated. The corresponding equation for

the risk rating variation is, $y = 0.0856x + 0.2455$ (from table 6).

Applying the equation to the project yields $y=0.0856*3+0.2455 = 0.5023$

Applying the equation to case 27 yields $y=0.0856*4+0.2455=0.5879$

The analogy is thus defined as the ratio $(0.5023) / (0.5879) = 0.854397$

The estimated duration is given by multiplying the analogy by the duration of case 27, i.e.

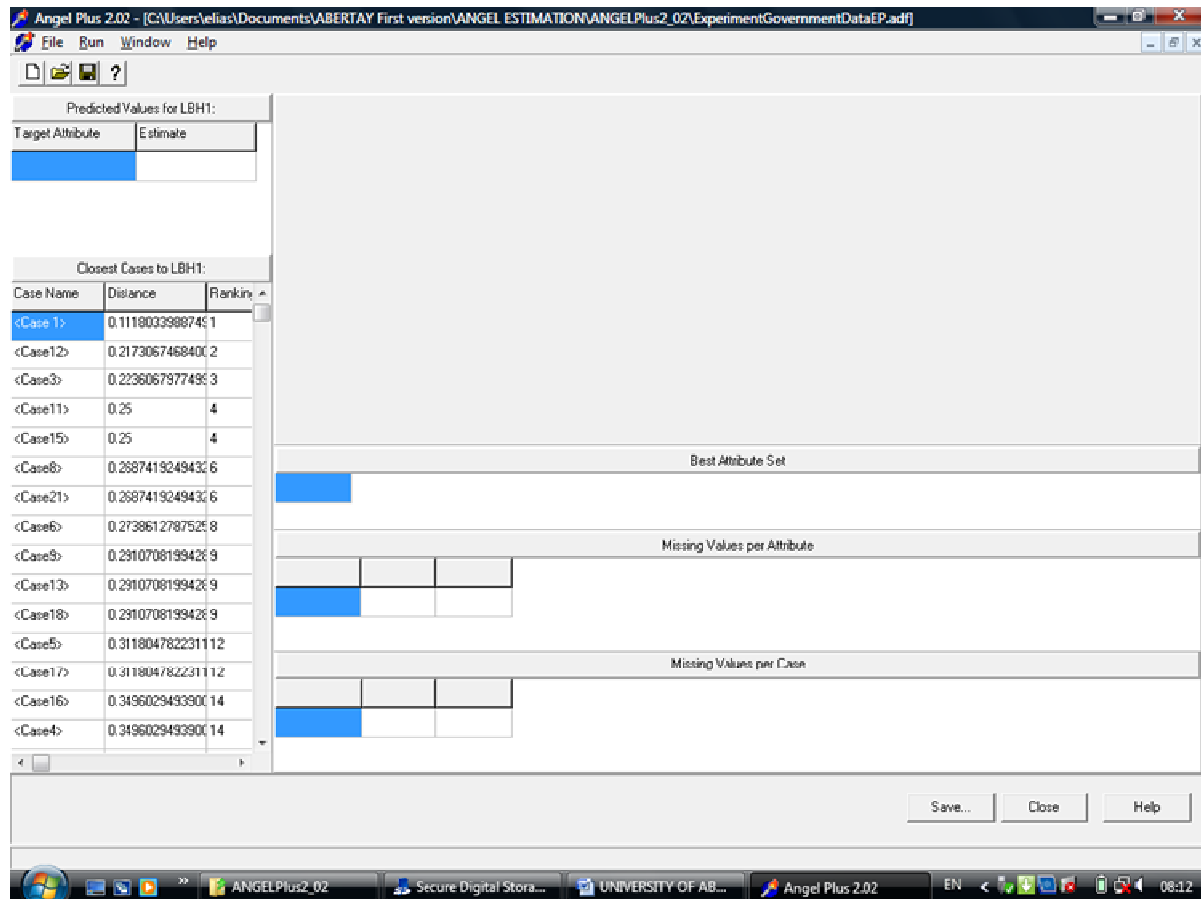$0.854397*250 = 214$ days (approximately)

# Estimating Project 2 of LB of Havering
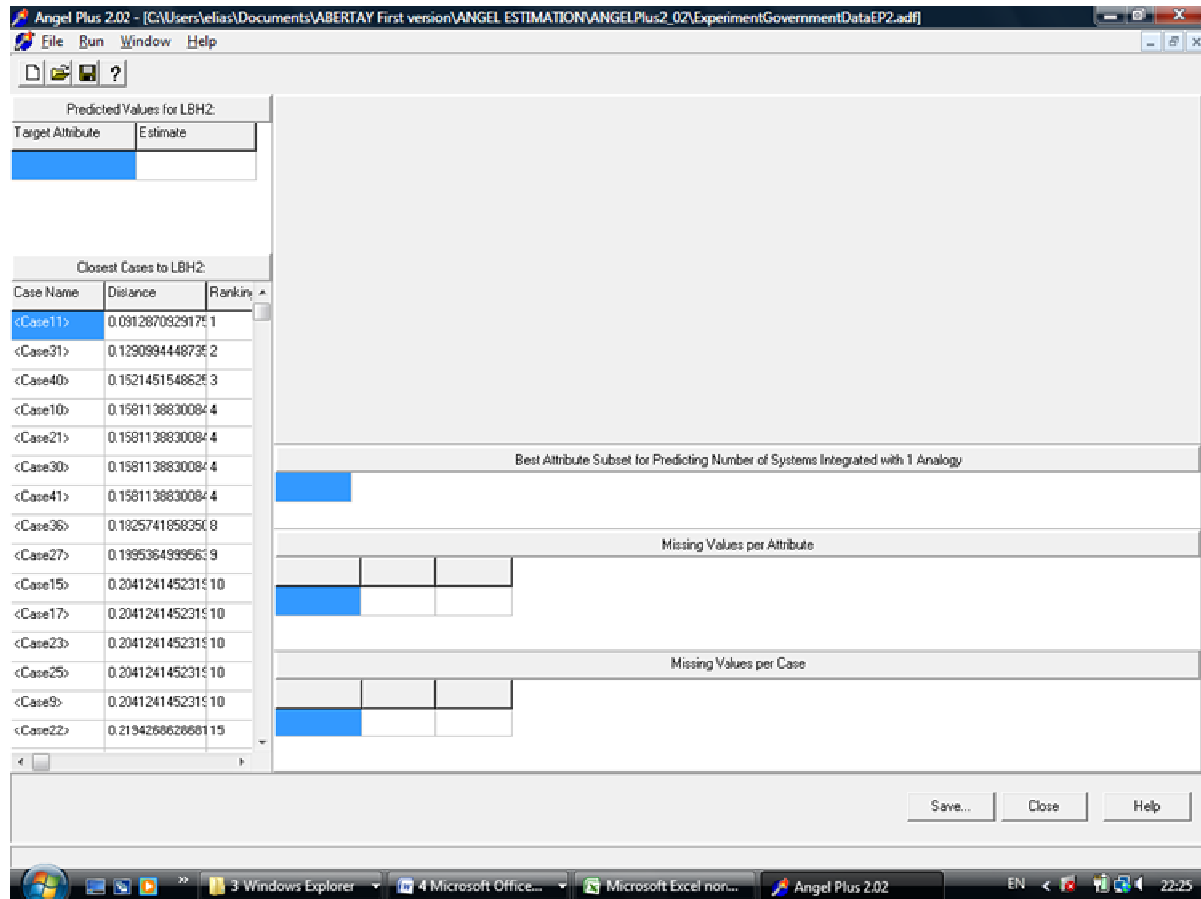


Figure C5 – Using the ANGEL tool to identify a matching case (project analogue) for the LB of Havering project 2.

Figure C6. Details of the identified project analogue for the LB of Havering project 2

## Calculating the estimated duration of project 2 of the LB of Havering

The non matching attribute for Project 2 of the LB of Havering and Case 11 of the Councils

data model is that of number of systems integrated. The corresponding equation for the risk

rating variation is, $y = 0.0824x + 0.1523$ (from table 6).

Applying the equation to the project yields $y = 0.0824*5 + 0.1523 = 0.5643$

Applying the equation to case 11 yields $y = 0.0824*4 + 0.1523 = 0.4819$

The analogy is thus defined as the ratio $(0.5643) / (0.4819) = 1.17099$

The estimated duration is given by multiplying the analogy by the duration of case 11, i.e.

$1.17099*391 = 458$ days (approximately)

# Estimating the "Protoporia" Project



Figure C7 – Using the ANGEL tool to identify a matching case (project analogue) for "Protoporia" project.

Figure C8. Details of the identified project analogue for "Protoporia" project

## Calculating the estimated duration of the "Protoporia" project

The non matching attribute for the "Protoporia" project and Case 33 of the Business data model is that of number of organizations involved. The corresponding equation for the risk rating variation is, $y = 0.0734x + 0.219$ (from table 6).

Applying the equation to the project yields $y = 0.0734*6 + 0.219 = 0.6594$

Applying the equation to case 33 yields $y = 0.0734*5 + 0.219 = 0.586$

The analogy is thus defined as the ratio $(0.6594) / (0.586) = 1.125256$

The estimated duration is given by multiplying the analogy by the duration of case 33, i.e.

$1.125256*388 = 437$ days (approximately)

# Estimating the "ELGEKA" project



Figure C9 – Using the ANGEL tool to identify a matching case (project analogue) for "ELGEKA" project.

Figure C10. Details of the identified project analogue for "ELGEKA" project

## Calculating the estimated duration of the "ELGEKA" project

The non matching attribute for the "ELGEKA" project and Case 55 of the Business data model is that of number of systems integrated. The corresponding equation for the risk rating variation is, $y = 0.0823x + 0.0705$ (from table 6).

Applying the equation to the project yields $y = 0.0823*5 + 0.0705 = 0.482$

Applying the equation to case 55 yields $y = 0.0823*4 + 0.0705 = 0.3997$

The analogy is thus defined as the ratio $(0.482) / (0.3997) = 1.205904$

The estimated duration is given by multiplying the analogy by the duration of case 55, i.e. $1.205904*204 = 246$ days (approximately)

# Estimating using own collected data independently

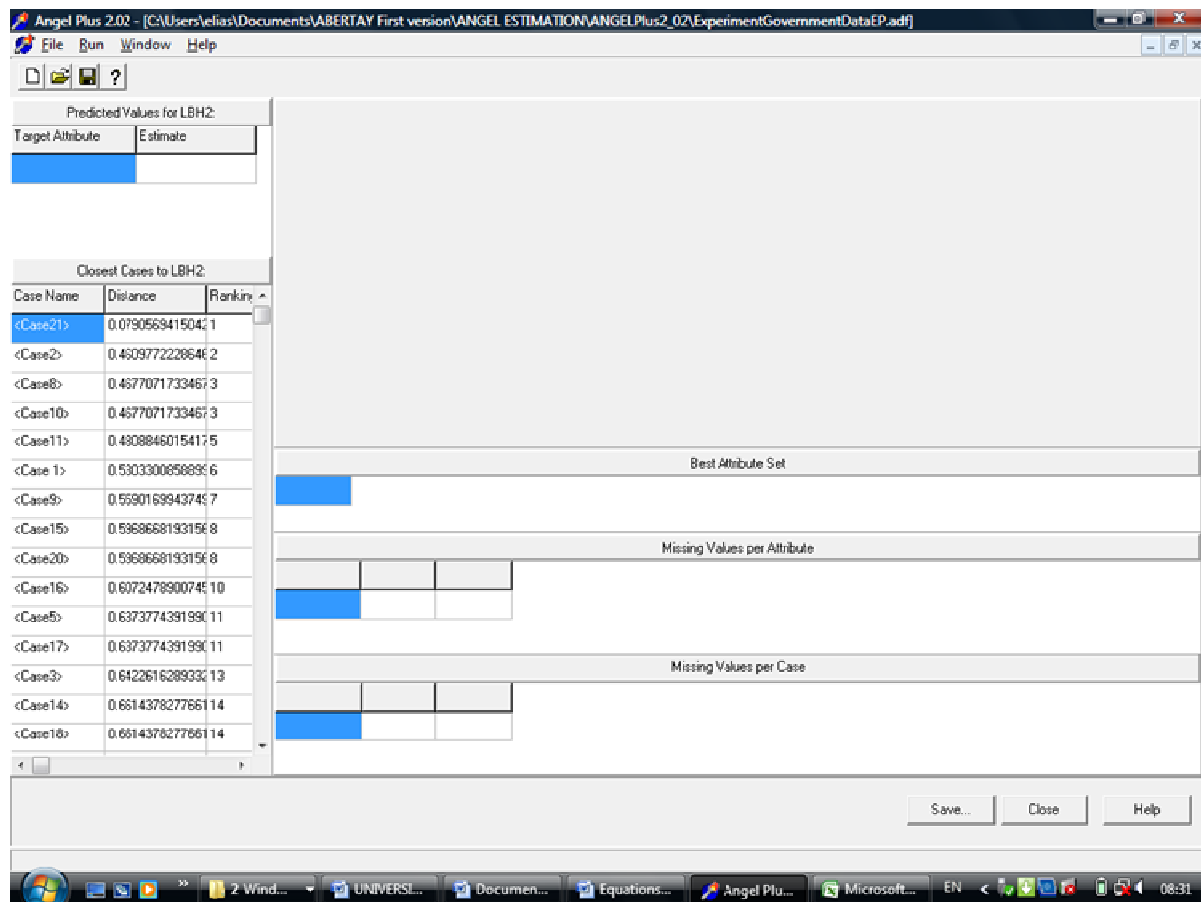## Estimating the 1st LB of Havering Project



Figure C11 – Using the ANGEL tool to identify a matching case (project analogue) for the LB of Havering project 1. Own dataset, independently collected.

Figure C12. Details of the identified project analogue for the 1st LB of Havering project; own dataset independently collected.

## Calculating the estimated duration of the 1st LB of Havering project

The non matching attribute for the 1st LB of Havering project and Case 1 of the "independent" councils data model is that of number of implementation teams involved. The corresponding equation for the risk rating variation is, $y = 0.0843x + 0.0752$ (from table 6).

Applying the equation to the project yields $y = 0.0843*4+ 0.0752= 0.4124$

Applying the equation to case 1 yields $y = 0.0843*3 + 0.0752=0.3281$

The analogy is thus defined as the ratio $(0.4124) / (0.3281) = 1.256934$

The estimated duration is given by multiplying the analogy by the duration of case 1, i.e.

$1.256934*168 = 211$ days (approximately)

# Estimating the 2<sup>nd</sup> LB of Havering Project



Figure C13 – Using the ANGEL tool to identify a matching case (project analogue) for the LB of Havering project 2. Own dataset, independently collected.

Figure C14. Details of the identified project analogue for the 2nd LB of Havering project; own dataset independently collected.

## Calculating the estimated duration of the 2nd LB of Havering project

The non matching attribute for the 2nd LB of Havering project and Case 21 of the "independent" councils data model is that of number of types of systems integrated. The corresponding equation for the risk rating variation is, $y = 0.0856x + 0.2455$ (from table 6).

Applying the equation to the project yields $y = 0.0856*5 + 0.2455 = 0.6735$

Applying the equation to case 21 yields $y = 0.0856*4 + 0.2455 = 0.5879$

The analogy is thus defined as the ratio $(0.6735) / (0.5879) = 1.145603$

The estimated duration is given by multiplying the analogy by the duration of case 21, i.e.

$1.145603*390 = 447$ days(approximately)

# APPENDIX D

## An Example of Applying FPA Calculations

# Function Point Calculation

## The Main Four Steps

1 - Get count-total, i.e. the number of features times complexity

2 - Get $\sum Fi$ – The total of the rating of 14 factors (0-5)

3 - FP = count-total X [0.65 + 0.01 X $\sum Fi$ ]

4 - Multiply historical averages per FP by this FP

## Step 1 – Get count total.

**Complexity Weighting**

| | simple | average | complex | product |
|---|---|---|---|---|
| Number of user inputs | __ × 3 + | __ × 4 + | __ x 6 | = ___ |
| Number of user outputs | __ × 4 + | __ × 5 + | __ x 7 | = ___ |
| Number of user inquiries | __ × 3 + | __ × 4 + | __ x 6 | = ___ |
| Number of files | __ × 7 + | __ × 10+ | __ x 15 | = ___ |
| Number of external interfaces | __ × 5 + | __ × 7 + | __ x 10 | = ___ |

An example of a Bank accounts record system – involving,

| | |
|---|---|
| 36 user inputs | simple complexity |
| 5 user outputs | average complexity |
| 20 user inquiries | simple complexity |
| 40 files accessed | simple complexity |
| 3 external interfaces | average complexity |

**Applying the Complexity Weighting formulae to the above example**

| | Simple; Average; Complex | product |
|---|---|---|
| 36 user inputs | 36 x 3 + __ x 4 + __ x 6 | = 108 |
| 5 user outputs | __ x 4 + 5 x 5 + __ x 7 | = 25 |
| 20 user inquiries | 20 x 3 + __ x 4 + __ x 6 | = 60 |
| 40 files | 40 x 7 + __ x 10+ __ x 15 | = 280 |

3 external interfaces   __ x 5 +  3  x 7 + __ x 10     =  21

<div align="center">TOTAL          494</div>

## Step 2 – Get $\sum F_i$

Fourteen factors (listed below) are rated between 0 and 5 depending on their significance to the specific project.

0 - Negligible

1- Incidental

2 - Moderate

3 - Average

4 - Significant

5 - Essential

For the Bank accounts example,

| | | | |
|---|---|---|---|
| F1 | require reliable backup & recovery? | Significant | 4 |
| F2 | data communications required? | Moderate | 2 |
| F3 | distributed processing functions? | Significant | 4 |
| F4 | performance critical? | Average | 3 |
| F5 | run on existing, heavily utilized environment? | Essential | 5 |
| F6 | require on-line data entry? | Essential | 5 |
| F7 | on-line data entry from multiple operations? | Incidental | 1 |
| F8 | master files updated on-line? | No influence | 0 |
| F9 | inputs, outputs, files, or inquiries complex? | Incidental | 1 |
| F10 | internal processing complex? | Incidental | 1 |
| F11 | code designed to be reusable? | Average | 3 |
| F12 | conversion and installation included in the design? | Average | 3 |
| F13 | system designed for multiple installations in different orgs? | No influence | 0 |
| F14 | application designed to facilitate change and ease of use? | No influence | 0 |

<div align="right">$\sum F_i = 32$</div>

## Step 3 – Calculate FP

FP = count-total x[0.65 + 0.01 x$\sum$ Fi ]

= 494 x [0.65 + 0.01 x 32] = 479.18 (Applied to the Bank Accounts example.

## Step 4 – Estimate Effort by Multiplying with Historical Averages

Estimated FP   479.18

Proportional to average 479.18/623 = 0.77

Estimated effort = 33 months x 0.77 =25.4 months

**APPENDIX E**

**Previously Published Work**

List of Papers attached on CD

Pimenidis, E. and Miller, C.J. (2006) 'Evaluating Risk in E-Business Development Projects': *the IADIS International Conference E-Commerce 2006.* Barcelona, Spain, 9-11 December 2006, pp. 81-88.

Pimenidis, E. and Miller, C.J. (2005) 'Risk and E-Business Development Projects: in Search of a New Metric': *the IADIS International Conference E-Commerce 2005.* Porto, Portugal, 15-17 December 2005, pp. 303-311.

Pimenidis E., Miller C. J. and MacEachen, C.F. (2004) 'Identifying Risk for e-Business Development': *the European Conference on Information Technology Evaluation (ECITE2004)*, Amsterdam, The Netherlands, 11-12 November 2004pp. 349-358.

Pimenidis E., Miller C. J. and MacEachen C.F. (2002) 'A New Look at Project Estimation for E-Business': *the International Conference on Project Management "Breakthrough with Project Management – In the Era of Global Revolution by IT" (ProMAC2002),* Singapore, 31 July – 2 August 2002, pp. 353-358