CAUSAL NETWORK METHODS FOR INTEGRATED PROJECT PORTFOLIO RISK ANALYSIS

A Dissertation

by

PAUL BRADFORD GOVAN

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Ivan D. Damnjanovic
Co-Chair of Committee,	Kenneth F. Reinschmidt
Committee Members,	Stuart Anderson
	Zenón Medina-Cetina
	Daren B. H. Cline
Head of Department,	Robin Autenrieth

August 2014

Major Subject: Civil Engineering

ABSTRACT

Corporate portfolio risk analysis is of primary concern for many organizations, as the success of strategic objectives greatly depends on an accurate risk assessment. Current risk analysis methods typically involve statistical models of risk with varying levels of complexity. Though, as risk events are often rare, sufficient data is often not available for statistical models. Other methods are the so-called expert models, which involve subjective estimates of risk based on experience and intuition. However, experience and intuition are often insufficient for expert models as well. Furthermore, neither of these approaches reflects the general information available on projects, both expert opinions and the observed data.

The goal of this dissertation is to develop a general corporate portfolio risk analysis methodology that identifies theoretical causal relationships and integrates expert opinions with the observed data. The proposed conceptual framework takes a resource-based view, where risk is identified and measured in terms of the uncertainty associated with project resources. The methodological framework utilizes causal networks to model risk and the associated consequences.

This research contributes to the field of risk analysis in two primary ways. First, this research introduces a new general theory of corporate portfolio risk analysis. This theoretical framework supports risk-based decision making whether through a formal analysis or heuristic measures. Second, this research applies the causal network methodology to the problem of project risk analysis. This methodological framework

provides the ability to model risk events throughout the project life-cycle. Furthermore, this framework identifies risk-based dependencies given varying levels of information, and promotes organizational learning by identifying which project information is more or less valuable to the organization.

DEDICATION

To my family

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my committee chair, Dr. Damnjanovic, for his immense inspiration, enthusiasm, and guidance. I consider it an honor and a privilege to have been his student.

I would also like to extend my appreciation to my committee co-chair, Dr. Reinschmidt, for his continuous support. I will always remember the valuable experiences he shared with me. Thanks also go to my committee members, Dr. Anderson, Dr. Medina-Cetina, and Dr. Cline, for their valuable comments and suggestions over the course of my study.

Finally, thanks to my mother and father for their encouragement and to my wife for her support and love.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER I

INTRODUCTION

The concept of risk plays an important part in everyone's life. Risk is involved when playing a game of chance, choosing a career, or investing in a new opportunity. Without an understanding of risk, people cannot properly assess decisions such as these.

Firms are certainly concerned with risk. If when a firm considers a new venture, management seeks answers to questions such as: what is the risk associated with the venture? More specifically, what is the risk to the firm associated with entering a new market or developing a new technology? If the firm manages a corporate portfolio of projects, then the firm is concerned with how risk affects projects.

Risk generally carries a negative connotation. The main reason for this view is that most people are concerned with the downside of their decisions and generally prefer less risk than more. In reality, an upside to risk often exists, as risk may imply an opportunity for attaining a certain objective. For example, skydivers risk their lives for the thrill of jumping out of airplanes, skiers risk injury to enjoy a ride through the slopes, and, of course, firms/projects risk initial investments in an effort to meet shareholder expectations (i.e. returns). Fundamentally, making decisions implies uncertain outcomes, some favorable and some not.

Risk analysis is concerned with identifying these sources of uncertainty and measuring them in a quantifiable way. These statements lead to a general definition of a risk event: an event or set of events that, if they occur, have one or more impacts, either positive or negative, on at least one strategic objective (Project Management Institute, 2010).

These ideas, of course, lead to the question of how to manage risk. How do firms make decisions under uncertainty? What options or alternatives do they have, and how do they choose alternatives? These questions are the subject of the expanding field of risk management.

Risk management is concerned with planning controlled responses in the likely case that risk events occur. While risk analysis is typically an objective assessment of the state or nature of risk, the response to risk is subjective. Decision making also depends on risk preference or how more or less risk an individual prefers.

1.1 Background and Motivation

Risk analysis is usually concerned with assigning a probability to risk. Hence probability theory and statistics are the framework for many risk analysis methods. Current methods usually involve statistical models of risk, with varying levels of complexity based on the specific problem. These models depend greatly on data; however, data collection is no small task for many problems, especially if certain risk events are rare.

In absence of data, risk-related decisions should instead be based on experience, judgment, and intuition. After all, risk models are only as good as the assumptions that go into them. However, gaining experience and learning about risk may be time consuming and is often very expensive. In fact, one of the main reasons why statistical and mathematical models were developed in the first place was because experience and judgment alone were deemed inadequate. Judgment is subjective, which is not incorrect by virtue of being subjective, but the possibility exists (Fenton & Neil, 2011). The purpose of risk models, therefore, should be to inform and to be consistent with experience and judgment. In other words, the model should reflect the information available, both expert judgment and data.

Another common problem with modeling risk is accounting for the fact that risks are interdependent. Current risk analysis methods often assume that risks are independent. These methods are not structural models, which attempt to explain why, in principle, risks are interrelated. However, risk interactions are important to identify and measures because they impact the coordination of project resources for project risk management.

1.2 Research Goal and Objectives

The goal of this dissertation is to develop a general risk analysis methodology that identifies causal relationships among events and conditions and integrates expert opinions with the observed data. This dissertation is intended to integrate the experiential and the analytical approaches to risk analysis.

The proposed theoretical framework is based on the Resource-based View (RBV) of the firm (Wernerfelt, 1984). The RBV theory proposes that value exists in viewing the firm in terms of its resources, and, in fact, a source of value to the firm is its resources. This dissertation shows that the RBV theory leads to a new theoretical framework of risk analysis, which leads to significant consequences for strategic project management.

The proposed methodological framework utilizes causal networks to model risk and the consequences of risk. Networks, in general, are valuable both as a graphical display of the problem area and as a way to model interdependencies between risks. Causal networks also have many practical applications that allow firms to assess different management strategies. This methodology also opens the problem of risk analysis to a multitude of popular statistical and network methods, many of which are explored here.

The specific objectives related to this overall research goal are as follows:

- Formulate a ***general theoretical framework of risk analysis*** based on the RBV theory. The theoretical framework should be logically valid, and, furthermore, the model must be consistent and complete.
- Develop a ***causal network model*** that can be used by firms to help assess project related risk. The developed model should be able to account for many different types of risk factors, data sources, and management strategies.
- Explore the use of ***network measures*** to support model development and analysis. Network measures are applicable in a number of fields. While some existing measures are applicable to the problem of risk, still others are introduced for the first time.
- Develop a Bayesian network model for ***corporate portfolio risk analysis***. Firms concerned with managing portfolios of projects and require an understanding of how different projects interrelate. The corporate portfolio

model should also follow from the proposed theoretical and methodological framework.

1.3 Research Contribution

In a broad sense, this research contributes to the field of project risk analysis in two major ways. First, this research introduces a new general theory of risk analysis. Firms may benefit from the risk-resource-based view whether or not they proceed with a formal analysis. This theoretical framework guides decision making and overall risk management. If firms then want to move into a more formal analysis, they have the theory and methods to do so.

The second major way this research contributes is the application of causal networks to the problem of risk. Traditionally, risk analysis methods are problem specific. Firms may benefit from a comprehensive methodological framework with the capability of modeling risks throughout the project life-cycle, from conception to operation. The methodology is also an integrative approach to risk analysis that is a fundamentally simple yet widely applicable model. Further, the model helps identify risk-related dependencies given varying levels of information, and promotes organizational learning by identifying what data to collect in order to create value for the organization. Some of the specific contributions related to these two major areas are as follows:

- Development of a comprehensive methodological framework for modeling life-cycle project risk

- Guidelines for model development that are based on existing popular project management tools
- Advancement of network measures that support both the process of model development and model analysis.
- Development of a corporate portfolio risk analysis methodology. Traditionally, research effort has focused on project risk analysis, ignoring dependency structures that occur on the portfolio level (e.g. shared resources).
- Design of project and firm options, alternatives, and recourse strategies using the network approach. Alternatives and strategies are certainly numerous, but this research shows how they can be adapted for specific problems.

1.4 Dissertation Outline

This dissertation is organized into eight chapters. Chapter I introduced the motivation, objectives and contributions of this dissertation. Chapter II is a review of the literature related to this research and is subdivided into five related topics: general project risk management, the Resource-based View (RBV) theory, network/graph theory, Bayesian and causal networks with an emphasis on Bayesian methods and networks, and corporate portfolio risk management.

Chapter III covers the proposed methodological framework for risk analysis. This chapter presents an introduction to the proposed risk-resource-based view, the formal theoretical model, and the research approach and methodology used for the remainder of the dissertation.

Chapter IV concerns the proposed methodology of causal networks for project risk analysis. The research methodology includes developing the model structure, validating the model, and showing how the model can be used to assess many problems related to risk.

Chapter V explores the use of network measures for the problem of project risk analysis. Network measures are useful both as tools for model development and for risk analysis; this chapter gives an overview of existing measures and explores the use of new ones.

Chapter VI is an extension of the project risk analysis model to the problem of corporate portfolio risk. The corporate portfolio model is consistent with both the proposed methodological framework and the principles of portfolio theory.

Chapter VII presents a case study of a compressor station project in order to illustrate the methodology presented in this dissertation. The case study demonstrates the process of specifying a model and the model parameters, validating the model, and using it for risk-related decision making.

Finally, chapter VIII is a summary of the major findings of this dissertation, the limitations of the work performed so far, and directions for future work. This research is the first step in a rich and exciting field of study.

CHAPTER II

LITERATURE REVIEW

This chapter presents an overview of the literature related to this research and is subdivided into four major areas: project risk management, the Resource-based View (RBV) theory, probabilistic networks, and corporate portfolio risk management. The first section provides an introduction to project risk management and existing risk analysis methods and procedures. This introduction identifies the main shortcomings of existing methods and how this research contributes to the field of project risk management. The second section discusses the RBV theory, related extensions of the theory, and common criticisms of the RBV. The RBV serves as the background for the conceptual framework of this research. The third section gives an overview of probabilistic networks, with an emphasis on Bayesian methods and Bayesian networks, and related network approaches to risk analysis. Finally, the fourth section discusses research related to corporate portfolio risk management. This overview identifies several areas where existing research is inadequate and how this research contributes to corporate portfolio risk analysis.

2.1 Project Risk Management

Risk is the effect of uncertainty, either positive or negative, on stakeholders' objectives. Risk management is the process of identifying, assessing, and managing the causes of uncertainty (ISO, 2009). In the context of projects, an example of risk management may be structuring a contract to transfer risk to another party or using an engineering design with which all parties are well familiar. In practice, management of the overall risk may be challenging. Some risk may go unnoticed, and some may be difficult to reduce. For this reason, risk management continues to be an important topic of research and practice.

Firms have been practicing risk management for over 50 years, and evidence of risk management principles may be traced to some of the earliest known projects (Covello & Mumpower, 1985). However, formal risk management methods are a relatively recent development. Many of these tools are still used today.

The U.S. Department of Defense (DoD) developed, among other things, three of the most popular risk management tools: the work breakdown structure (WBS), the Program Evaluation and Review Technique (PERT), and the risk register. The DoD created the WBS as part of the Polaris mobile submarine launched ballistic missile project (Department of Defense, 1968). The WBS is a hierarchical tree-like structure of the work required to complete a project. The WBS is built by dividing the deliverables from the most general to the most specific. This structure is often used as a basis for management to budget, schedule, and assign resources to the project. Figure 1 shows a simple example of a WBS. Level 1 indicates the overall project, in this case, an aircraft system. Level 2 breaks the project into subtasks, and level 3 breaks the subtasks into packages of work.

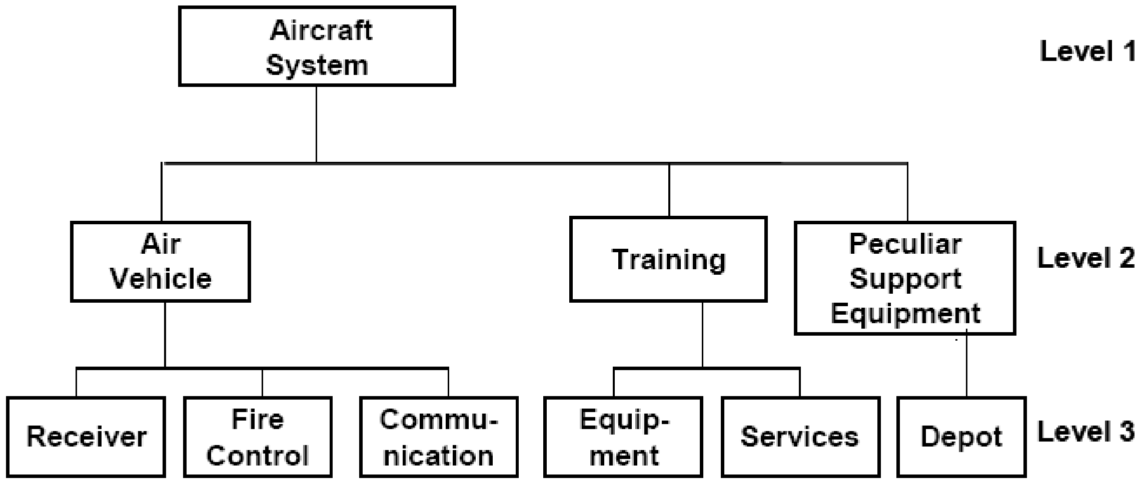


Figure 1: Example WBS (Adapted from Department of Defense, 1968)

The DoD also developed the Program Evaluation and Review Technique (PERT) for project cost and schedule analysis as part of the Polaris project (Department of Defense and National Aeronautical Space Administration, 1962). PERT is a method for analyzing the uncertainty in tasks involved in completing a project. Given the uncertainty in project tasks and the relationships between them, the PERT method provides an estimate of the uncertainty of the total project network. Figure 2 illustrates the general PERT method. The packages of work shown in Figure 1 are each associated with some probability distribution. With knowledge of the interrelationships between the tasks, PERT provides an estimate of the probability distribution of the overall project network.

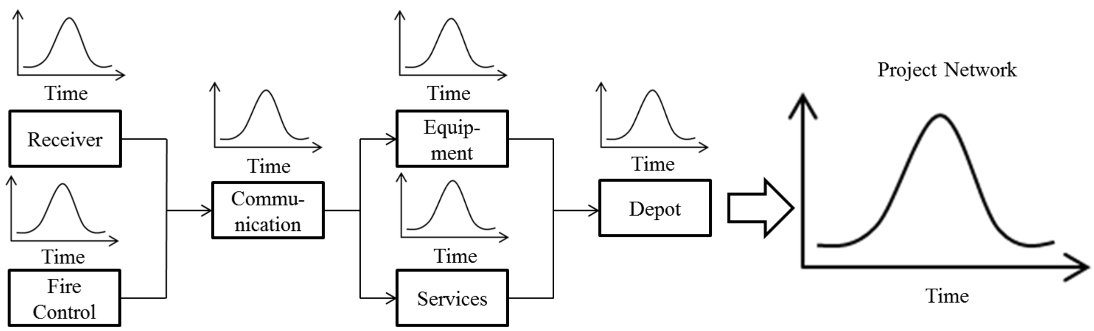


Figure 2: Illustration of PERT Method

The risk register is a tool to help qualify risks and rank them in order of greatest priority (Department of Defense, 2006). A risk register usually includes an ordinal risk matrix, which "bins" risks into priority categories. The risk matrix has two scales to represent the two aspects or dimensions of risk, the probability and the impact/consequence. The higher the probability, the more likely the occurrence of risk, and the higher the impact, the greater the consequence on strategic objectives (Garvey, 2009). The risk register also includes other relevant information such as the risk title and description, the root causes or early warning signs, the specific tasks impacted, the occurrence probability, the severity or impact on the overall project, and mitigation and management strategies. An example risk matrix is shown in Figure 3. The figure indicates the consequence (x-axis) and likelihood (y-axis). The darkest cells indicate the category of highest priority, or the category for which management should be most concerned.

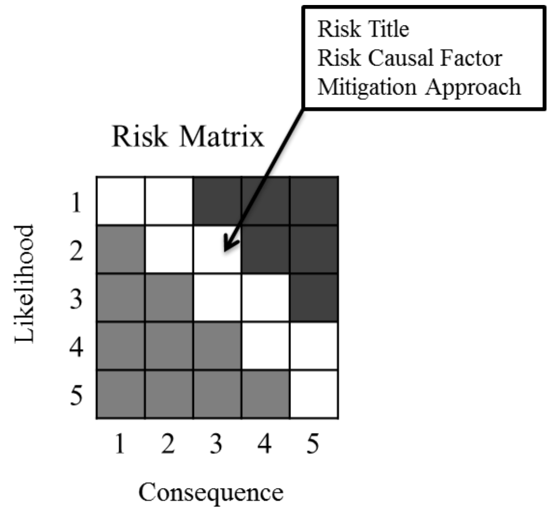


Figure 3: Example Risk Matrix (Adapted from Department of Defense 2006)

One of the primary applications of the risk register is to develop the so-called risk statement. A risk statement follows the Condition-If-Then construct. The condition represents the early warning sign or root cause of the identified risk event(s), while risk events are probabilistic events that may occur because the condition is present, and the consequence(s) is the impact of the risk event occurring (Garvey, 2009). Take, for example, the illustration shown in Figure 4. Suppose the condition is that high traffic volume is present in and around the project site. A risk event might be that access to the site is inadequate, and the consequences of the risk event include delays, which could cause an increase in the required resources, namely construction management and labor.

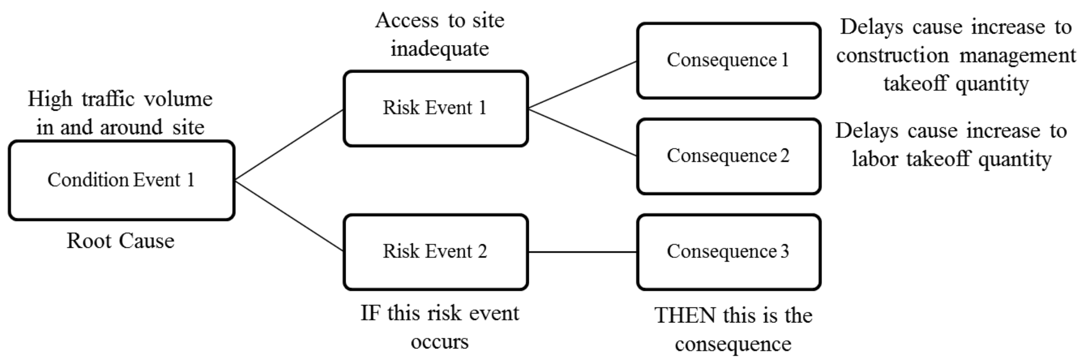


Figure 4: The Condition-If-Then Risk Statement (Garvey, 2009)

Since its formal development in the 1960's, numerous contributions have been made to the field of project risk assessment. Several authors provided guidance on estimating statistics for project risk analysis based on judgment (Moder & Rodgers, 1968), (Davidson & Cooper, 1976), (Perry & Greig, 1975), (Keefer & Bodily, 1983). Hillier (1963) showed how to derive the probability distributions of present worth, annual cost, and internal rate of return (IRR). Fault tree analysis (FTA) for deductive failure analysis was developed by Bell Laboratories in 1962 (Bedford & Cooke, 2003). The six sigma process for quality control was introduced by Motorola in 1986 and later popularized by General Electric in 1995 (Eckes, 2001). However, much less research has focused on integrating subjective estimates of risk based on judgment with statistical methods for risk analysis (Fenton & Neil, 2011). This dissertation contributes to the field by introducing a general methodology that integrates estimates of risk based on judgment with observational data.

In addition, many contributions have been made to project risk management and project risk-related decision making. Multiple authors have studied managerial perspectives or attitudes on risk (March & Shapira, 1987), (Kim & Reinschmidt, 2011). Several authors explored the use of real options for project management under risk (Trigeorgis, 1993), (Huchzermeier & Loch, 2001).

Many government agencies publish risk management guidelines. The DoD publishes the Risk Management Guide for DoD Acquisition (2006), and the National Aeronautics and Space Administration (NASA) publishes the Risk Management Handbook (2011). Several nonprofit project management organizations also publish risk management guidelines. In 1987, the Project Management Institute (PMI) first published the Project Management Body of Knowledge (PMBOK), which attempted to document and standardize, among other project related methods, risk management information and practices. The National Institute of Standards and Technology (NIST) publishes the Guide for Applying the Risk Management Framework to Information Systems (2010) and the Guide for Conducting Risk Assessments (2012).

Existing methods and guidelines attempt to identify possible risk events. However, no list of possible risk events could possibly be comprehensive, as potential risk events are numerous. However, firms are usually concerned with financial risk events. Financial risk events have the consequence of an impact on financial cost, where financial cost is expressed in monetary terms.

Project finance-related risk events may be divided into three main categories: commercial, macro-economic, and political risk events (Yescombe, 2002). Commercial risk events are specific to the project or the market in which the firm operates. Commercial risk events include, but are not limited to, completion, environmental, operating, revenue, supply, force majeure, contract, and sponsor risk events. Macro-economic risk events have to do with external economic conditions such as inflation, interest rates, and currency exchange rates. These risk events are not directly related to the project but have a broad impact on the economic environment in which it operates. Political risk events relate to government decisions or actions such as war or civil disturbance. Political risk events are especially prevalent when projects involve international cross-border financing. Political support for a project is often an ex-ante necessity. Governments may, moreover, change throughout the course of the project, thus affecting projects in a number of ways. Even during operation, political agreements may change or reverse. Once the project is complete, usually the facility cannot be taken out of the country, and sometimes the country uses this as leverage to change political agreements.

Engineers are concerned with technical risk assessment. Technical risk events are related to the performance requirements of engineering systems. (Garvey, 2009). Engineering systems and subsystems may have many technical risk events, and engineers have many different ways to assess technical risk events. As such, engineers do not have a way to integrate these independent risk measures into an overall risk assessment of the system. Performance measures of engineering systems include dependability, reliability (i.e. remaining time to failure, mean time to failure (MTTF), failure rate, etc.), and maintenance requirements (e.g. degradation) (Weber, Medina-

Oliva, Simon, & Iung, 2012). For example, structural engineers are concerned with the failure rate of welds, transportation engineers are interested in the degradation of roadways, and geotechnical engineers are concerned with the settlement rate of foundations.

Project risk management may be divided into three primary steps: risk identification, risk analysis, quantification and assessment, and risk mitigation and management (Nicholas & Steyn, 2008). The first step in risk management is to identify possible risks. This step should typically be performed as early in the project as possible. In general, risk management at the front end of the project is more beneficial than at the back end. Often the scope of work (SOW) and project management tools such as the WBS, RBS, and so on are used to help identify possible sources of risk.

The next step is to analyze and quantify risk. Due to the inherent uncertainty related to risk, most risk analysis methods are based on the theory of probability. If risk is expressed as a probability, then the problem is open to the broad field of probabilistic methods (Fenton & Neil, 2011). Classical probabilistic methods, however, have certain difficulties in practice. Classical probabilities are based on relative frequencies derived from historical data. When data is not available, which is often the case with projects, these methods break down. In the case when data is not available, probabilities must be subjective and are usually set by an expert or someone familiar with the project. Many analysts, however, have difficulty expressing uncertainty as a probability. Conversely, analysts who do have experience with probability theory may not have relevant experience with projects.

The final step is risk mitigation and management. In some cases, risk can be eliminated or reduced. In many cases, however, risk management requires unique strategies. Risk management strategies typically include four options: avoidance, reduction, sharing, and retention. (1) Risk avoidance is the elimination of risk. (2) Risk reduction is the optimization or mitigation of risk. (3) Risk sharing refers to sharing or transferring risk to another party. (4) Risk retention is the acceptance of the loss or gain posed by risk. Figure 5 presents the general risk management process, where risk management is the overall process of risk identifications, analysis, evaluation, and mitigation.

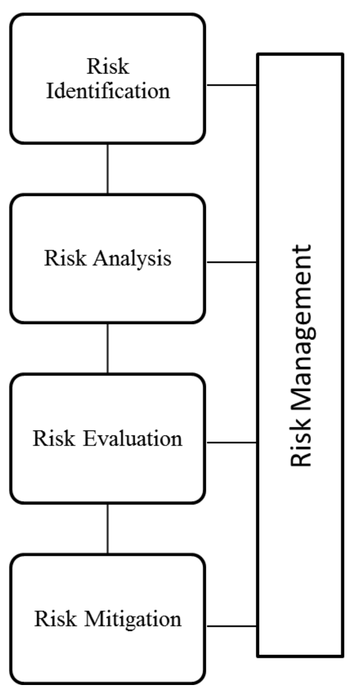


Figure 5: Risk Management Process

2.2 The Resource-based View

A fundamental question in strategic management research is how firms create value. A common answer to this question is the term, a competitive advantage, or, more specifically, a long-term or sustained competitive advantage (SCA). Barney (1991) defines a competitive advantage as a value creating strategy, one not currently or possibly employed by competitors.

Many firms consider the product-market as the source of SCA, and thus many analysis methods operate from the product-market view (Wernerfelt, 1984). See, for example, Porter's (1980) five competitive forces model, which attempts to describe the external factors of an attractive industry. The product-market view makes simplifying assumptions about firms. First, firms are assumed to be identical or homogenous in terms of strategic assets and current or possible strategies (Barney J. , 1991). Second, if resource heterogeneity does exist, it does not last long because resources are assumed to be highly mobile and can be bought or sold in product-markets.

The management theory so-called the Resource-based View (RBV) looks at the other side of the coin. The RBV proposes that value exists in a firm's resources, and, in fact, the source of SCA is the mix of firm-related resources. The idea of viewing the firm as a mix of resources may be traced back to earlier research in strategic management, most notably by Penrose (1959) and Rubin (1973). The term RBV was coined by Wernerfelt (1984) in his seminal paper, "A Resource-based View of the Firm".

Wernerfelt argued that value exists in viewing the firm in terms of its resources. By specifying the size of the firm's activity in different product-markets, the firm may infer the required resources. Conversely, by specifying the firm's available resources, the firm may find the optimal product-market activities. While the product-market and resource views should, in principle, lead to the same solutions, firms may find one perspective easier to apply than the other.

A resource is anything that could be a strength or weakness of the firm (Wernerfelt, 1984). Most people think of resources as tangible assets, where such examples include capital, labor, machinery, and natural resources such as land. However, in a broad sense, resources also include intangible assets. Intangible assets include the company brand, technology, education, or skill set. Amit and Shoemaker (1993) referred to intangible assets as capabilities. In this sense, tangible assets are tradable commodities and non-specific to the firm, whereas capabilities are specific to the firm or organization (Amit & Schoemaker, 1993), (Makadok, 2001).

Figure 6 illustrates the RBV and the link to SCA. The firm consists of a set of corporate resources, both tangible assets such as capital, labor, equipment, and materials as well as capabilities such as technology, skills, and company brand. To achieve SCA, firms must specify a certain mix of corporate resources.

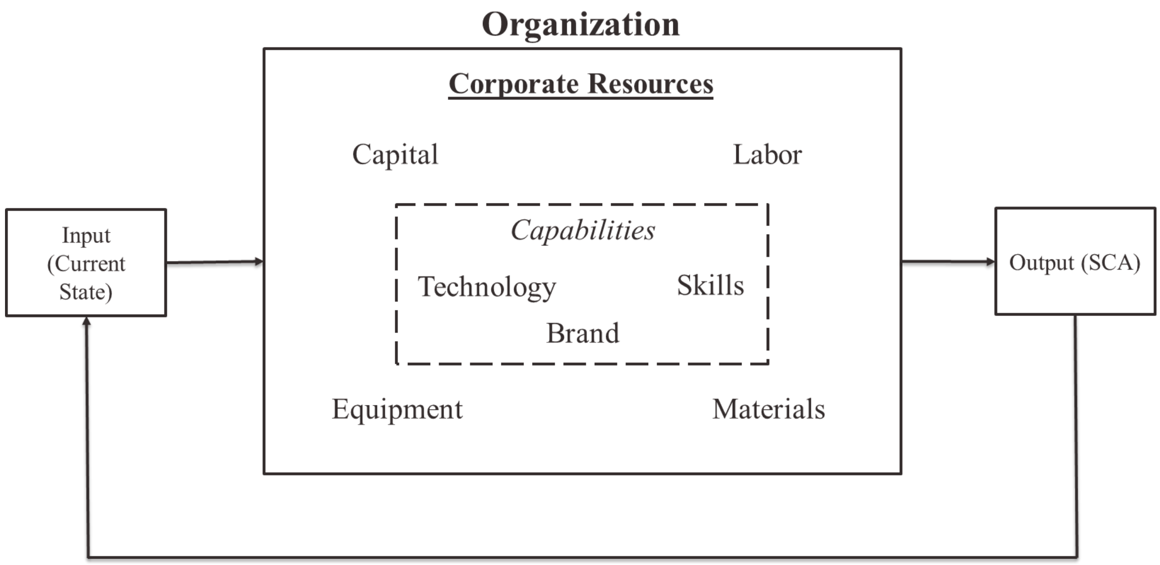


Figure 6: The Resource-based View of the Firm

However, not all resources are a source of SCA. If one person's skill set happens to be the same as another's skill set, all else being equal, then neither has a competitive advantage. Barney (1991) argued that resources must possess key characteristics in order to be a source of SCA. These characteristics are as follows:

- Valuable – The resource must allow the firm to implement a value-creating strategy
- Rare – The resource must not be owned by a large number of firms
- Inimitable – Other firms must not be able to obtain the resource
- Non-substitutable – There must not be any equivalent resources that are neither rare nor inimitable

These characteristics have the popular acronym VRIN. Barney claimed these characteristics are empirical factors of how useful resources are for generating SCA.

Barney and Hesterly (2010) went on to argue that the firm must organize resources in order to exploit value creation. This idea leads to resource-portfolio theory, where resources may be a source of firm diversification. At one time or another certain resources may be a source of value while others are not. Rather, it is important how the value of resources change relative to other resources in the portfolio. Thus, a diversified portfolio of resources lowers the risk of changing market forces. In the same light, Wernerfelt (1984) argues that firms should balance the use of existing resources and the development of new ones.

The RBV has, in fact, been applied to project management. Consider the firm as a set of resources. Logically any part of the firm is a subset of resources. Projects are part of the firm; then, projects are simply a subset of firm-related resources. Projects may lead to or contribute to a competitive advantage; therefore, firms should invest in practices and develop assets relevant to positioning projects strategically (Jugdev, 2004).

Project management often views the project from a resource perspective. Consider, for example, the resource breakdown structure (RBS) and the bottom-up cost estimate. The RBS is similar to the WBS, except that it is a hierarchical structure of the project resources (Project Management Institute, 2010). The RBS has applications in many project management techniques, such as leveling resources, scheduling, and identifying needs for planning and control. The RBS is also often used to help identify risks by assessing resources at different levels of the breakdown. At the bottom level, the RBS is often used to estimate project cost. This type of cost estimate is referred to as a bottom-up cost estimate, which is a fairly detailed estimate of the project work.

Bottom-up cost estimating is a process of breaking each task into smaller components, estimating the resource costs required to complete each task, and summing up the estimates to get the total project cost estimate (Project Management Institute, 2010). The project cost estimate is, therefore, the sum of the products of two numbers, the resource quantities required and the corresponding unit costs. For a vector of resource unit costs $\boldsymbol{u}$ and a vector of resource quantities $\boldsymbol{q}$ both of length n, the project cost p is the following:

$$p = \sum_{i=1}^{n} u_i q_i \tag{1}$$

From the RBV, the project is defined as a set of resources. Therefore, a source of value to the firm is the project resources or, more specifically, the certain mix of project-related resources. By specifying the project's available resources, the firm may, in theory, find the optimal project activities.

The RBV theory influenced many additional extensions and related research. Rumelt (1984) studied market imperfections, the heterogeneity of firms, and limited transferability of resources. Dierickx and Cool (1989) linked SCA to the lack of substitutability and imitability. Conner (1991) compared the RBV to other theories of the firm in the context of industrial organization economics. Mahoney and Pandian (1992) considered the RBV and its implications from different research perspectives. Amit and Schoemaker (1993) connected industry related factors at the market level to strategic assets at the firm level. Peteraf (1993) claimed that SCA requires heterogeneous resources within an industry, ex-post limits to competition, imperfect resource mobility, and ex-ante limits to competition. Conner and Prahalad (1996) extended the RBV to a knowledge-based theory of the firm. Makadok (2001) integrated the resource-based and dynamic capability views. Since its initial development, many articles have been written about the RBV and the theory continues to be a topic of current research (Kraaijenbrink, Spender, & Groen, 2009). The RBV theory has also been extended to many other fields, such as information systems (Wade & Hulland, 2004) and organizational networks (Lavie, 2006).

The RBV theory has also been criticized. Priem and Butler's (2001a), (2001b) critiques of the RBV and Barney's (2001) rebuttal are commonly cited. Probably most notably, Priem and Butler (2001) argued that the RBV is tautological. The idea that a resource is a source of value creation because it is valuable is a tautology. Barney (2001) argued that, in fact, any theory could be rephrased as a tautology. Furthermore, Priem and Butler (2001) argued that while the VRIN characteristics are necessary individually, they are not sufficient conditions for SCA. Barney & Hesterly (2010) later revised his characteristics and replaced the acronym N with O, where O stands for the organization of the resources.

Kraaijenbrink et al. (2009) argue that criticism has contributed to the theory's development. They review the development of the RBV theory and related criticism. A common criticism is that the RBV has no managerial implications. They argue that not all theories have managerial implications, and the extensive references to RBV are evidence of its impact. Another common criticism is that VRIN/O is neither necessary nor sufficient for SCA. They argue that the RBV does not sufficiently consider how judgment and mental models affect SCA. Kraaijenbrink et al. conclude that the RBV can withstand most criticism but that more theory and research is still required. The authors suggest that the RBV theory may be improved by moving into an inherently dynamic framework, “by incorporating time, space, and uncertainty resolution into the RBV’s axiomatic base.”

This dissertation contributes to the RBV in three primary ways. First, this research shows that the RBV has managerial implications in the context of project risk management. Second, the research illustrates how judgment and mental models impact risk-based decision making and, in turn, organizational resources. Third, the research introduces the concept of uncertainty related to risks and the corresponding impact on organizational resources.

2.3 Network Theory

This section provides a brief introduction to network theory and related applications to project risk management. For the purposes of this dissertation, the terms “network” and “graph” are used interchangeably. Leonard Euler is generally credited with introducing graph theory in a 1736 monograph titled “Seven Bridges of Konigsberg.” Euler developed a formula relating the number of edges, vertices, and faces of a convex polyhedron (Biggs, Lloyd, & Wilson, 1968). James Joseph Sylvester is generally credited with first introducing the term “graph” (Sylvester, 1878).

A graph is a mathematical model of systems that involve a binary relation, where a set of vertices (nodes) are interconnected by edges (arcs) (Ibe, 2011). For a graph G, the set of vertices V, and the set of edges E, $G = (V, E)$. Two vertices x and y are *adjacent* if they have an edge between them. The adjacency matrix A of a finite graph G on n vertices is the $n*n$ matrix where the non-diagonal entry a_{ij} is the value 1 if vertex i is adjacent to vertex j and 0 otherwise. The *degree* of a vertex x is the number of edges that are incident with it. The degree of vertex x given adjacency matrix A is the following:

$$d(x) = \sum_{y} A(x, y) \tag{2}$$

A *path* is a set of distinct connected nodes in a graph. A graph is *connected* if every pair of nodes is joined by a path, and a *bridge* is an edge whose removal would disconnect the original graph into separate subgraphs. Finally, a *directed* graph consists of edges with ordered vertex pairs or a direction from one vertex to another. Directed graphs are often used to model relationships between vertices. The direction is usually indicated by an arrow in the direction from one vertex to another. The preceding definitions may, in addition, be extended to take into account the directionality of the edges.

Network theory has been applied to a number of fields, including engineering and risk management applications. Complex projects are often characterized as a network of components that share interfaces to function as a whole (Sosa, Eppinger, & Rowles, 2007). Sosa et al. (2003) use a network approach to identify modular and integrative systems in engineering designs. They claim that complex products may be decomposed into systems of modular or integrative components, where modular components are defined as having design interfaces with other systems that are clustered among a few physically adjacent systems, and integrative systems have design interfaces that span over most or all of the systems. They use the matrix-based network tool called the design structure matrix (DSM) to sequence dependencies between product components (Steward, 1981). Sosa et al. (2007) define modularity as inversely related to the centrality measures of degree, distance, and path. Thus, they use a network approach to measure the level of component disconnectedness with other product components.

Network measures have also been applied to project risk management. Fang et al. (2012) develop a project risk network to model project risks and interdependencies. They use the DSM to capture risk interdependencies, where the DSM is the basis for a project network. Marle et al. (2013) use clustering methods to group risks, such that project interactions are maximum within clusters and minimal outside of them. They identify project interactions through project resources (actors), which require coordination for project risk management. However, the methods of Fang et al. and Marle et al. are not structural models and thus do not explain why, in theory, risks are interrelated. This dissertation extends applications of network theory to engineering and management by proposing a general theoretical model of project risks and their interactions.

2.4 Bayesian and Causal Networks

This section provides an overview of Bayesian and causal network methods. The terms Bayesian, probabilistic, and belief networks are used interchangeably to describe the branch of network methods where nodes represent random variables and arcs represent conditional dependencies between variables. The term causal networks is used to describe the larger branch of network methods where nodes may or may not represent random variables, and arcs represent causation between nodes. Thus, Bayesian networks are a special case of causal networks. These ideas are made clearer in the sections that follow. The following subsections discuss Bayesian methods, Bayesian networks, and causal networks, with an emphasis on applications for risk analysis.

2.4.1 Bayesian Methods

Bayesian methods are a way of doing statistical inference using Bayes' Theorem (Hoff, 2009), (Gelman, Carlin, Stern, & Rubin, 2004). The methods involve updating prior beliefs based on information contained in data. The term Bayesian is named for Thomas Bayes, a British mathematician and Presbyterian minister, who introduced his theorem while studying how to compute the probability distribution for the parameter of the binomial distribution. Bayes' ideas were later published and extended by Laplace (Laplace, 1986). Bayes' Theorem is a consequence of the Law of Conditional Probability, where for any events *A* and *B*, Bayes' Theorem states the following:

$$P(B|A) = \frac{P(B,A)}{P(A)} = \frac{P(A|B)P(B)}{P(A)} \tag{3}$$

Bayes' Theorem notably applies to statistical models. Let $\boldsymbol{Y}$ be an observed data vector, a realization of which is $\boldsymbol{y}$. $\boldsymbol{Y}$ has a probability distribution that depends on an unknown vector of parameters $\boldsymbol{\theta}$. The probability distribution of $\boldsymbol{Y}$ given $\boldsymbol{\theta}$, or more commonly the likelihood of $\boldsymbol{Y}$ given $\boldsymbol{\theta}$, is $p(\boldsymbol{y}|\boldsymbol{\theta})$, where the likelihood of $\boldsymbol{Y}$ is conditional on $\boldsymbol{\theta}$ being the true parameter value. The prior distribution of $\boldsymbol{\theta}$ is a probability distribution that represents the a priori opinion or belief about the unknown parameters prior to observing the data. The prior distribution of $\boldsymbol{\theta}$ is represented as $p(\boldsymbol{\theta})$. Bayes' Theorem applied to statistical models leads to the following:

$$P(\boldsymbol{\theta}|\boldsymbol{y}) = \frac{p(\boldsymbol{y},\boldsymbol{\theta})}{m(\boldsymbol{y})} = \frac{p(\boldsymbol{y}|\boldsymbol{\theta})\,p(\boldsymbol{\theta})}{m(\boldsymbol{y})} \propto p(\boldsymbol{y}|\boldsymbol{\theta})\,p(\boldsymbol{\theta}) \tag{4}$$

where $p(\boldsymbol{\theta}|\boldsymbol{y})$ is the posterior distribution of $\boldsymbol{\theta}$, and $m(\boldsymbol{y})$ is the marginal distribution of $\boldsymbol{y}$. The posterior is proportional to the numerator because $m(\boldsymbol{y})$ is a known quantity and thus a constant of proportionality.

The posterior distribution represents the updated belief about $\boldsymbol{\theta}$ after observing the data $\boldsymbol{y}$. Bayesian inference is this process of updating beliefs based on evidence. The data may change opinions about the parameters and often sharpen them.

Bayesian inference is subjective in that the prior distribution is the prior opinion or belief about $\boldsymbol{\theta}$. Different people may arrive at different conclusions about $\boldsymbol{\theta}$, even though they observed the same data $\boldsymbol{y}$. The subjectivity of Bayesian methods is the source of criticism, the argument being that it seems "unscientific" for personal opinions to affect the conclusions of a study.

Hoff (2009) offers these counters to the criticism of subjectivity. First, in "unscientific" analysis, it seems natural that prior opinions affect one's conclusions. Also, when a relatively large amount of data is available, the prior has little effect on the posterior, unless the prior is very informative. Last, in scientific studies, an objective approach is to use the so-called noninformative class of priors. Noninformative priors are analogous to an experimenter expressing ignorance about the unknown parameters.

The posterior distribution of $\boldsymbol{\theta}$ may either be derived analytically using the so-called conjugate family of models or approximated using Markov Chain Monte Carlo (MCMC) methods. The posterior distribution has a standard or known form if the conjugate family of models is used. A prior is a conjugate prior if the posterior density has the same form as the prior.

To introduce the conjugate family of models, consider the simplest possible models: single parameter models. An example of a single parameter model is the binomial experiment. The binomial experiment consists of a sequence of independent trials, each of which results in either a 1 or 0, for "success" or "failure." It is assumed that the probability of a success is constant from trial to trial, and this probably is $\boldsymbol{\theta}$.

In the binomial experiment, there are independent and identically distributed observations $Y_1, \ldots, Y_n$ where Y_i has the following Bernoulli distribution:

$$P\left(Y_i = y\right) = \theta^y \left(1-\theta\right)^{1-y} I_{\{0,1\}}\left(y\right)$$

where $I_{\{0,1\}}(y)$ is the indicator function. The joint conditional distribution of the data given θ is:

$$P\left(y_1,\ldots,y_n|\theta\right) = \prod_{i=1}^{n}\theta^{y_i}\left(1-\theta\right)^{1-y_i} I_{\{0,1\}}\left(y_i\right) = \exp\left[\sum_{i=1}^{n} y_i log\theta + \left(n-\sum_{i=1}^{n} y_i\right)\log\left(1-\theta\right)\right] I_n$$

where $I_n = \prod_{i=1}^{n} I_{\{0,1\}}\left(y_i\right)$. If the prior distribution of θ is $p(\theta)$, then the posterior distribution is:

$$p\left(\theta|y_1,\ldots,y_n\right) = \frac{p\left(\theta\right)\theta^y\left(1-\theta\right)^{n-y}}{\int_0^1 p\left(t\right)t^y\left(1-t\right)^{n-y}dt} \propto p\left(\theta\right)\theta^y\left(1-\theta\right)^{n-y}$$

where $y = \sum_{i=1}^{n} y_i$. The posterior distribution always depends on the data only through a sufficient statistic. The sufficient statistic contains all the information in the data set. In the binomial experiment, the sufficient statistic is *y*, the total number of successes.

Now suppose that the prior distribution is the commonly used beta(*a*, *b*) density, which has the following form:

$$p(\theta|a,b) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \theta^{a-1} (1-\theta)^{b-1} I_{(0,1)}(\theta)$$

where $a > 0$, $b > 0$ and Γ is the gamma function. The posterior distribution is now:

$$p(\theta|y_1,\ldots,y_n) \propto \theta^{y} (1-\theta)^{n-y} \theta^{a-1} (1-\theta)^{b-1} = \theta^{y+a+1} (1-\theta)^{n-y+b-1}$$

where the last equation is proportional to a beta($y + a$, $n - y + b$). So if the beta prior is used in the binomial experiment, the resulting posterior is also beta, but with different parameters. The beta prior is an example of a conjugate prior for the binomial experiment.

Multivariate models are analogous to single parameter models. Take, for example, the normal distribution, N(μ, σ^2), which is very prevalent in practice. The normal distribution has the density:

$$f(y \mid \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma}} exp\left(-\frac{1}{2}\left(\frac{y-\pi}{\sigma} \right)^2 \right)$$

Let $\boldsymbol{Y} = (Y_1, \ldots, Y_n)$, where $Y_1, \ldots, Y_n$ is a random sample from N(θ_1, $1/\theta_2$) and $\Theta = \{(\theta_1, \theta_2): -\infty < \theta_1 < \infty, \theta_2 > 0\}$. Here θ_1 is the mean of the distribution and θ_2 is the precision, or the reciprocal of the variance. Now the normal-gamma family is a conjugate family of priors for the normal model.

Let θ_2 have a gamma(a, b) distribution and $\theta_1 | \theta_2$ have a N(μ, $(\tau\,\theta_1)^{-1}$). Then the joint distribution of (θ_1, θ_2) is normal-gamma:

$$p(\theta_1,\theta_2) \propto \sqrt{\theta_2}\,exp\left(-\frac{\tau\theta_2}{2}(\theta_1-\mu)^2\right)\theta_2^{a-1}e^{-b\theta_2}$$

which has four parameters, $a > 0$, $b > 0$, $\tau > 0$, and μ, which is unrestricted. Assuming a normal-gamma prior, the posterior is:

$$p(\theta_1,\theta_2 \mid y) \propto \sqrt{\theta_2}\,exp\left(-\frac{(\tau+n)\theta_2}{2}(\theta_1-\mu')^2\right)\theta_2^{a+n/2-1}e^{-b'\theta_2}$$

$$\mu' = \frac{\tau\mu + n\bar{y}}{\tau+n}$$

$$b' = b + \frac{1}{2}\sum_{i=1}^{n}(y_i-\bar{y})^2 + \frac{\tau n(\bar{y}-\mu)^2}{2(\tau+n)}$$

In summary, the posterior is normal-gamma as well. The marginal posterior distribution of θ_2 is gamma and the marginal posterior of θ_1 is normal.

If a conjugate model is not available or not known, the posterior distribution may be approximated using MCMC methods. MCMC methods are a class of algorithms for sampling from a target distribution based on forming a Markov chain that has the target distribution as the equilibrium distribution.

The most common MCMC algorithms are the Gibbs and Metropolis-Hastings algorithms. The Gibbs algorithm is a method of sampling from a distribution when it is known how to sample from each of the full conditional distributions. If the full conditionals are not known, the more general Metropolis Hastings algorithm may be used. The Metropolis Hastings algorithm generates a Markov chain from a proposal density and uses an acceptance ratio to accept or reject proposed values. Figure 7 shows an MCMC simulation run which is produced using the Metropolis Hastings algorithm. The plot above shows a Markov chain of simulated values for θ, while the plot below shows the density estimate of θ for the simulated values.

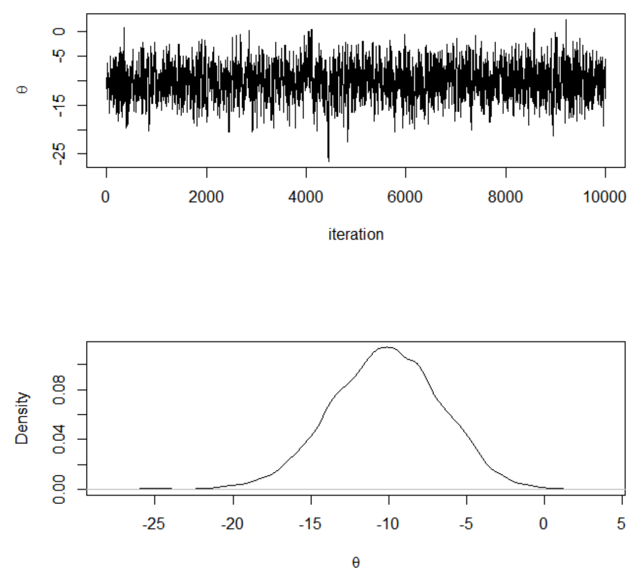


Figure 7: Example MCMC Simulation

2.4.2 Bayesian Networks

Bayesian networks are directed acyclic graphs (DAG) in which nodes represent random variables, and arcs represent conditional dependencies between variables (Pearl, Bayesian Networks: A Model of Self-Activated Memory for Evidential Reasoning, 1985). The strengths of these dependencies are expressed by conditional probabilities.

Bayesian networks are essentially a combination of graph theory and Bayesian theory (Ibe, 2011). These networks represent conditional dependency structures, prior beliefs about systems of variables, and observational events which are used to update beliefs.

Judea pearl first coined the term “Bayesian networks” in his influential 1985 paper, Bayesian Networks: A Model of Self-Activated Memory for Evidential Reasoning. Pearl was motivated “by attempts to devise a computational model for humans’ inferential reasoning, namely, the mechanism by which people integrate data from multiple sources and generate a coherent interpretation of that data.” Other texts that appeared shortly after Pearl’s article, like Probabilistic Reasoning in Intelligent Systems (1988) and Probabilistic Reasoning in Expert Systems (1989), helped further develop Bayesian networks and lead to the application of the methodology to a wide range of fields.

Bayesian networks have often been applied to specific risk analysis problems. See, for example, Weber et al. (2012) overview of applications for risk analysis, dependability, and maintenance. Frequently, Bayesian networks are considered for assessing financial risk (Neil, Fenton, & Tailor, 2005), (Shenoy & Shenoy, 1999). Research has also shown that Bayesian networks may be integrated with other risk analysis methods, such as fault tree analysis and event tree analysis (Roed, Mosleh, Vinnem, & Aven, 2009), (Groth, Wang, & Mosleh, 2010). Weber et al. (2012) argue that one of the primary weak points of these applications is that no formal guide to model development exists, which ensures model coherence. Further, a current research need is for the development of tools to help formalize Bayesian network models and integrate different dimensions of information, such as technical, organizational, information, decision, and finance considerations. This research contributes to this need by introducing a formal model of project risks and their interrelationships as well as tools to guide formal model development.

To illustrate Bayesian networks, consider the following simple example with events *A*, *B*, and *C,* shown in Figure 8.

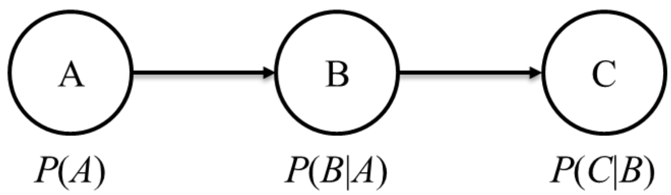


Figure 8: Bayesian Network for *A*, *B*, and *C*

where *C* is conditionally dependent on *B*, and *B* is conditionally dependent on *A*. Further, *A* and *C* are said to be conditionally independent. It is also reasonable to say that event *A* affects *B* directly, while *A* affects *C* indirectly, through *B*. Bayes' Theorem applies to the simple network in the following way:

$$P(C|B) = \frac{P(C,B)}{P(B)} = \frac{P(B|C)P(C)}{P(B)}$$

If the prior probability of *C* is known, and later *B* is observed, then the revised likelihood of *C*, or posterior probability, is $P(C|B)$. Similarly, the following is true:

$$P(B|A) = \frac{P(B,A)}{P(A)} = \frac{P(A|B)P(B)}{P(A)}$$

$$P(C|A \cap B) = P(C \mid B)$$

To obtain the probability of *A*, *B*, and *C* or the probability of the network, the Law of Total Probability gives the following:

$$P(A,B,C) = P(C|B)P(B|A)P(A) = \frac{P(B|C)P(C)}{P(B)}\frac{P(A|B)P(B)}{P(A)}P(A)$$

As another simple example, consider the following Bayesian network for events *D*, *E*, and *F*, shown in Figure 9.

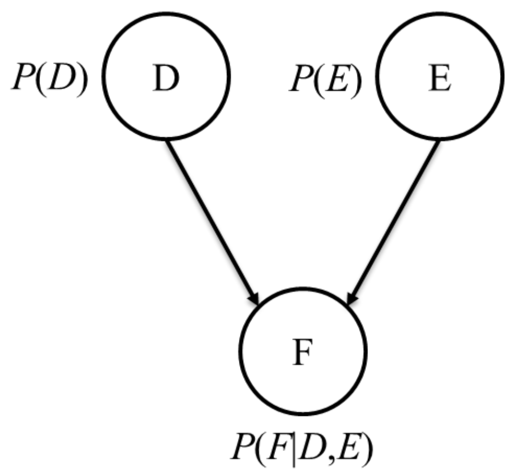


Figure 9: Bayesian Network for *D, E, F*

Now the probability of the network follows from the concepts of conditional dependence and independence:

$$P(D,E,F) = P(F|E \cap D)P(E)P(D) = P(F|E)P(F|D)P(E)P(D)$$

$$P(D,E,F) = \frac{P(E|F)P(F)}{P(E)}\frac{P(D|F)P(F)}{P(D)}P(E)P(D)$$

In general, consider n random variables X_1, X_2, …, X_n and a directed graph with n numbered nodes. The graph is a Bayesian network if:

$$P\left(X_1, X_2, \ldots, X_n\right) = \prod_{j=1}^{n} P(X_j \mid parents\left(X_j\right)) \tag{5}$$

The parents(X_j) denote the set of all variables X_i, such that a directional edge exists from node i to node j in the graph (Pourret, Naim, & Marcot, 2008). This notation indicates that the joint probability of the network is the product of the probabilities of each variable given the parent values. If the variable X_i has no parents, then:

$$P(X_j \mid parents\left(X_j\right) = P\left(X_i\right)$$

Furthermore, any joint probability distribution may be represented by a Bayesian network since the following is true (Pourret, Naim, & Marcot, 2008).

$$P\left(X_1, X_2, \ldots, X_n\right) = P\left(X_1\right) P\left(X_2, \ldots, X_n \mid X_1\right)$$

$$P(X_1, X_2, ..., X_n) = P\left(X_1\right) P\left(X_2 \mid X_1\right) \ldots P(X_n \mid X_1, \ldots, X_{n-1}) \tag{6}$$

Classifications of Bayesian Networks

Bayesian networks are classified in a number of ways. First, in regards to network structure, Bayesian networks are either singly connected or multiply connected (Ibe, 2011). Singly connected networks are also referred to as polytrees and have at most one path between any two nodes. A multiply connected network, on the other hand, has at least one pair of nodes with more than one path between them.

Bayesian networks are also classified by the types of random variables in the model. These types include discrete models, continuous models, and hybrid models, which consist of both discrete and continuous random variables.

Finally, Bayesian networks may be either static or dynamic networks. Static networks have nodes with constant values over time, whereas dynamic networks have at least some nodes that change over time. Dynamic networks are usually time-invariant with respect to network structure so that the topology is a repeating structure.

The two primary applications of Bayesian networks are Bayesian inference and Bayesian learning, which are discussed further below.

Bayesian Inference

As alluded to in the previous section, conditional independence is an important concept in Bayesian inference. Suppose, for example, a Bayesian network consists of the variable or node set $X = \{X_1, \ldots, X_n\}$, and suppose the variables $Y_1, \ldots, Y_k$ are observed to have the values $y_1, \ldots, y_k$. The conditional probability of X given the observations is:

$$P\left(X|Y_1 = y_k, \ldots, Y_k = y_k\right) = \frac{P(Y_1|Y_2, \ldots, Y_k, X)P(Y_k|X)P\left(X\right)}{\sum_{v=domain\left(X\right)} P(Y_1|Y_2, \ldots, Y_k, X)P(Y_k|X)P\left(X\right)} \tag{7}$$

Due to conditional independence, a variable can be removed from the equation if that variable is not a child of X.

Bayesian inference, however, is difficult with complex models. For this reason, researchers have developed inference algorithms. Inference algorithms may be divided into two classes: exact and approximate algorithms. Exact algorithms store belief distributions of the possible values on each node of the network. If the belief distribution of a node changes, messages are propagated to adjacent nodes in order to update the belief distributions of the nodes that receive them. This process is repeated until convergence is reached for all nodes in the network. Some common types of exact algorithms include: variable elimination, belief propagation, and clique tree propagation, which is also called the junction tree algorithm (Ibe, 2011), (Darwiche, 2009), (Neapolitan, 2003).

Approximate algorithms use the Bayesian network as a number generator to produce a sample from which the result is estimated using the relative frequency of the generated values. Common approximate algorithms include: stochastic sampling and MCMC, variational methods, and loopy belief propagation (Ibe, 2011), (Darwiche, 2009), (Neapolitan, 2003).

Bayesian Learning

Bayesian learning refers to the process of learning the Bayesian network from the observed data. Learning the network follows directly from Bayesian methods, where the prior knowledge is combined with data to produce improved knowledge.

Bayesian learning can be classified into two types: structural learning and parameter learning. Structural learning involves selecting the arcs between a given set of variables, which usually requires approximate methods (Ibe, 2011), (Darwiche, 2009), (Neapolitan, 2003). This dissertation does not explicitly explore the use of structural learning, and this problem is left for future work.

Parameter learning is an exercise in Bayesian statistical modeling. Parameter learning refers to the procedure of estimating the posterior of $\boldsymbol{\theta}$. One can then use the posterior to make inference about $\boldsymbol{\theta}$. A common type of Bayesian inference is point estimation, and popular types of Bayesian point estimates are the mean of the posterior, and the posterior mode, which is analogous to the maximum likelihood estimate (MLE).

Challenges with Bayesian Networks

Bayesian networks have a number of common challenges. First and foremost, decisions must be made as to what variables to include in the network and what conditional dependencies exists between the variables. The network topology is particularly challenging when no theoretical model of the study area exists (Rad & Cioffi, 2000). Another common problem with Bayesian networks and Bayesian methods in general is the selection of prior distributions. If the distribution is not known, and data are not available, priors are usually based on belief measures. However, transforming someone's belief about an event into a prior distribution is a challenge in itself. Finally, a common challenge is to obtain the data necessary to update the network. Data should be gathered to test the network topology and parameter probability distributions.

While the network topology remains a challenge, this dissertation proposes a theoretical model of the study area, project risk management, which guides model development and provides guidance on selecting prior distributions and data collection, which follows from current risk analysis methods.

2.4.3 Causal Networks

The causal network is a model of how humans reason with causes (Neapolitan, 2003). If one identifies direct cause-effect relationships (edges), draws a causal DAG using the edges identified, and assumes the probability distribution of the variables satisfies the Markov condition, then the model is a causal network. In other words, when the DAG in a Bayesian network is a causal DAG, the network is called a causal network. Causal networks are the branch of network methods where nodes may or may not represent random variables, and arcs represent causation between nodes. Bayesian networks, furthermore, are a special case of causal networks. However, in this dissertation, the terms Bayesian and causal are typically used interchangeably.

2.5 Corporate Portfolio Risk Management

Portfolio theory holds that assets in an investment portfolio should not be selected independently because returns are often correlated. It is important to measure how the return of each asset changes relative to every other asset in the portfolio. Portfolio theory explains how to select assets with the highest possible expected return for a given amount of risk. Conversely, the theory describes how to select assets with the lowest possible risk for a given expected return. The concept of diversification is that in order to balance risk and return, the portfolio should hold the right kinds and amounts of assets.

Harry Markowitz (1952) introduced portfolio theory in his classic paper, “Portfolio Selection.” He later expanded his ideas in the book, *Portfolio Selection: Efficient Diversification of Investments* (1959). Markowitz held that “the first stage [of the process of selecting a portfolio] starts with observation and experience and ends with beliefs about the future performances of available securities. The second stage starts with the relevant beliefs about future performances and ends with the choice of portfolio”

(Markowitz, 1952). Portfolio theory inspired numerous extensions and applications, and remains an important topic in many fields.

Portfolio theory assumes that investors are rational and risk averse, meaning that given the option between two portfolios with the same expected returns, the investor selects the one with less risk. Thus, an investor must be compensated by higher returns in order take on additional risk. The exact tradeoff between risk and return, however, depends on an investor's utility or risk aversion characteristics.

In general, the expected return of the portfolio $E(R_p)$ is:

$$E\left(R_p\right)=\sum_{i=1}^{n} w_i E\left(R_i\right) \tag{8}$$

where R_p is the return of the portfolio, R_i is the return of asset i, w_i is the weight or proportion of asset i, and n is the number of assets in the portfolio. The variance of the portfolio σ_p^2 is:

$$\sigma_p^2=\sum_{i=1}^{n} w_i^2 \sigma_i^2+\sum_{i}\sum_{j\neq i} w_i w_j \sigma_i \sigma_j \rho_{ij} \tag{9}$$

where ρ_{ij} is the correlation coefficient between the returns on assets i and j. The standard deviation is often referred to as the portfolio return volatility:

$$\sigma_p=\sqrt{\sigma_p^2} \tag{10}$$

Corporate portfolio management is the application of investment portfolio theory to project management. The analogy is that projects are firm-related assets, and like financial instruments, projects are often correlated. Furthermore, multi-project management infers project portfolio management and not simply the management of multiple projects independently (Olsson, 2008). Firms should select and manage projects while considering the other projects in the portfolio. A diversified corporate portfolio is one which balances risk and return with the right types and amounts of projects.

The corporate portfolio depends on the organization of the firm as well as the environment in which it operates. The firm organization consists of a set of resources and capabilities that combine for the overall strategic objectives (Sanchez, Robert, & Pellerin, 2008). The corporate portfolio is the approach by which the firm organization implements these strategic objectives. Furthermore, the corporate portfolio is dynamic in that projects are continuously selected, reprioritized, or terminated to meet objectives.

The organization of the firm should also be considered in context of the environment in which it operates. The environment refers to external conditions that do not directly impact the firm but have a broad impact on the environment in which it operates. Firms must consider how these conditions impact the corporate portfolio and how conditions change over time. Figure 10 illustrates the general corporate project portfolio organization and environment. The figure indicates the corporate portfolio is influenced by organizational factors, such as stakeholders, technology, resources, and capabilities, and the organization is impacted by the external environment in which it operates, where the environment includes factors such as the market, government, resources, and capabilities.

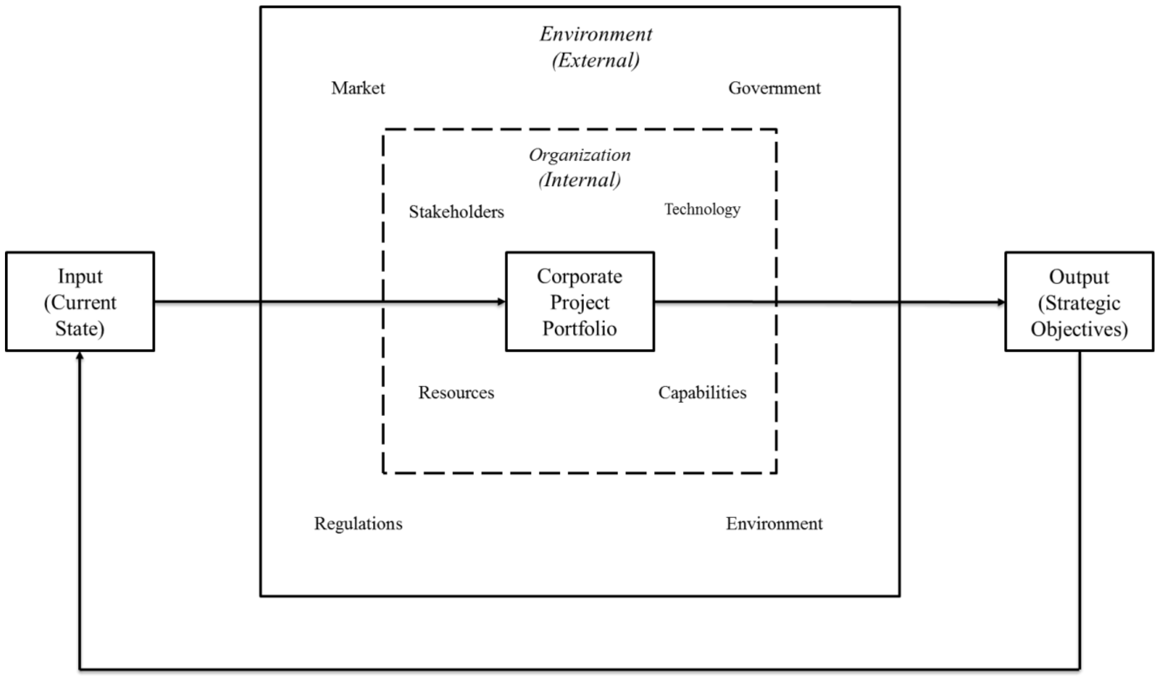


Figure 10: The Corporate Portfolio Organization-Environment (Adapted from Sanchez et al. 2008)

Corporate risks are events that affect the strategic objectives of the corporate portfolio. Risks may be environmental and materialize outside of the firm (i.e. market, government, regulatory, or environmental-related) or internal and relate to the organization (e.g. technical risks). Project portfolio risk management involves identifying and measuring these corporate risks.

The broad view of corporate portfolio management includes the management of interdependencies between projects, the coordination of multiple projects, the management of resources and constraints, and the link to strategic objectives (Olsson, 2008). Firms must allocate a limited number of resources to projects in a way that balances risk, reward, and strategic objectives (Dickinson, Thornton, & Graves, 2001).

Firms may benefit from leveraging key processes and procedures among multiple projects by considering interactions among projects (Teller, 2013). Firms are not only concerned with maximizing profit, moreover, and the long term goal is to grow with balanced risk and return (Han, Diekmann, Lee, & Ock, 2004). The balance between risk and return is firm specific and depends on the firm's risk aversion characteristics (Dickinson, Thornton, & Graves, 2001).

Project portfolios and traditional financial portfolios have some key differences. First, financial instruments are divisible, while projects are discrete (Hubbard, 2009). Portfolio theory may explain the optimal shares of stocks, but the optimal proportions of projects may not be practical. It may not be easy or even possible to change the amounts and types of projects. Projects have logical units, which cannot always be separated. Second, financial instruments are liquid, and can be measured at any time (Hubbard, 2009). Projects, on the other hand, are limited and are usually only valuable when complete. An unfinished project offers little to no recovery or salvage value. Project portfolio management, furthermore, should consider these additional constraints that do not apply to traditional financial portfolios.

An important difference also exists between traditional portfolio risk analysis and project risk analysis. A project risk analysis is usually a structural model that seeks to model project components and their interrelationships. For example, if component *A* fails, then component *B* is affected and so on. Portfolio theory, however, does not explain the cause of changes in the prices of stocks. Assets are assigned probabilities based on the history of the assets. If historical data does not exist, then it is not possible to assign probabilities. If project risk analysis operated in the same way, then it would rarely provide any useful insights (Hubbard, 2009).

Historically, project risk management is concerned with managing individual projects (Olsson, 2008). As such, relatively little research exists on project portfolio risk management. However, the firm-related risk may be considerably different than the sum of individual project risks (Han, Diekmann, Lee, & Ock, 2004). If projects are managed independently, then the firm must rely on the experience of the organization to identify links between projects. Furthermore, the firm may fail to choose the optimal combination of risk and return.

The methods for corporate portfolio risk management that do exist range from the quantitative (e.g. return on investment) to the qualitative (e.g. alignment with strategic objectives) (Dickinson, Thornton, & Graves, 2001). Vergara (1977) and Vergara and Boyer (1977) suggested using portfolio theory to support the decision to bid or not bid on projects. Vergara and Boyer (1977) proposed four steps for portfolio selection: (1) individual analysis of the existing projects; (2) analysis of the all of the existing projects in the portfolio; (3) analysis of possible new project(s) on which the firm could bid; (4) choice of the best portfolio of projects for a firm and determination of the optimal bid price for the new project(s).

Minato (1994) first discussed the concept of corporate risk in the project portfolio. He argued that corporate risk could be controlled using strategies at the corporate level. Minato suggested using the common financial metric beta to assess risk. Beta is defined as:

$$\beta = \frac{Cov(R_a, R_b)}{Var(R_b)} \tag{11}$$

where R_a is the return of asset *a* and R_b is a benchmark. Minato defined R_b as the return of the overall project portfolio and R_a as the return of a single new project. Minato's method, however, is limited due to the limitation of historical data, the difficulty of collecting data, and the lack of qualitative assessment.

Dickinson et al. (2001) proposed a method to qualify the interdependencies between projects and optimize project portfolios. First, they used a so-called project dependency matrix (i.e., DSM) to qualify the interdependencies between projects. The project dependency matrix is a square matrix of size n_p, the number of projects. Each project has one column and one row ordered sequentially. Each element in the matrix, d_{ij} varies from 0 to 1, where the value of d_{ij} represents the level of dependency between projects *i* and *j*. A value of 1, for example, implies that project *i* is completely dependent on project *j*. A value of 0, conversely, indicates that project *i* is independent of project *j*. Thus, the project dependency matrix is analogous to the correlation matrix. The matrix is a scalable and flexible method in that it can be used to evaluate all different types of portfolios for a single period or across multiple periods. An example of a dependency matrix is shown in Figure 11.

	A	B	C
Project A	A	0.5	1
Project B	0.5	B	0
Project C	1	0	C

Figure 11: Project Dependency Matrix

Second, Dickinson et al. developed a nonlinear, integer optimization model to select optimal portfolios. The optimization model combines the dependency matrix with user defined financial metrics to estimate the performance of a portfolio. The model is subject to multiple qualitative balance and budget constraints, which are also set by the user. Once an initial portfolio is specified, the model can be used to explore the impact of different changes to a portfolio.

Han et al. (2004) proposed a multi-criteria approach to project portfolio risk management. Their approach consists of three criteria: maximize expected value, minimize risk variability, and maximize efficiency. First, the expected value is taken as the expected net present value (NPV) of the cash flows. The NPV of a series of cash flows, both incoming and outgoing, is the sum of the present values of the cash flows. The present value is discounted assuming a rate of return (ROR). The NPV is:

$$NPV = \sum_{t=1}^{n} \frac{R_t}{(1+i)^t} \tag{12}$$

where R_t is the cash flow at time t, either in or out, i is the ROR, and n is the number of periods. Han et al. assume that the probability distribution of NPV is approximately the Pearson-Tukey three point estimate (median, 0.05, and 0.95 quantiles).

Second, risk variability is measured using the concept of value at risk (VaR). Given the probability distribution of a return, R_p, on a given investment, the VaR at level α (VaR_α) for the return R_p is the value x, for which the probability of obtaining a return less than x is α. The VaR approach for normally distributed returns is given as the following (J.P. Morgan, 1996):

$$VaR = W_o \alpha \sigma \sqrt{\Delta t} \tag{13}$$

where W_o is the initial investment, α is the z value at the specified confidence level, σ is the standard deviation of the return, and Δt is the time interval. Figure 12 gives an example of VaR.

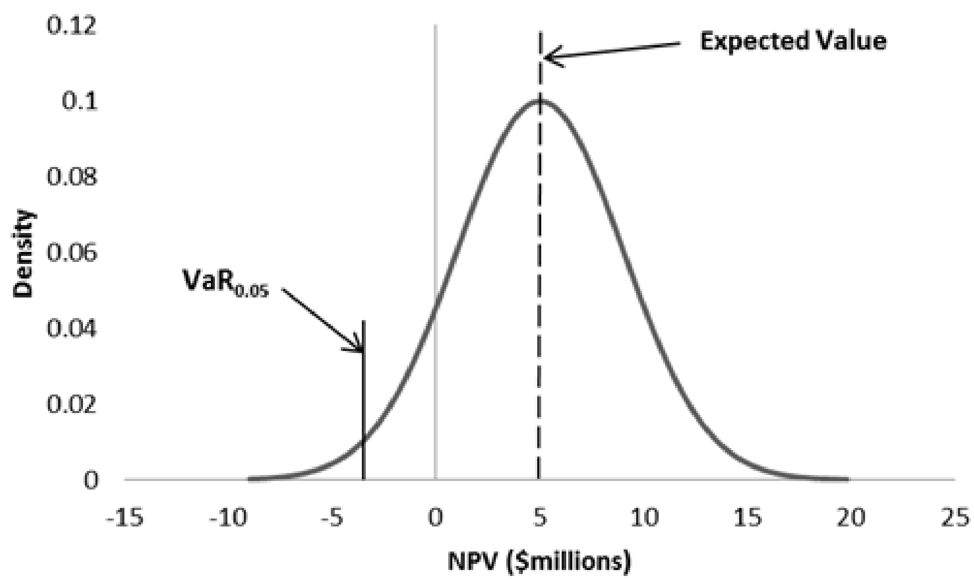


Figure 12: Example of VaR

Third, the efficiency is measured using the return on investment (ROI). The ROI is a common performance indicator of the so-called effectiveness of assets. Han et al. define ROI as the expected NPV divided by the estimated total project cost.

$$ROI = \frac{NPV}{project\,cost} \tag{14}$$

Han et al. outline their procedure for portfolio selection in five steps: (1) gather data on the current portfolio's risk and return measures, including the variance of currency exchange rate, discount rate, and depreciation rate for each project; (2) evaluate the risk and return of each potential project using the NPV, VaR, and ROI criteria; (3) produce possible new sets of portfolios; (4) evaluate the possible portfolios based on the three criteria and the firm's key targets; (5) select the portfolio that is in line with the firm's strategic objectives; (6) provide feedback/monitoring over time to the portfolio analysis cycle.

Caron (2007) also used a value at risk approach to project portfolio management. They extend the application of VaR to another popular concept, conditional value at risk (CVaR). CVaR is defined as the expected loss corresponding to VaR_{α}. Caron used Monte Carlo (MC) simulation to estimate the probability distribution of NPV and then use CVaR to estimate the conditional net present value at risk.

Sanchez et al. (2008) proposed a risk/opportunity identification framework. They took a so-called systems theory perspective and modeled the portfolio of projects as the system and the environment as all factors interacting with the system. Sanchez et al. argue that resources (including knowledge and strategy) are fundamental to maximizing the value of the portfolio. Furthermore, it is possible to evaluate risk/opportunity by considering resource interdependencies. They proposed a general framework for identifying these resource interdependencies and linking them to strategic objectives. See Figure 13 for a simple example of their proposed framework. In the figure, *P* indicates project, *R* indicates resource, and *B* indicates benefit.

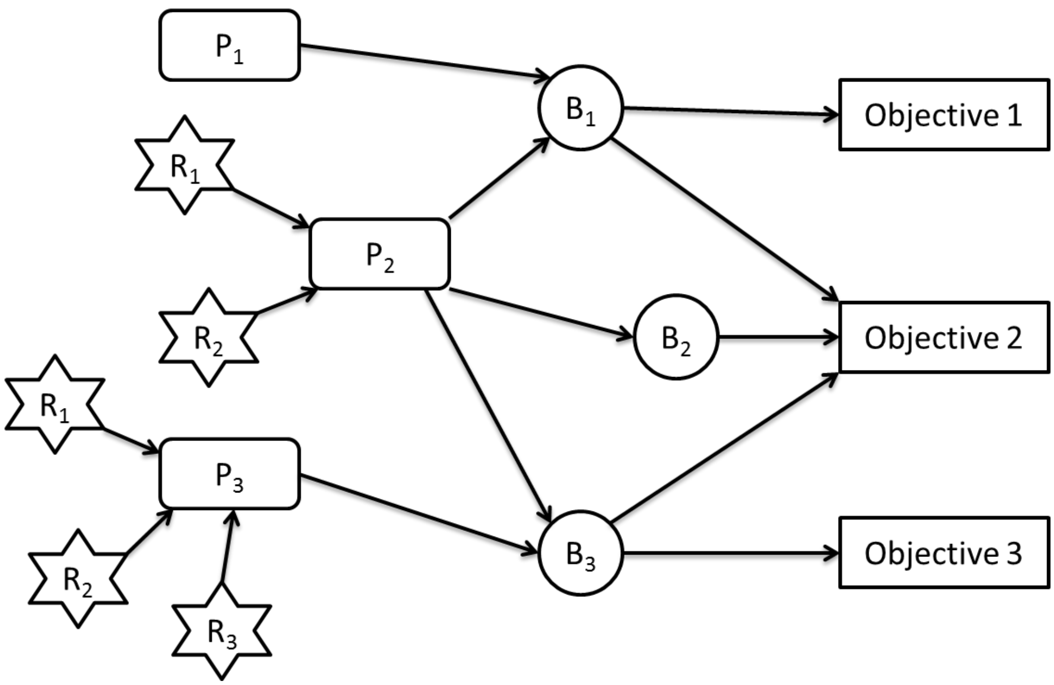


Figure 13: Model of Project Interdependencies and Link between Projects and Objectives (Adapted from Sanchez et al. 2008)

Various other studies in the area of project portfolio risk analysis should be noted. Other notable research includes (Pellegrinelli, 1997), (Olsson, 2008), (Teller & Kock, 2013), and (Teller, 2013). The Project Management Institute (PMI) publishes the guidelines called The Standard for Portfolio Management (2008). PMI proposes a four step process for managing risk in project portfolios and defines different categories of portfolio risks. However, current methods do not include a structural model that accounts for inter-dependencies between projects and integrates expert opinions and historical data. This dissertation proposes a general structural model for project portfolio risk management by extending the project risk model.

2.6 Summary

This chapter presented an overview of the literature related to the overall research objectives and provided the necessary background to develop an integrated risk analysis methodology. The literature review revealed research needs in each area discussed: the RBV, project risk management, causal networks, and project portfolio risk management. The following chapter provides an overview of the methodology that is used for the remainder of the dissertation.

CHAPTER III

METHODOLOGY

This chapter presents an overview of the proposed methodological framework for integrated project risk analysis, and includes four primary sections: the conceptual framework, the methodological framework, the research approach, and the research methodology. The first section presents the conceptual framework, which is based on the Resource-based View (RBV) of the firm. The second section presents the methodological framework, which consists of utilizing causal networks to model risk events and the associated consequences. The third section provides a general overview of the research approach, which serves as the outline for the remaining chapters. The fourth section presents the research methodology of causal networks for project risk analysis. This research methodology provides an outline for model development, validation, and analysis.

3.1 The Conceptual Framework

Chapter 2 presented an overview of the RBV, the theory that value exists in viewing the firm in terms of resources. In this section, the RBV is integrated with a risk-based view. If the firm is defined as a set of resources, then, logically, a risk event is an event or set of events that, if they occur, have one or more impacts, either positive or negative, on at least one resource. In other words, certain firm-related resources are at risk, and risk impacts the firm through the resources. Fundamentally, firm-related risk is associated with resources, and this association has significant consequences for strategic management.

Now consider this concept from the point of view of project-based firms, such as engineering, procurement, and construction companies. If the project(s) is defined as a subset of firm-related resources, then risk events impact the project(s) through project-related resources. This concept applies both at the individual project level as well as the corporate portfolio level, which leads to the conceptual framework for corporate portfolio risk analysis.

The conceptual framework is a structural model of risk events and their interrelationships. The model assumes a direct relationship between risk events and project-related resources, which leads to an overall impact on the corporate portfolio. Therefore, the general risk statement is the following: risk events are probabilistic events that may occur because a condition(s) is present, and the consequence(s) is the impact of the risk event on a resource(s). The conceptual framework for corporate portfolio risk analysis serves as the framework for development of the causal network methods.

The proposed methodological framework utilizes causal networks to identify and measure risk events that explain the uncertainty in project resources and, in turn, the uncertainty in projects. The causal network methodology provides a graphical explanation of risk events and the interdependencies between projects as well as a probabilistic model of project uncertainty. Firms may investigate different risk mitigation and management strategies by simply selecting a different set of causal network model parameters and/or a different structure and measuring the overall effect of these changes on project uncertainty.

The conceptual framework may be subject to criticism. First, some people may find it difficult to view all project cost as incurred by utilizing a resource or a set of resources. After all, some project cost is not often considered to be resource costs. For example, construction projects incur cost related to indirect expenses, right of way (ROW) acquisition, and taxes. However, given the definition of a resource, anything that is a strength or weakness of the firm, these components are, in fact, resources (Wernerfelt 1984). Indirect expenses may be linked to other resources, such as administrative and office expenses, ROW is a material resource, and taxes are actually a monetary resource.

Second, the model assumes a direct relationship between risk events and project resources. Current risk analysis methods do not provide a general framework for identifying the consequence of a risk event occurring. As such, risk events may not be linked to resources. For example, macro-economic risk events, such as inflation risk events, are typically associated with a broad impact on the economic environment in which the firm operates. However, inflation risk events may be linked to the unit costs of certain resources. Thus, a direct link exists between inflation risk events and resources.

Third, the concept is limited in that it only applies to risk event cost. The argument may be that not all risk events have the consequence of an impact on cost. Risk events also impact other project objectives, such as the project schedule. For example, if a supplier's delivery is late, then the project may be delayed. However, the risk event of a late delivery is associated with an impact on cost, as resources must be paid additionally for the delay. A suggestion for future research is to extend the general model to account for schedule risk events (i.e. availability, productivity, etc.). This extension is a logical consequence of the RBV: if risk events impact resources, and the schedule depends on resources, then the schedule is impacted through resources.

3.2 Methodological Framework

This conceptual framework may be implemented by utilizing causal networks. Consider a simple project that utilizes two resources with costs a_1 and a_2. Project management identifies three risk events r_1, r_2, and r_3. r_1 and r_2 impact a_1, and r_2 and r_3 impact a_2. The project cost is p. These dependencies are modeled in Figure 14 part A, where:

$$a_1 \subseteq r_1 \cup r_2$$

$$a_2 \subseteq r_2 \cup r_3$$

$$p \subseteq a_1 \cap a_2$$

More specifically, if risk events r_1 and r_2 impact the a_1 unit cost u_1, and risk events r_2 and r_3 impact the a_2 quantity q_2, then the previous graph now looks like Figure 14 part B, where:

$$u_1 \subseteq r_1 \cup r_2$$

$$a_2 \subseteq r_2 \cup r_3$$

$$p \subseteq (u_1 \cap q_2) \cap (u_2 \cap q_2)$$

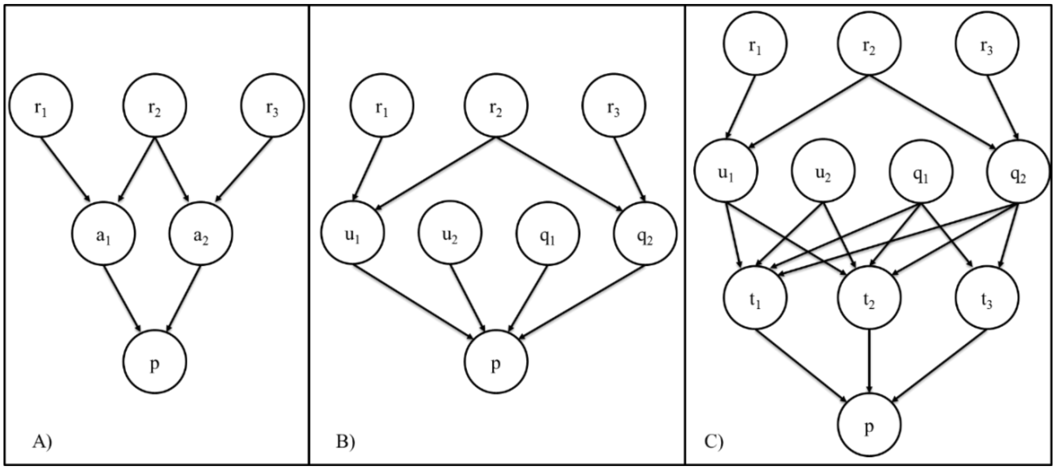


Figure 14: Example Project A (Left), B (Center), and C (Right)

Furthermore, projects are broken into components of work or tasks. Tasks require or depend on resources. Resources are often shared between tasks, just as risks are often shared between resources. This project consists of three tasks with costs t_1, t_2, and t_3. t_1 and t_2 depend on a_1 and a_2, and t_3 depends on a_2 only. The example project now looks like Figure 14 part C, where:

$$t_1 \subseteq (u_1 \cap q_1) \cap (u_2 \cap q_2)$$

$$t_2 \subseteq (u_1 \cap q_1) \cap (u_2 \cap q_2)$$

$$t_3 \subseteq (u_2 \cap q_2)$$

$$p \subseteq t_1 \cap t_2 \cap t_3$$

Actually, this theory holds for any work breakdown. All tasks depend on resources and larger tasks are simply combinations of smaller tasks.

The causal network methodology may be applied to any project or vector of projects (i.e., corporate portfolio) $\boldsymbol{p}$, vector of tasks $\boldsymbol{t}$, resource unit costs $\boldsymbol{u}$, quantities $\boldsymbol{q}$, risk events $\boldsymbol{r}$, and root causes $\boldsymbol{c}$. Figure 15 illustrates a hypothetical project risk network. The nodes represent random variables and the arcs represent causal dependencies between the variables. This causal network may be classified as a multiply connected, hybrid causal network. The network is multiply connected because at least one pair of nodes has more than one path between them. For example, in Figure 15, r_2 may impact t_2 through u_2 or q_2. Furthermore, the network is hybrid because it consists of both discrete and continuous variables. Root causes $\boldsymbol{c}$ and risk events $\boldsymbol{r}$ are discrete events with probability p of occurring and probability $1 - p$ of not occurring. Variables such as unit costs $\boldsymbol{u}$ and quantities $\boldsymbol{q}$ are continuously distributed over a certain range. The figure indicates discrete nodes $\boldsymbol{c}$ and $\boldsymbol{r}$ with a box and continuous nodes $\boldsymbol{r}$, $\boldsymbol{t}$, and p with a circle. Each variable in the causal network is defined as follows:

- **Root Cause**: the condition of the identified risk event. If the root cause is present, then the risk event is possible. The root cause is a discrete variable.
- **Risk Event**: the uncertain future event that may occur because the condition is present. If the risk event occurs, then the consequence is the impact on the resources. The risk event is a discrete variable.
- **Resource Unit Cost**: the cost per unit of resource, which depends on the resource type. The resource unit cost is a continuous variable.
- **Resource Quantity**: the amount of the resource required for the task, which depends on the task. The resource quantity is a continuous variable.

- **Task**: component of project work. The task is a continuous variable.

The parameters of the project risk network are conditional probability distributions. Risk events ***r*** may be conditionally dependent on root causes or early warning signs ***c***. Whether or not root causes are observed, risk events are uncertain, but observation of root causes may lower the uncertainty. Similarly, unit costs ***u*** and quantities ***q*** may be conditionally dependent on risk events ***r***, meaning that the probability distributions depend on whether or not risk events occur. Again, evidence of risk events does not necessarily eliminate the uncertainty associated with resource costs, but the evidence may lower the uncertainty.

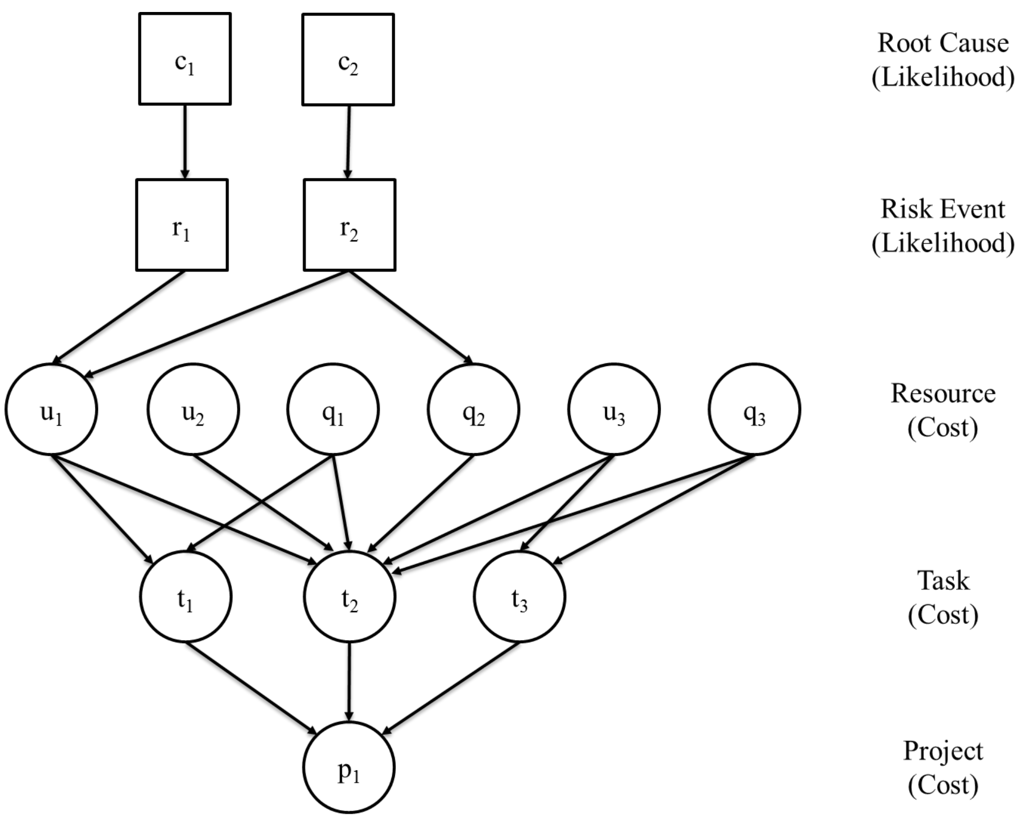


Figure 15: Hypothetical Project Risk Network

The causal network of project risk contributes to the field in two primary ways. First, the model is a graphical display of project risk events and a method to model interdependencies between risk events. Second, the model opens the field of project risk analysis to the broad field of causal network methods. The causal network represents the prior distributions of the project-related parameters, which may be updated based on information contained in data.

The causal network methodology also contributes in another unique way. The project risk network represents the joint impact of multiple risk events as probability distributions. Conversely, existing methods typically assume risk events are additive, meaning that the joint impact of multiple risk events is simply the summation of the impacts of the individual risk events. However, this joint impact may, in theory, not be additive. By representing the impacts of risk events as probability distributions, this methodological framework may, in fact, model the joint impact of multiple risk events whether or not they are additive.

3.3 Research Approach

The overall plan of research tasks and the proposed methods required to accomplish the research is presented in Figure 16. The dissertation is subdivided into six main research concepts, which are indicated by dashed boxes. First, the focus of research concept 1 is to define the project or projects from a resource-based view. As previously mentioned, existing project management tools support this concept. For instance, the project statement of work (SOW) links to the work breakdown structure (WBS), which, in turn, links to the resource breakdown structure (RBS), where the RBS is a hierarchical view of the project in terms of the resources. Second, the topic of research concept 2 is to link risk to project-related resources. Ex-ante risk first may be identified by utilizing existing project management tools and then linked to project resources. Third, research concept 3 is concerned with developing causal network methods for project risk assessment. Chapter 4 presents the formal causal network methodology for project risk analysis. Fourth, the purpose of research concept 4 is a network topology analysis. This concept proposes that the project risk network structure provides valuable information to project managers. Chapter 5 presents network measures for project risk analysis. Fifth, the purpose of research concept 5 is to extend the causal network methodology for corporate portfolio risk analysis. Finally, research concept 6 consists of a case study of the causal network methods for project risk management and mitigation. The purpose of the case study is to illustrate the general research methodology, to provide empirical support for the theoretical framework, and to evaluate the methodological framework for practical implications.

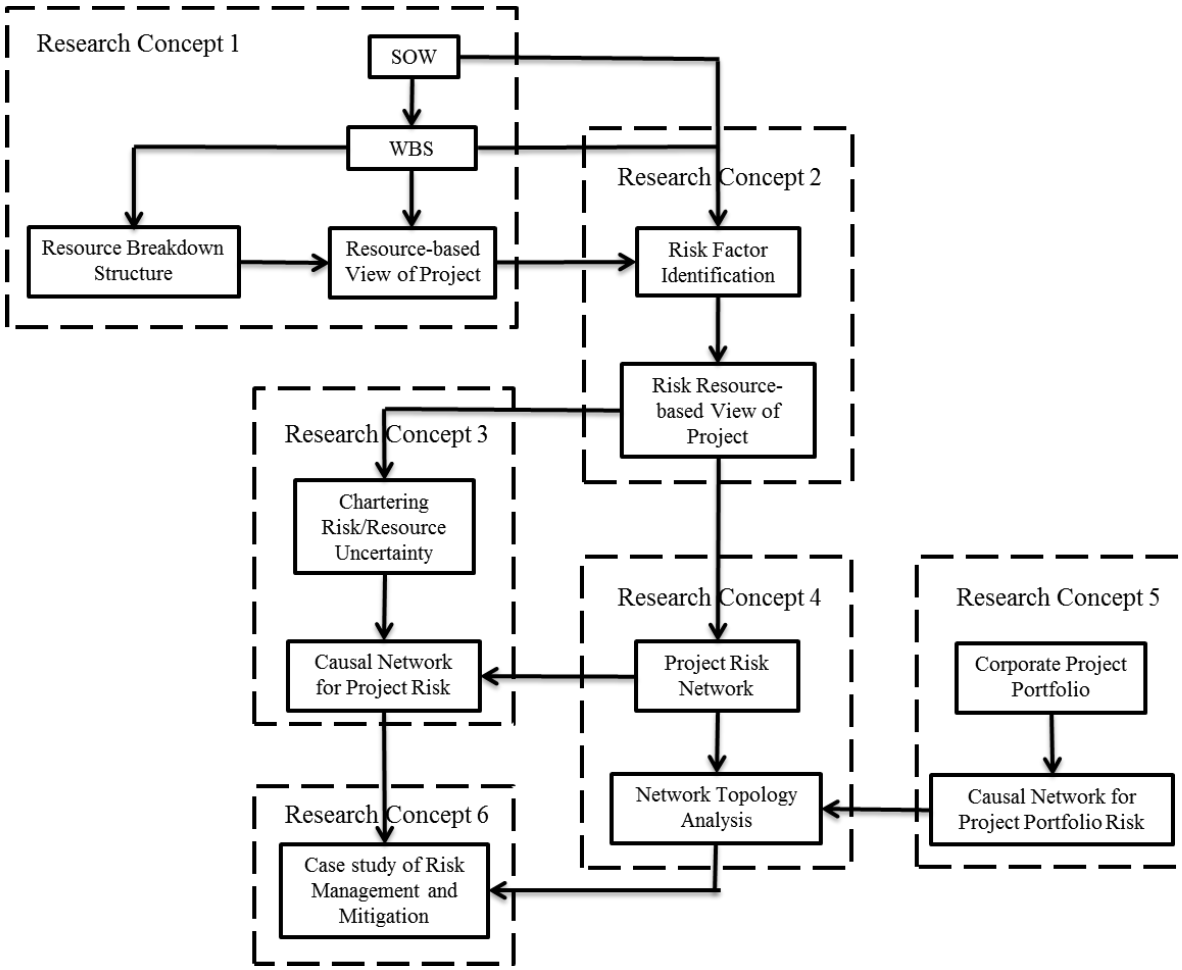


Figure 16: Research Approach

3.4 Research Methodology

The research methodology is a general outline for model development, validation, analysis, and use. Figure 17 illustrates the research methodology for the causal network method approach to project risk analysis. First, the model inputs are identified, such as work tasks, resources, costs, and risk events. Section 4.1 presents guidelines for selecting these model inputs from existing project management tools, such as the WBS, RBS, bottom-up cost estimate, and risk register. Second, the conditional dependencies and model parameters are specified. Section 4.1 discusses specifying the conditional dependencies and model parameters. The conditional dependencies are specified in the adjacency matrix shown in equation (15), and the model parameters are specified as conditional probability distributions shown in equations (16) through (21). Third, the model is validated using sensitivity analysis. Section 4.2 describes general Bayesian methods and sensitivity analysis for model validation. Fourth, the model is analyzed using Bayesian inference and network measures. Bayesian inference for project risk analysis is discussed in section 4.3 (equations (26) through (29) and algorithms 1 and 2). Network measures for project risk analysis are described in Chapter 5. Fifth, if data is gathered on the model parameters, then the next step is Bayesian learning. The learning process is a feedback loop to the model specification step. Bayesian learning is explained in section 4.4 (equations (30) through(33)). Finally, the model is utilized to assess different management and mitigation alternatives/options. This step is illustrated in Chapter 7, the case study of a compressor station project.

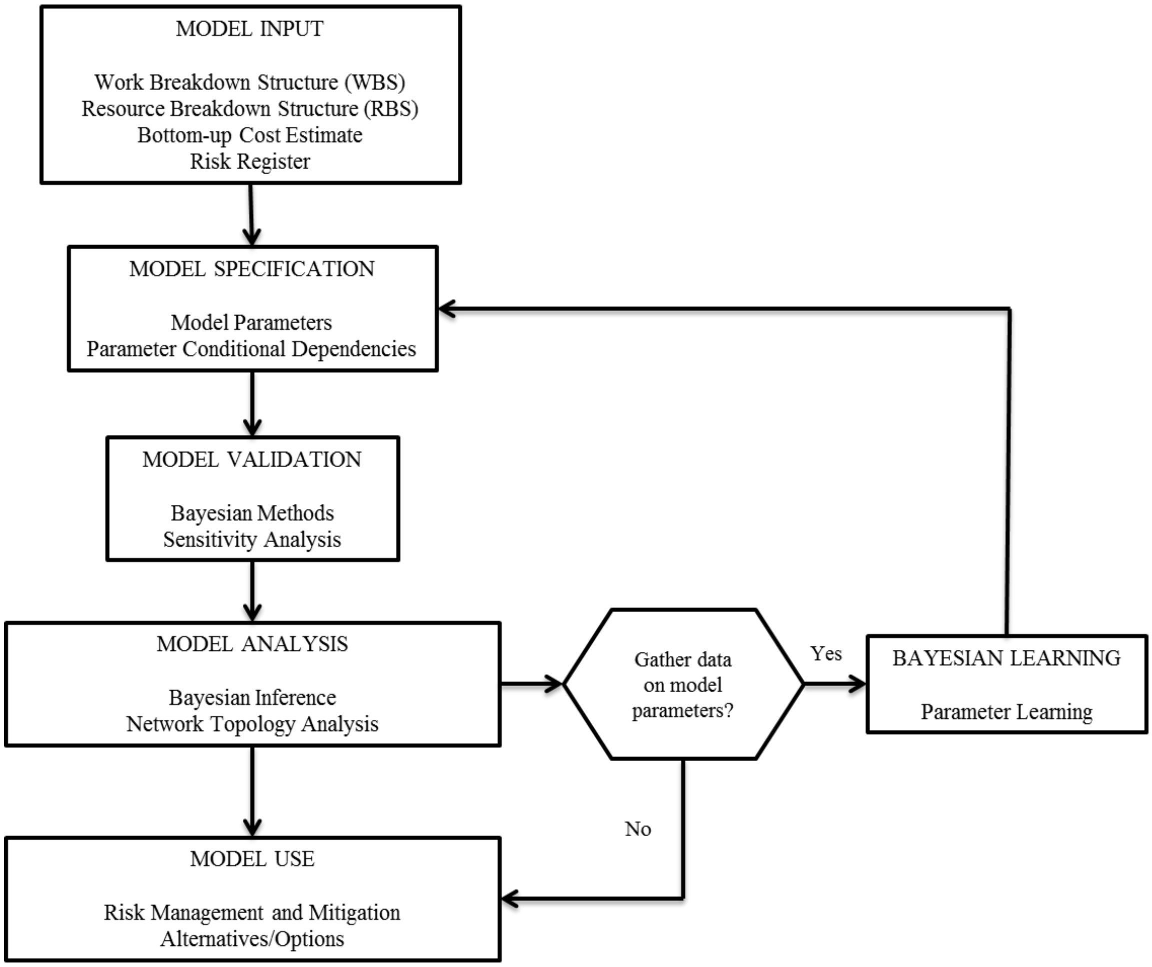


Figure 17: Research Methodology

3.5 Summary

This chapter presented an integrated framework for project risk analysis. First, the conceptual framework was introduced, which is based on the RBV. Second, the conceptual framework was implemented by utilizing causal networks to model risk events and the associated consequences. Third an overall research approach and methodology was developed for project risk analysis. The following chapter further develops the application of causal network methods for project risk analysis.

CHAPTER IV

CAUSAL NETWORKS FOR PROJECT RISK ANALYSIS

This chapter presents the causal network methodology for project risk analysis. The chapter includes four main sections: model development, model validation, Bayesian inference, and Bayesian learning. The first section presents the development approach, which involves identifying the model inputs and specifying the model parameters and conditional dependencies. The model inputs include the vectors of root causes, risk events, resource costs, and tasks. The model inputs are selected from existing project management tools, such as the WBS, RBS, bottom-up cost estimate, and risk register. The second section discusses the model validation approach, which consists of Bayesian methods and sensitivity analysis. A sensitivity analysis is a study of how the uncertainty in model outputs may be explained by different sources of uncertainty in model inputs. The third section demonstrates Bayesian inference for project risk analysis. Bayesian inference has many relevant applications, such as estimating the probability distributions and correlations of the variables of the project risk network. The fourth section shows Bayesian learning of the project risk model, which involves updating the model based on the observed project data.

4.1 Model Development

The first step in model development is to specify the directed acyclic graph (DAG) of the project risk network. The DAG may be represented by the adjacency matrix $\boldsymbol{A}$, where $\boldsymbol{A}$ is as follows:

$$
\boldsymbol{A} = \begin{array}{c|cccccccccccccccc}
 & c_1 & \cdots & c_m & r_1 & \cdots & r_l & u_1 & \cdots & u_k & q_1 & \cdots & q_k & t_1 & \cdots & t_j & p \\
\hline
c_1 & & & & & & & & & & & & & & & & \\
\vdots & & & & & & & & & & & & & & & & \\
c_m & & & & & & & & & & & & & & & & \\
r_1 & \bullet & \bullet & \bullet & & & & & & & & & & & & & \\
\vdots & \bullet & \bullet & \bullet & & & & & & & & & & & & & \\
r_l & \bullet & \bullet & \bullet & & & & & & & & & & & & & \\
u_1 & & & & \bullet & \bullet & \bullet & & & & & & & & & & \\
\vdots & & & & \bullet & \bullet & \bullet & & & & & & & & & & \\
u_k & & & & \bullet & \bullet & \bullet & & & & & & & & & & \\
q_1 & & & & \bullet & \bullet & \bullet & & & & & & & & & & \\
\vdots & & & & \bullet & \bullet & \bullet & & & & & & & & & & \\
q_k & & & & \bullet & \bullet & \bullet & & & & & & & & & & \\
t_1 & & & & & & & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & & & & \\
\vdots & & & & & & & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & & & & \\
t_j & & & & & & & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & & & & \\
p & & & & & & & & & & & & & \bullet & \bullet & \bullet &
\end{array} \qquad (15)
$$

$\boldsymbol{A}$ is the lower truncated matrix of size n_p, the number of nodes in the project risk network. The nodes in the project risk network consist of root causes $\boldsymbol{c}$, risk events $\boldsymbol{r}$, resource unit costs $\boldsymbol{u}$ and quantities $\boldsymbol{q}$, tasks $\boldsymbol{t}$, and project p. $\boldsymbol{A}$ has a value of 1 in its (i, j) cell if i is conditionally dependent on j. A single bullet in equation (15) indicates a potential conditional dependency, where risk events $\boldsymbol{r}$ may be conditionally dependent on root causes $\boldsymbol{c}$, resource unit costs $\boldsymbol{u}$ and quantities $\boldsymbol{q}$ may be conditionally dependent on risk events $\boldsymbol{r}$, tasks $\boldsymbol{t}$ are a set of resource unit costs $\boldsymbol{u}$ and quantities $\boldsymbol{q}$, and project p is a set of tasks $\boldsymbol{t}$.

The second step in model development is to specify the conditional probability distributions of the project risk network. From equation (5), given root causes ***c***, the joint distribution $P(\boldsymbol{c})$ is as follows:

$$P(\boldsymbol{c}) = \prod_{j=1}^{m} P(c_j) \tag{16}$$

Similarly, given root causes ***c*** and risk events ***r***, the joint distribution $P(\boldsymbol{r})$ is as follows:

$$P(\boldsymbol{r}) = \prod_{j=1}^{l} P\left(r_j \mid c(r_j)\right) \tag{17}$$

$\boldsymbol{c}(\boldsymbol{r})$ denotes the set of all variables ***c***, such that a conditional dependency exists from root cause *i* to risk event *j* in the project risk network.

Furthermore, given risk events ***r*** and resource costs ***a***, the joint distribution $P(\boldsymbol{a})$ and is as follows:

$$P(\boldsymbol{a}) = \prod_{j=1}^{k} P\left(a_j \mid r(a_j)\right) \tag{18}$$

$\boldsymbol{r}(\boldsymbol{a})$ denotes the set of all variables ***r***, such that a conditional dependency exists from risk event *i* to resource cost *j* in the project risk network.

More specifically, given risk events ***r*** and unit costs ***u*** (quantities ***q***), the joint distributions $P(\boldsymbol{u})$ and $(P(\boldsymbol{q}))$ is as follows:

$$P(\boldsymbol{u}) = \prod_{j=1}^{k} P\left(u_j \mid r(u_j)\right) \tag{19}$$

$$P(\boldsymbol{q}) = \prod_{j=1}^{k} P\left(q_j \mid r(q_j)\right) \tag{20}$$

$\boldsymbol{r}(\boldsymbol{u})$ ($\boldsymbol{r}(\boldsymbol{q})$) denotes the set of all variables $\boldsymbol{r}$, such that a conditional dependency exists from risk event i to unit cost (quantity) j in the project risk network.

Finally, from equation (1), given unit costs $\boldsymbol{u}$, quantities $\boldsymbol{q}$, and task costs $\boldsymbol{t}$, the project cost p is as follows:

$$p = \sum_{i=1}^{j} t_i = \sum_{i=1}^{k} u_i q_i \tag{21}$$

The model inputs defined in equations (16) - (21) are selected from existing project management tools: the work breakdown structure (WBS), resource breakdown structure (RBS), bottom-up cost estimate, and risk register. First, task costs $\boldsymbol{t}$ are selected from the task level of the WBS. In order to specify both task costs $\boldsymbol{t}$ and the conditional dependencies between resources and tasks, the WBS must link to resource requirements (i.e., RBS).

Second, resource unit costs $\boldsymbol{u}$ and quantities $\boldsymbol{q}$ are selected from the lower level of the RBS, and the conditional probabilities in equations (19) and (20) are taken from the bottom-up cost estimate. Thus, the bottom-up cost estimate must be modified in order to specify resource costs in the form of probability distributions. In order to specify the conditional dependencies between risks and resources, moreover, the bottom-up cost estimate must link to risk events (i.e., risk register).

Third, risk events $\boldsymbol{r}$ and root causes $\boldsymbol{c}$ are taken from the risk register. The risk register must be modified to link risk events to resources and to specify the impact on resources as probability distributions. Furthermore, the risk register must specify the joint impact of multiple risk events as probability distributions. The current risk register assumes that risks are additive, meaning the joint impact of multiple risks is simply the summation of the discrete risks. However, if the joint impact of multiple risks is described by probability distributions, then the risk register must be modified.

The result of integrating these common project management tools is a fully specified causal network of project risk. Figure 18 illustrates this integrative approach for model development. The model parameters of the causal network link to different project management tools, while these tools may link as well. The WBS and RBS, for instance, must be integrated in order to specify the resource requirements for each task.

4.2 Model Validation

4.2.1 Model Selection

The model development approach leads to the problem of model selection. The model selection problem may be viewed in two primary ways, which depends on whether or not the data is observed. First, given the prior of the project risk model, select the best set of parameters. Second, given the prior and the observed data, select the best subset of model parameters.

The first problem is prevalent during model development. The number of model parameters is otherwise known as the network granularity. Network granularity is one of the central issues of causal networks (Darwiche, 2009). The research methodology, in principle, applies at any level of network granularity. One approach is to include all of the model parameters. However, following this approach, large projects may lead to models with hundreds of thousands of parameters, and a model of this size may not be practical to develop and/or use. The preferred approach is to choose the model among all competing models with the fewest model parameters.

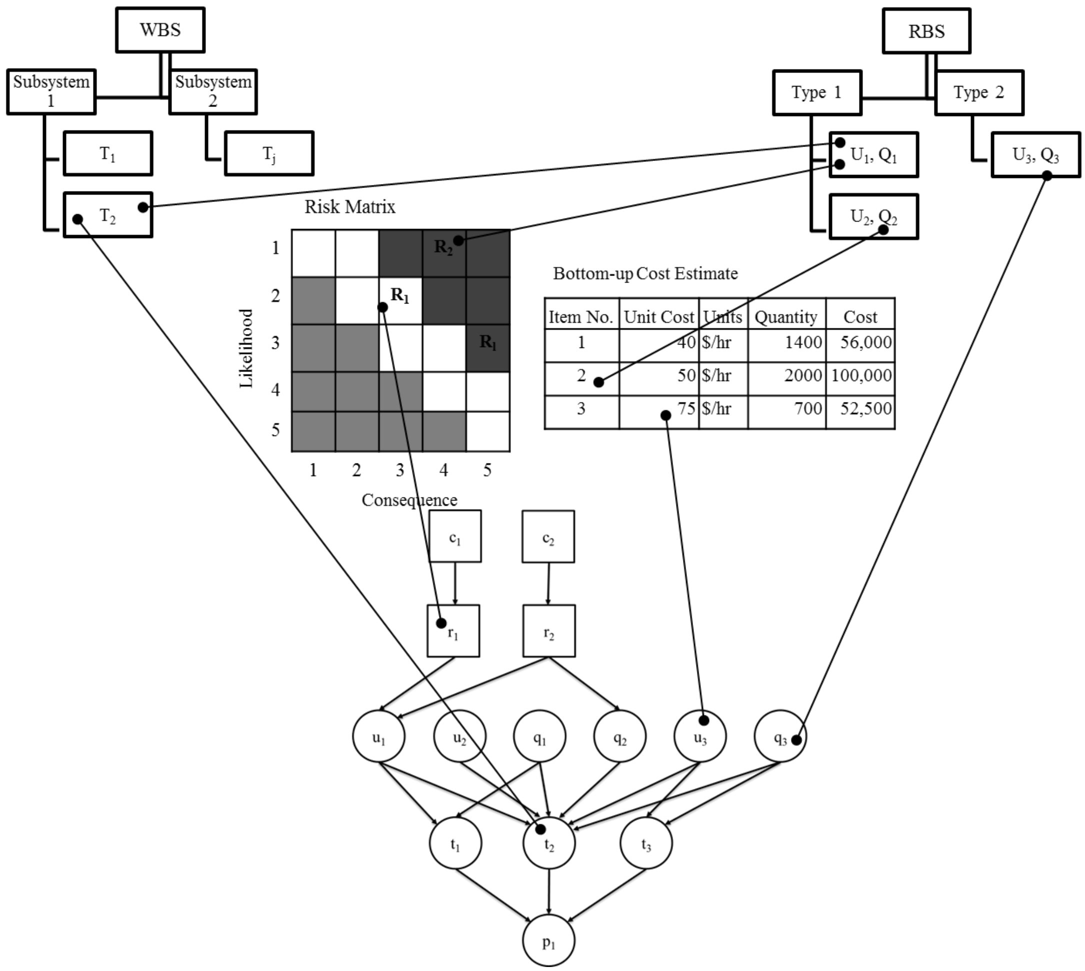


Figure 18: Model Development Approach

Suppose $P(\bullet)$ is the probability distribution of the project risk model with vector of parameters $\boldsymbol{x}$, which follows from (5). $P(\bullet)$ represents the prior distribution of the project risk model. In reality, of course, $P(\bullet)$ is certainly not a “true” prior, since the user undoubtedly specifies model parameters that are based on empirical evidence. Nonetheless, $P(\bullet)$ is a prior distribution in a pragmatic sense, as the model is not updated according to the Bayesian methods. Suppose $P'(\bullet)$ is the probability distribution of the project risk model with vector of parameters $\boldsymbol{x}_p$, where $\boldsymbol{x}_p$ is the subset of $\boldsymbol{x}$ whose elements have indices *p*. In general, $P'(\bullet)$ is "preferred" to $P(\bullet)$ if $P(\bullet)$ and $P'(\bullet)$ agree on every inference formulated. Therefore, the general approach to address the first problem of model selection is as follows:

1. Generate simulated data from the model $y_{1,i} = P(\boldsymbol{x}_i)$ and $y_{2,i} = P'(\boldsymbol{x}_{p,i})$
2. Estimate the empirical distribution functions $F_1(y_1)$ of $y_{1,i}$ and $F_2(y_2)$ of $y_{2,i}$
3. Test $H_0 : F_1(y_1) = F_2(y_2)$, $H_1 : F_1(y_1) \neq F_2(y_2)$ at level α
4. If the null hypothesis is not rejected at level α, select $P'(\bullet)$ as the project risk model

The second problem follows from Bayesian methods, which provide a means of estimating the best subset of model parameters. Suppose x_k is the vector of parameters for project risk model M_k, $P(x_k \mid k)$ is the prior distribution of x_k assuming that M_k is the correct model, p_k is the prior probability of M_k, and $P(y \mid x_k, k)$ is the likelihood of the data $\boldsymbol{y}$. assuming true model is M_k and parameters are x_k. The posterior is as follows:

$$P\left(x_k, k \mid y\right) = \frac{P\left(y \mid x_k, k\right) P\left(x_k \mid k\right) p_k}{m\left(y\right)} \tag{22}$$

Therefore, the general approach to address the second problem of model selection is to select the model that maximizes the posterior probability:

$$P\left(M_k \mid y\right) = \int_{\theta_k} p\left(\theta_k, k \mid y\right) d\theta_k \tag{23}$$

During model development, model parameters should, in principle, be selected for which data is available. In practice, of course, data may be incomplete or unavailable. Therefore, the Bayesian approach to model selection may not always be applicable. However, Bayesian learning may still apply at a lower level, such as the level of the single parameter. Moreover, data may only be available at different phases of the project, as less information is available initially, but over the course of the project more becomes available. Bayesian learning of the project risk model is then more of a dynamic process, which takes place over the course of the project or multiple projects as information becomes available to support the model.

4.2.2 Sensitivity Analysis

As the parameters of the project risk model are uncertain, the response of the model to changes in the parameters may not be apparent. Sensitivity analysis is a study of how the uncertainty in model outputs may be explained by different sources of uncertainty in model inputs. Sensitivity analysis may also explain the practical implications of the project risk model, such as for which risks the model is more or less "sensitive".

Sensitivity analysis is broadly classified as either local or global sensitivity analysis. For a project risk model represented by $\boldsymbol{y} = P(\boldsymbol{x})$, where $\boldsymbol{x}$ is a vector of parameters, local sensitivity analysis involves deriving the derivatives of $P(\bullet)$, evaluated at $\boldsymbol{x} = \boldsymbol{x}_0$ (Oakley & O'Hagan, 2004). The local sensitivity indicates how $P(\bullet)$ changes for a given change in parameters. In practice, however, managers are interested in the range of possible outputs, not a particular perturbation. In this case, global sensitivity analysis assesses changes in $P(\bullet)$ as $\boldsymbol{x}$ varies over all possible values.

Common probabilistic measures of sensitivity are expectation and variance-based methods (Oakley & O'Hagan, 2004). Expectation-based methods quantify the sensitivity of $\boldsymbol{y}$ to the model parameters in terms of a change in the expectation of $\boldsymbol{y}$. Consider the following basic expectation-based measure:

$$z_i = E(\boldsymbol{y} \mid \boldsymbol{x}_i) - E(\boldsymbol{y}) \tag{24}$$

This measure is the expected amount by which the expectation of $\boldsymbol{y}$ will change if one learns the true value of $\boldsymbol{x}_i$.

Variance-based methods quantify the sensitivity of $\boldsymbol{y}$ to the model parameters in terms of a reduction in the variance of $\boldsymbol{y}$. Consider the following variance-based measure:

$$v_i = \mathrm{var}\left(E(\boldsymbol{y} \mid \boldsymbol{x}_i)\right) \tag{25}$$

This measure is the expected amount by which the variance of $\boldsymbol{y}$ will be reduced if one learns the true value of $\boldsymbol{x}_i$.

In addition to explaining the relative importance of model parameters, these sensitivity measures indicate where to direct resources in order to reduce the uncertainty of the project risk model. Suppose risk events $\boldsymbol{r}$ may be observed for a given cost, where the cost of observation is the same for all $\boldsymbol{r}$. All else being equal, a rational person would choose to observe r_i corresponding with the greatest reduction in the variance of the model. In reality, of course, risks are rarely if ever ex-ante observable. Still, these measures support managers by indicating where further research is required.

4.3 Bayesian Inference

Following model validation, the next step in the research methodology is model analysis through Bayesian inference. Bayesian inference is the process of updating beliefs based on evidence of events. If the events represent actual evidence, then Bayesian inference produces updated probability distributions. Conversely, if the events represent hypothetical evidence, then Bayesian inference produces possible scenarios. Thus, Bayesian inference also serves as a scenario analysis, which managers may use to evaluate different management strategies.

Bayesian inference also serves as both a prognosis and diagnosis tool for projects. Bayesian inference serves as a prognosis tool to evaluate possible outcomes of the project and as a diagnosis tool to evaluate possible causes of the project success or failure.

Bayesian inference may be classified as either exact or approximate. Exact Bayesian inference applies to risk events $\boldsymbol{r}$ and resource costs $\boldsymbol{a}$. First, given risk event r and conditionally dependent root causes $\boldsymbol{c}$, the probability distribution $P(r)$ is as follows:

$$P(r) = \sum_{i=1}^{j} P(r \mid c_i) P(c_i) \tag{26}$$

Example 1: Consider the hypothetical project shown in Figure 15. Suppose $P(c_1) = 0.5$, $P(r_1 \mid c_1 = T) = 0.75$, and $P(r_1 \mid r_1 = F) = 0.65$. The prior probability $P(r_1)$ is as follows:

$$P(r_1) = P(r_1 \mid c_1 = T) P(c_1 = T) + P(r_1 \mid c_1 = F) P(c_1 = F)$$

$$P(r_1) = (0.75)(0.5) + (0.65)(0.5) = 0.7$$

Example 2: Consider Figure 15. Suppose $P(c_2) = 0.25$, $P(r_2 \mid c_2 = T) = 0.95$, and $P(r_2 \mid c_2 = F) = 0.75$. The prior probability $P(r_2)$ is as follows:

$$P(r_1) = P(r_2 \mid c_2 = T) P(c_2 = T) + P(r_2 \mid c_2 = F) P(c_2 = F)$$

$$P(r_1) = (0.95)(0.25) + (0.75)(0.75) = 0.8$$

Similarly, given resource cost a and conditionally dependent risk events $\boldsymbol{r}$, the prior probability $P(a)$ is as follows:

$$P(a) = \sum_{i=1}^{j} P(a \mid r_i) P(r_i) \tag{27}$$

More specifically, given resource unit cost (quantity) u (q) and conditionally dependent risk events $\boldsymbol{r}$, the prior probability $P(u)$ ($P(q)$) is as follows:

$$P(u)=\sum_{i=1}^{j}P(u \mid r_i)P(r_i) \tag{28}$$

$$P(q)=\sum_{i=1}^{j}P(q \mid r_i)P(r_i) \tag{29}$$

Example 3: Consider Figure 15. Suppose u_1 is normally distributed and is conditionally dependent on both r_1 and r_2. Suppose $P(u_1 \mid r_1 = T, r_2 = F) = N(10,\ 10)$, $P(u_1 \mid r_1 = F, r_2 = T) = N(20,\ 10)$, $P(u_1 \mid r_1 = T, r_2 = T) = N(25,\ 15)$, and $P(u_1 \mid r_1 = F, r_2 = F) = N(8,\ 5)$. The prior probability $P(u_1)$ is as follows:

$$P(u_1)=P(u_1 \mid r_1 = T, r_2 = F)P(r_1 = T)P(r_2 = F)$$

$$+P(u_1 \mid r_1 = F, r_2 = T)P(r_1 = F)P(r_2 = T)$$

$$+P(u_1 \mid r_1 = \mathrm{T}, r_2 = T)P(r_1 = T)P(r_2 = T)$$

$$+P(r_1 \mid r_1 = F, r_2 = F)P(r_1 = F)P(r_2 = F)$$

$$P(u_1)=(0.24)N(10,10)+(0.14)N(20,10)+(0.56)N(25,15)+(0.06)N(8,5)$$

Approximate Bayesian inference is required for task costs $\boldsymbol{t}$ and project cost p.

Algorithm 1: given task cost t and conditionally dependent resource unit costs $\boldsymbol{u}$ and quantities $\boldsymbol{q}$, an MC algorithm to generate samples $\{t^{(r)}\}_{r=1}^{R}$ from $P(t)$ is as follows:

1. Randomly select samples $\boldsymbol{u}_1$, $\boldsymbol{u}_2$, …, $\boldsymbol{u}_j$ from conditionals $P(u_1)$, $P(u_2)$, …, $P(u_j)$ respectively, where $\boldsymbol{u}_i$ is a vector of length R from $P(u_i)$.

2. Randomly select samples $\boldsymbol{q}_1$, $\boldsymbol{q}_2$, …$\boldsymbol{q}_j$ from conditionals $P(q_1)$, $P(q_2)$, …, $P(q_j)$ respectively, where $\boldsymbol{q}_i$ is a vector of length R from $P(q_i)$.

3. $\left\{t^{(r)}\right\}_{r=1}^{R} = \sum_{i=1}^{j} \boldsymbol{u}_i \boldsymbol{q}_i^T$

Example 4: Consider Figure 15. Given task cost t_1 is conditionally dependent on resource unit cost u_1 and quantity q_1, an MC algorithm to generate samples $\left\{t_1^{(r)}\right\}_{r=1}^{R}$ from $P(t_1)$ follows from ***Algorithm 1***. Figure 19 shows the histogram of t_1 for a sample R of size 1000.

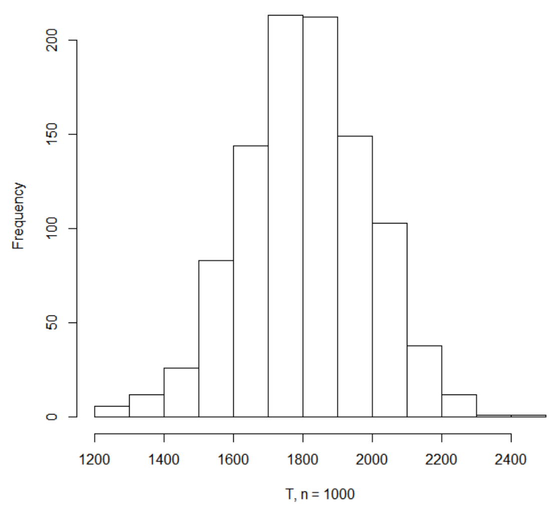


Figure 19: Histogram of T_1

Finally, project managers are concerned with how task costs change relative to one another. As such, the task cost correlation matrix is common project management tool.

Algorithm 2: given task costs ***t***, an MC algorithm to estimate the correlation matrix of the j random variables $t_1, \ldots, t_j$ is as follows:

1. Randomly select samples from $P(t_1), P(t_2), \ldots, P(t_j)$ using ***Algorithm* 1**, and call the samples $\boldsymbol{n}_1, \ldots, \boldsymbol{n}_j$, respectively.
2. Calculate the correlation matrix of the j random samples $\boldsymbol{n}_1, \ldots, \boldsymbol{n}_j$, which is a j * j matrix whose i, j entry is $\mathrm{corr}(\boldsymbol{n}_i, \boldsymbol{n}_j)$.

Example 5: Consider Figure 14. Given task cost t_1 and t_2, an MC algorithm to estimate the correlation of t_1 and t_2 follows from ***Algorithm 2***. For $n_1 = n_2 = 1000$, the estimated correlation of t_1 and t_2 is 0.36.

4.4 Bayesian Learning

While Bayesian inference involves updating beliefs based on evidence of events, Bayesian learning consists of combining beliefs with the observed data. Bayesian learning may be classified as either parameter learning or structural learning. This section presents Bayesian methods for parameter learning of the project risk network. Parameter learning may be further subdivided into discrete and continuous variable learning. Discrete variable learning applies to risk events ***r***.

Given risk event r and conditionally dependent root causes ***c*** that are observed to have the values $c_1, c_2 \ldots, c_j$. The conditional probability of R given the observations is:

$$P\left(r|c_1, c_2, \ldots, c_j\right) = \frac{P(c_1|r)P(c_2|r)\ldots P\left(c_j \mid r\right)P\left(r\right)}{\sum_{v=domain(r)} P(c_1|r)P(c_2|r)\ldots P\left(c_j \mid r\right)P\left(r\right)} \tag{30}$$

Example 6: Consider Figure 15, the general project risk model. Suppose risk event r_1 is conditionally dependent on root cause c_1, and suppose $P(r_1 = T) = 0.25$, $P(c_1 = T \mid r_1 = T) = 0.75$, and $P(c_1 = T \mid r_1 = F) = 0.35$. Now suppose that for a sample of size three, c_1 does not occur (*FFF*). The conditional likelihood $P(c_1 = FFF \mid r_1 = T)$ is simply:

$$P(c_1 = FFF \mid r_1 = T) = 0.25 \times 0.25 \times 0.25 = 0.0156$$

Similarly, $P(c_1 = FFF \mid r_1 = F)$ is as follows:

$$P(c_1 = FFF \mid r_1 = F) = 0.65 \times 0.65 \times 0.65 = 0.2746$$

Seeing this new evidence that c_1 does not occur on three separate occasions, the objective is to revise the posterior probability of r_1 as follows:

$$P(r_1 = T \mid c_1 = FFF) = \frac{P(c_1 = FFF \mid r_1 = T)P(r_1 = T)}{\sum_{v = domain(r_1)} P(c_1 = FFF \mid r_1 = T)P(r_1 = T)}$$

$$P(r_1 = T \mid r_1 = FFF) = \frac{0.0156 \times 0.25}{0.0156 \times 0.25 + 0.2746 \times 0.75} = 0.0186$$

Example 7: Consider Figure 15. Suppose risk event r_2 is conditionally dependent on root cause c_2, and suppose $P(r_2 = T) = 0.5$, $P(c_2 = T \mid r_2 = T) = 0.75$, and $P(c_2 = T \mid r_2 = F) = 0.15$. Now suppose that for a sample of size three, c_2 occurs twice and does not occur once (*TTF*). The conditional likelihood $P(c_2 = TTF \mid r_2 = T)$ is simply:

$$P(c_2 = TTF \mid r_2 = T) = 0.75 \times 0.75 \times 0.25 = 0.1406$$

Similarly, $P(c_2 = TTF \mid r_2 = F)$ is as follows:

$$P(c_2 = TTF \mid r_2 = F) = 0.15 \times 0.15 \times 0.85 = 0.0191$$

The posterior probability of r_2 as follows:

$$P(r_2 = T \mid c_2 = TTF) = \frac{P(c_2 = TTF \mid r_2 = T)P(r_2 = T)}{\sum_{v=domain(r_2)} P(c_2 = TTF \mid r_2 = T)P(r_2 = T)}$$

$$P(r_2 = T \mid c_2 = TTF) = \frac{0.1406 \times 0.5}{0.1406 \times 0.5 + 0.0191 \times 0.5} = 0.8804$$

Continuous variable learning applies to resource costs $\boldsymbol{a}$. Given resource cost a and j conditionally dependent risk events $\boldsymbol{r}$, $\boldsymbol{r}$ may impact a in $j*j$ possible ways. Suppose a has parameters $\boldsymbol{\theta}$, and a_1, a_2, …, a_{j*j} are taken from a. From (4), the posterior distribution is given as follows:

$$P(\boldsymbol{\theta}|\boldsymbol{a}) = \frac{p(\boldsymbol{a},\boldsymbol{\theta})}{m(\boldsymbol{a})} = \frac{p(\boldsymbol{a}|\boldsymbol{\theta})\,p(\boldsymbol{\theta})}{m(\boldsymbol{a})} \propto p(\boldsymbol{a}|\boldsymbol{\theta})\,p(\boldsymbol{\theta}) \qquad (31)$$

Similarly, suppose u (q) has parameters $\boldsymbol{\theta}_u$ ($\boldsymbol{\theta}_q$), and u_1, u_2, …, u_{j*j} are taken from u (q). From (4), the posterior distribution is given as follows:

$$P(\boldsymbol{\theta}|\boldsymbol{u}) = \frac{p(\boldsymbol{u},\boldsymbol{\theta})}{m(\boldsymbol{u})} = \frac{p(\boldsymbol{u}|\boldsymbol{\theta})\,p(\boldsymbol{\theta})}{m(\boldsymbol{u})} \propto p(\boldsymbol{u}|\boldsymbol{\theta})\,p(\boldsymbol{\theta}) \qquad (32)$$

$$P(\boldsymbol{\theta}|\boldsymbol{q}) = \frac{p(\boldsymbol{q},\boldsymbol{\theta})}{m(\boldsymbol{q})} = \frac{p(\boldsymbol{q}|\boldsymbol{\theta})\,p(\boldsymbol{\theta})}{m(\boldsymbol{q})} \propto p(\boldsymbol{q}|\boldsymbol{\theta})\,p(\boldsymbol{\theta}) \qquad (33)$$

Example 8: Consider Figure 15. Suppose resource unit cost u_2 is normally distributed with unknown mean μ and unknown variance σ^2. Recall that the normal distribution has the density:

$$\left(y \mid \mu, \sigma^2\right) = \frac{1}{\sqrt{2\pi\sigma}} exp\left(-\frac{1}{2}\left(\frac{y-\pi}{\sigma}\right)^2\right)$$

Further, suppose 20 random observations are taken from u_2. The following noninformative and improper prior is used:

$$p(\mu, \sigma) \propto \sigma^{-1} I_{(0,\infty)}(\sigma) I_{(-\infty,\infty)}(\mu)$$

The posterior depends on the data only through sufficient statistics, which in this case are:

$$\bar{y} = 22 \text{ and } \sum_{i=1}^{20}\left(y_i - \bar{y}\right)^2 = 325$$

The posterior distribution is given by:

$$p(\mu, \sigma \mid y) \propto \sigma^{-21} \exp\left(-\frac{325}{2\sigma^2}\right) \exp\left[-\frac{20}{2\sigma^2}(\mu - 22)^2\right]$$

The mode of the posterior is:

$$(\hat{\mu}, \hat{\sigma}) = \left(22, \sqrt{\frac{325}{21}}\right) = (22, 3.93)$$

Figure 20 shows the density of u_2 for N = 1000 and $(\hat{\mu}, \hat{\sigma})$ = (22, 3.93).

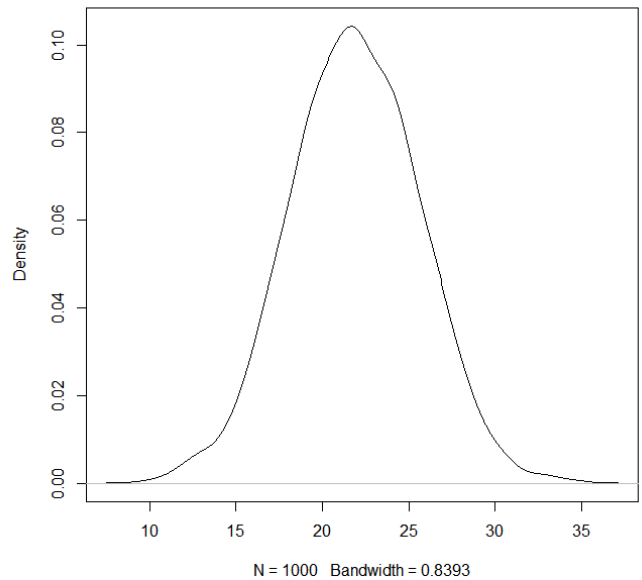


Figure 20: Density of U_2 for $(\hat{\mu}, \hat{\sigma})$ = (22, 3.93)

Example 9: Consider Figure 15. Suppose resource unit cost u_1 is conditionally dependent on risk events r_1 and r_2. Therefore, u_1 may be impacted by r_1 and r_2 in four possible ways: r_1 may occur, r_2 may occur, r_1 and r_2 both may occur, or r_1 and r_2 both may not occur. Suppose u_1 has a multivariate normal distribution with mean vector $\boldsymbol{\theta}$ and covariance matrix $\boldsymbol{\Sigma}$. In this case, u_1 has the density:

$$f\left(Y \mid \theta, \Sigma\right) = \frac{1}{\left(2\pi\right)^{p/2} \left|\Sigma\right|^{1/2}} \exp\left(\frac{1}{2}\left(x-\theta\right)^{T} \Sigma^{-1}\left(x-\theta\right)\right)$$

Further, suppose that (y_1, y_2, y_3, y_4) are measured on each of n = 175, where:

$$y_1 = R_1 \text{ occurs}$$

$$y_2 = R_2 \text{ occurs}$$

$$y_3 = R_1 \text{ and } R_2 \text{ occur}$$

$$y_4 = R_1 \text{ and } R_2 \text{ do not occur}$$

The sufficient statistics are:

$$\left(\bar{y}_1, \bar{y}_2, \bar{y}_3, \bar{y}_4\right) = \left(12, 18, 26, 8\right)$$

$$S^2 = \begin{bmatrix} 925 & 0 & 0 & 0 \\ 0 & 875 & 0 & 0 \\ 0 & 0 & 1050 & 0 \\ 0 & 0 & 0 & 650 \end{bmatrix}$$

The off-diagonal terms of S^2 are zero due to the property of conditional independence. In other words, R_1 and R_2 are conditionally independent. A noninformative prior is the following:

$$p\left(\theta, \Sigma\right) \propto \left|\Sigma\right|^{-(p+1)/2}$$

where p is the number of parameters, which is 4 in this case. The posterior distribution is given by:

$$\Sigma \mid Y \sim \text{inverse-Wishart}\left(174, \left(175 S^2\right)^{-1}\right)$$

$$\theta \mid \Sigma, Y \sim N\left(\bar{y}, \Sigma/175\right)$$

The posterior may be used to generate samples in order to: (1) obtain point estimates of the parameters, (2) construct kernel density approximations of the marginal posteriors

The kernel density estimates for each of the marginal posteriors (N = 10,000) are shown in Figure 21.

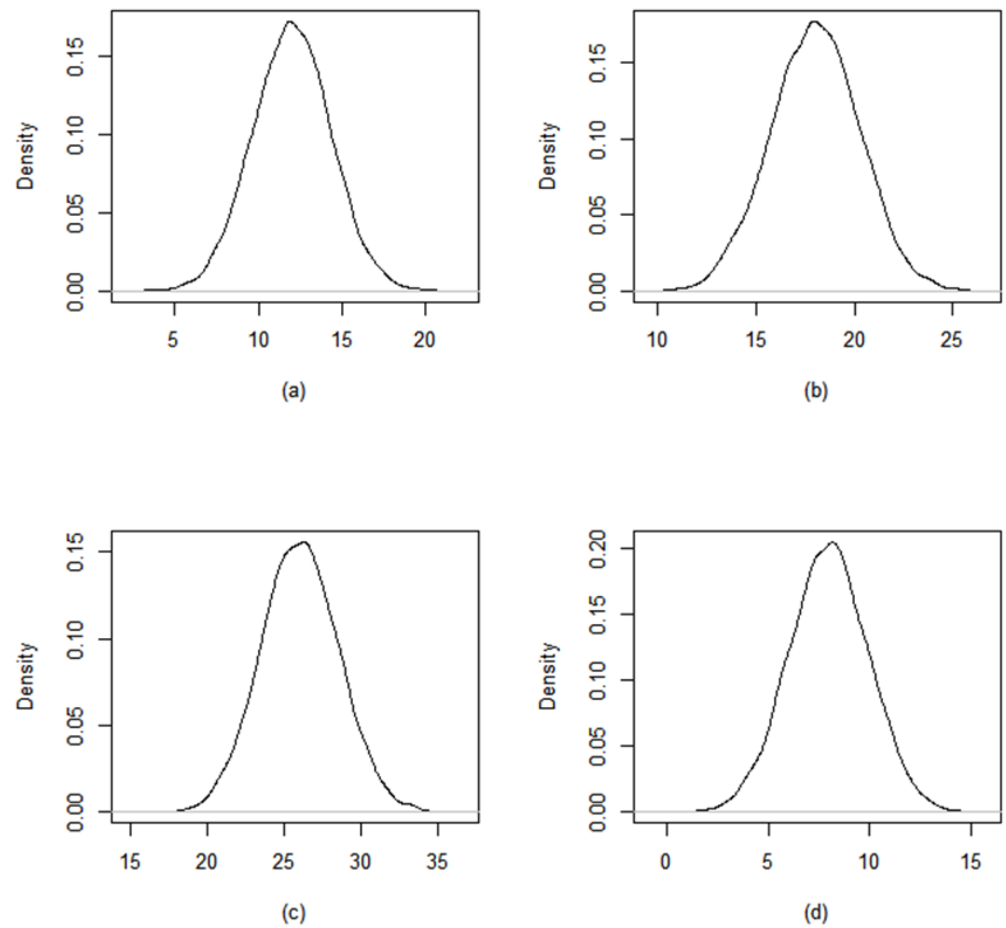


Figure 21: Posterior Density of θ_1 (a), θ_2 (b), θ_3 (c), and θ_4 (d) for N = 10,000

4.5 Summary

This chapter presented the methodology for developing causal networks for project risk analysis. The first section showed that the model may be fully specified by integrating existing project management tools: the WBS, RBS, bottom-up cost estimate, and risk register. The second section demonstrated how techniques from sensitivity analysis test query robustness and query control. The third section showed how Bayesian inference and MC simulation help estimate probability distributions and correlations between the variables. The final section explained how Bayesian methods may be used to learn the model parameters from the observed data. The following chapter illustrates how network measures may also be used for model analysis.

CHAPTER V

NETWORK MEASURES FOR PROJECT RISK ANALYSIS

This chapter presents an overview of network measures for project risk analysis. The network measures explain the dependency structure of projects. Projects may be decomposed into subsystems (i.e., tasks), where tasks share interfaces with other tasks. These interfaces may be identified through risks and resources, which require coordination for project risk analysis.

The network analysis measures have multiple practical implications. First, the analysis measures may support managers in risk-related decision analysis. For example, managers should assess and select mitigation strategies with an accurate understanding of task-related interdependencies. Managers may test strategies by changing parameters or breaking dependencies and observing the effect on the overall network. Second, the measures may help identify gaps between expected task-related dependencies and actual ones. Firms may use these measures as learning tools to help improve communication concerning task-related risks and strategies. Measures such as task modularity help break the project into interrelated parts and facilitate coordination among similar tasks.

The chapter includes three main sections: measures of the general project risk model, task dependency measures, and task modularity measures. The first section proposes a basic matrix representation of the general project risk model. The second section develops two primary types of task dependency measures: risk and resource-related dependency measures. The third section develops two primary types of task modularity measures: degree and path modularity measures.

5.1 Measures of the General Project Risk Model

The measures introduced in this section explain the dependency structure of the project and serve as the framework for the task dependency measures introduced in the following section. The general project risk model may be represented in matrix notation. Figure 22 shows the risk matrix ***R*** for the hypothetical project shown in Figure 15. The risk matrix ***R*** is defined as the $m*n$ matrix where m is the number of resources and n is the number of risk events, and the entry r_{ij} is the value 1 if resource i is conditionally dependent on risk event j and 0 otherwise. For example, in Figure 22, risk event 2 impacts both resource 1 and resource 2. The primary value of the risk matrix is that it captures the dependency structure of the risks-resources, which is one of the first steps of the model development approach.

The values in the risk matrix may be weighted to represent the qualified impact of a risk event on a resource. Instead of a cell value of 1, the risk matrix may be weighted with numbers based on the severity of the risk event. The risk register, for example, often uses the notation: 1 = low impact, 2 = medium impact, 3 = high impact, and the risk matrix may use a similar notation.

	Risk 1	Risk 2
Resource 1	1	1
Resource 2		1
Resource 3		

Figure 22: Risk Matrix *R* for Hypothetical Project

Similarly, Figure 23 shows the resource matrix ***A*** for the hypothetical project shown in Figure 15. The resource matrix ***A*** is the $k*m$ matrix where k is the number of tasks and m is the number of resources, and the entry a_{ij} is the value 1 if task i is affiliated with or depends on resource j and 0 otherwise. The main advantage of the resource matrix is that it captures the dependency structure of the resources-tasks. For example, in Figure 23, resource 1 is required for both task 1 and task 2.

Similar to the risk matrix, the entries in the resource matrix may be weighted to represent the relative resource requirement for each task. The resource matrix may be weighted with numbers based on the amount of a resource required for a task. For example, the resource matrix may use the notation: 1 = small requirement, 2 = medium requirement, 3 = high requirement.

	Resource 1	Resource 2	Resource 3
Task 1	1		
Task 2	1	1	1
Task 3			1

Figure 23: Resource Matrix *A* for Hypothetical Project

Finally, Figure 23 shows the risk-task matrix ***T*** for the hypothetical project shown in Figure 15. The risk-task matrix ***T*** is the matrix product of the risk matrix and the resource matrix as follows:

$$\boldsymbol{T} = \boldsymbol{A}\boldsymbol{R} \tag{34}$$

The risk-task matrix ***T*** is the $k*n$ matrix where k is the number of tasks and n is the number of risk events, and the entry r_{ij} is nonzero if task i is potentially, indirectly impacted by risk event j and 0 otherwise. The cell value indicates the number of ways in which a risk event may impact a task or, in other words, the number of paths between a risk event and a task. The primary value of the risk-task matrix is that it captures the dependency structure of the risks-tasks. For example, one can observe from Figure 15 that risk event 1 impacts task 1 and 2 through resource 1. Similar to the risk matrix, the values in the risk-task matrix may be weighted to represent the qualified impact of a risk event on a task.

	Risk 1	Risk 2
Task 1	1	1
Task 2	1	2
Task 3		

Figure 24: Risk-Task Matrix *T* for Hypothetical Project

Figure 25 illustrates the proposed measures for the hypothetical project shown in Figure 15. The risk matrix ***R*** captures the dependency structure of the three resources and three possible risk events. The resource matrix ***A*** captures the resource requirements for the three project tasks. The product ***AR*** produces the risk-task matrix, where nonzero cells capture the possible indirect impacts that risk events have on tasks. For example, risk event 2 may impact task 2 in two different ways, through unit cost 1 and/or quantity 2.

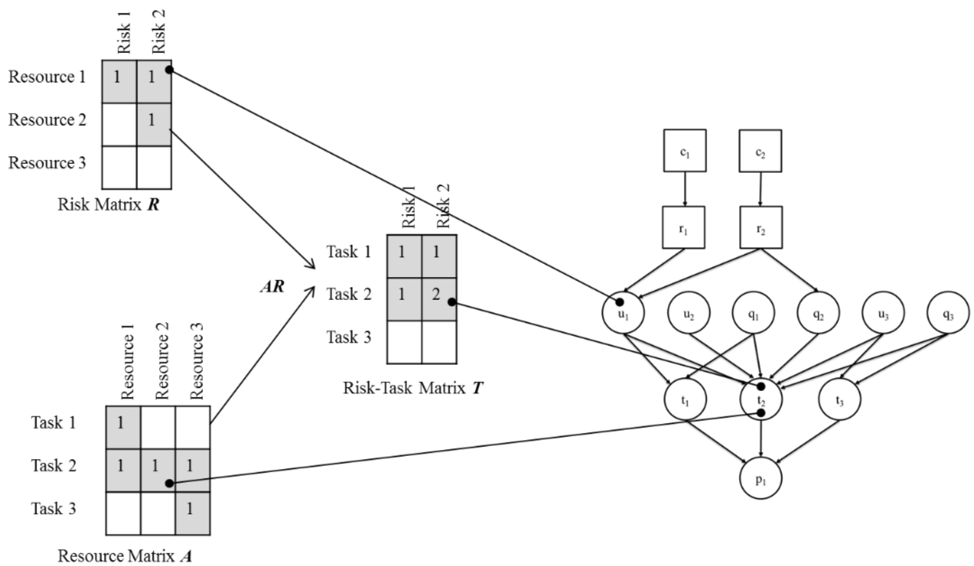


Figure 25: Measures of the General Project Risk Model

The risk matrix ***R***, the resource matrix ***A***, and the risk-task matrix ***T*** have multiple important properties. If the risk matrix is a binary matrix in which $r_{ij} = 1$ if resource i is impacted by risk event j, then the row totals of ***R***, $r_{i+} = \sum_j r_{ij}$ are equal to the number of risk events on which resource i is potentially impacted. Moreover, a row total equal to zero specifies a resource that is not at risk. The risk matrix is an important design tool during the planning phase of projects, as the tool could be used to evaluate different project designs based on risk. The column totals of ***R***, $r_{+j} = \sum_i r_{ij}$ are equal to the total number of resources on which risk event j has an impact. Unlike row totals, every column total logically must be nonzero because a column total equal to zero indicates a risk event that does not have an impact on the overall project.

Similar to the previous case, if the resource matrix $\boldsymbol{A}$ is a binary matrix in which $a_{ij} = 1$ if task i is affiliated with or depends on resource j, then the row totals of $\boldsymbol{A}$, $a_{i+} = \sum_j a_{ij}$ are equal to the number of resources in which task i is potentially dependent. Again, every row total is nonzero because a row total equal to zero indicates a task equal to an empty set. However, a task, by definition must equal some subset of resources. The column totals of $\boldsymbol{A}$, $a_{+j} = \sum_i a_{ij}$ are equal to the total number of tasks in which resource j is required. Finally, every column total must be nonzero because a column total equal to zero indicates a resource that is not required on the overall project.

If the risk-task matrix $\boldsymbol{T}$ is a binary matrix in which $t_{ij} = 1$ if task i is indirectly impacted by risk event j, then the row totals of $\boldsymbol{T}$, $t_{i+} = \sum_j t_{ij}$ are equal to the number of risk events on which task i is potentially impacted. Moreover, a row total equal to zero indicates a task that is not at risk. The column totals of $\boldsymbol{T}$, $t_{+j} = \sum_i t_{ij}$ are equal to the total number of tasks on which risk event j has an impact. Every column total must be nonzero because a column total equal to zero indicates a risk event that does not have an impact on the overall project.

5.2 Task Dependency Measures

Following the measures introduced in the previous section, this section introduces measures of task-related dependencies. The task dependency measures explain the interfaces between tasks, which may be identified through risks and resources. As different tasks are often managed by different groups within the organization, these interfaces require coordination in order to capture the overall impact of risk on the project. This section introduces three network measures of task dependencies: the task parent, grandparent, and resultant dependency matrices.

5.2.1 Task Parent Matrix

The task parent matrix gives the number of resources any two tasks has in common. Two tasks have the same resource (parent) if each has a nonzero value in the same column of *A*, or $a_{ki} = a_{li} = 1$ so that resource *i* is required by both tasks *k* and *l*. The number of times two tasks each have a nonzero value in the same column of ***A*** is the total number of common resources between tasks *k* and *l*. Therefore, the entry p_{kl} in the task-parent matrix is $p_{kl} = \sum_{i=1}^{m} a_{ki} a_{li}$. If tasks *k* and *l* do not have any common resources, then $p_{kl} = 0$, and if all of the resources contribute to tasks *k* and *l*, then $p_{kl} = m$. Therefore, the task parent matrix ***P*** is the matrix product of the resource matrix and the resource matrix transpose as follows:

$$\boldsymbol{P} = \boldsymbol{A}\boldsymbol{A}^{T} \tag{35}$$

The task parent matrix is a square, symmetric matrix of size *k*, where nonzero off-diagonal elements specify the total number of resources that a pair of tasks have in common. In other words, the task parent matrix indicates the number of common resource dependencies. The diagonal elements indicate the total number of resources required for each task. Figure 23 shows the task parent matrix ***P*** for the hypothetical project shown in Figure 15. Tasks 1 and 2 have one resource in common, while task *j* does not have any common resources. Task 2 requires 2 resources, while tasks 1 and 2 each require 1 resource.

	Task 1	Task 2	Task 3
Task 1	1	1	
Task 2	1	2	1
Task 3		1	1

Figure 26: Task Parent Matrix *P* for Hypothetical Project

5.2.2 Task Grandparent Matrix

While the task parent matrix captures the dependency structure between resources and tasks, it does not capture the dependencies between risk events and tasks. The task grandparent matrix gives the number of risk events any two tasks have in common. Two tasks have the same risk event (grandparent) if each has a nonzero value in the same column of $\boldsymbol{T}$, or $t_{ki} = t_{li} = 1$ so that risk event i may impact both tasks k and l. The total number of times two tasks have a nonzero value in the same column of $\boldsymbol{T}$ is the total number of risk events that impact both tasks k and l. Therefore, the entry g_{kl} in the task grandparent matrix is $g_{kl} = \sum_{i=1}^{n} \boldsymbol{T}_{ki}\boldsymbol{T}_{li}$. If tasks k and l do not have any risk events in common, then $g_{kl} = 0$, and if all of the risk events impact these two tasks, then $g_{kl} = n$. Thus, the task grandparent matrix $\boldsymbol{G}$ is the matrix product of the risk-task matrix and the risk-task matrix transpose as follows:

$$\boldsymbol{G} = \boldsymbol{T}\boldsymbol{T}^{T} \tag{36}$$

The task grandparent matrix is a square, symmetric matrix of size k, where nonzero off-diagonal elements specify the total number of risk events that a pair of tasks have in common. In other words, the task grandparent matrix indicates the number of common risk dependencies. The diagonal elements indicate the total number paths with which risk events may impact each task.

Figure 27 illustrates the proposed approach to derive the task grandparent matrix. From Figure 25, the risk-task matrix ***T*** is derived from the risk matrix ***R*** and the resource matrix ***A***. To obtain the task grandparent matrix ***G***, one must multiply the risk-task matrix by its transpose. Again, because binary matrices are used, the off diagonal cells in the task grandparent matrix capture the number of risk events that two tasks share, while diagonal cells specify the number of possible risk paths by which tasks may be impacted. For example, task 2 has 2 risk events in common with task 1 and may be impacted by risk events in three possible ways.

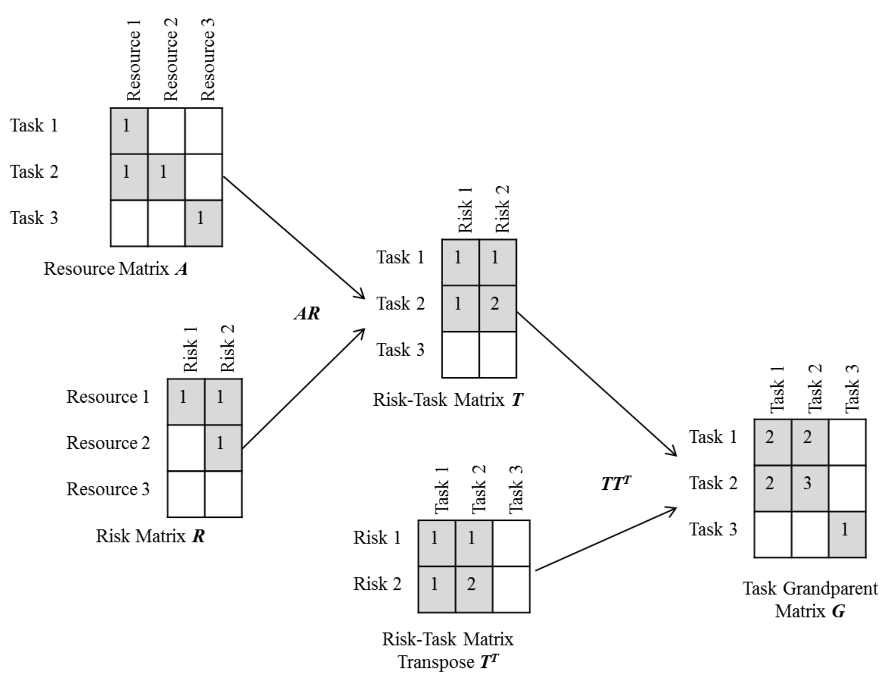


Figure 27: The Approach to Derive the Task Grandparent Matrix

5.2.3 Task Dependency Matrix

The resultant task dependency matrix integrates the task parent and grandparent matrices. The task dependency matrix is a variation of the design structure matrix (DSM) since it describes the structure of project interdependencies. The task dependency matrix is a square, symmetric matrix of size k, where off-diagonal elements with a value of Y indicate a resource-related dependency between a pair of tasks, while a value of Z indicates a risk-related dependency between a pair of tasks. The value of the task dependency matrix is that it provides a concise description of the task-related dependencies and illustrates the overall project design.

The task dependency matrix of the hypothetical project is shown in Figure 28. Not surprisingly, task 1 and task 2 share both resource and risk-related dependencies, while task 2 has a common resource with task 3. Note that the task dependency matrix does not indicate the number of task-related dependencies simply that a dependency does exist. Similar to all of the network measures introduced in this chapter, the task dependency matrix may be weighted to represent either the strength or the number of dependencies.

	Task 1	Task 2	Task 3
Task 1	-	Z	
Task 2	Z	-	Y
Task 3		Y	-

Figure 28: The Task Dependency Matrix for Hypothetical Project

5.3 Task Modularity Measures

Following the task dependency measures introduced in the previous section, this section develops measures of task modularity in the context of risk events and resources instead of requirements or inputs. The task modularity measures explain how relatively connected or disconnected tasks are relative to other tasks in the project. This information is important for project risk management. Highly connected tasks, for example, require a higher degree of coordination than disconnected tasks. This section introduces two measures of task modularity: task degree and path modularity.

5.3.1 Task Degree Modularity

Task degree modularity relates negatively to the number of resources a given task has in common with other tasks. The larger the number of resources task i has in common with other tasks, the less modular task i is. The degree of a task is the number of resources that are required by the task. Task degree modularity has a range from a minimum of 0, where no resources are common between tasks, to k, the total number of resources. The definition of task degree modularity T_D is as follows:

$$T_D = 1 - \frac{p_{i+}}{k} \tag{37}$$

where $p_{i+} = \sum_{j\equiv 1, j\neq i} p_{ij}$. T_D range over [0,1]. The maximum task degree modularity is when a task does not have any resources in common with any other tasks. The minimum task degree modularity occurs when a task has the maximum resource dependencies k with other tasks of the project. Task degree modularity increases linearly as the number of common resources decreases. If no common resource dependencies exist, then the task is completely disconnected from the other tasks, and the task degree modularity is 1.

5.3.2 Task Path Modularity

Task path modularity relates negatively to the number of risk paths a given task has in common with other tasks. The larger the number of risk events that indirectly affect task *i* as well as other tasks, the less modular task *i* is. A risk path is a path by which a risk event may impact a task. Task path modularity ranges from a minimum of 0, where a task has no common risk paths, to $l\left(k^2 - 1\right)$, where l is the number of risk events and k is the number of resources. The definition of task path modularity T_D is as follows:

$$T_P = 1 - \frac{g_{i+}}{l\left(k^2 - 1\right)} \tag{38}$$

where $g_{i+} = \sum_{j\equiv 1, j\neq i} g_{ij}$. T_G range over [0,1]. The maximum task path modularity is when a task does not have any common risk paths with any other tasks. The minimum task path modularity occurs when a task has the maximum risk dependencies $l\left(k^2 - 1\right)$ with other tasks in the project. The value of task path modularity increases linearly as the number of common risk paths decreases. If no common risk dependencies exist, then the task is completely disconnected from the other tasks, and the path modularity is 1.

The proposed measures of task modularity are complementary yet describe distinct types of dependencies. A risk dependency is also a resource dependency, but a resource dependency is not necessarily a risk dependency. To define task modularity, the project risk network must be divided into each task's sub network. In order to illustrate these concepts, Figure 29 shows the task modularity measures for the hypothetical project.

	T_D	T_P
Task 1	0.667	0.875
Task 2	0.333	0.875
Task 3	0.667	1.000

Figure 29: Task Modularity Measures for Hypothetical Project

By inspection, the task 3 sub network is the most modular based on its lack of connectivity, while the task 3 sub network is the most integral based on the relatively high connectivity. Furthermore, since tasks 1 and 3 only have one common resource, they have the highest task degree modularity of 0.667, while task 2 has 2 common resources with a task degree modularity of 0.333. Conversely, since task 3 does not have any common risks events, it has a task path modularity of 1.000, while tasks 1 and 2 each have 1 common risk path with a task path modularity of 0.875.

5.4 Summary

This chapter presented an overview of network measures for project risk analysis. The chapter included three main sections: measures of the general project risk model, task dependency measures, and task modularity measures. In chapter 7, a case study illustrates network measures of project designs and management strategies.

CHAPTER VI

CAUSAL NETWORKS FOR PROJECT PORTFOLIO RISK ANALYSIS

The purpose of this chapter is to extend the causal network methods from the application on the project level to the application on the corporate portfolio level. The chapter includes three primary sections: model development, project dependency measures, and project modularity measures. The first section presents the model development approach, where the general corporate portfolio risk model is comprised of multiple project-level models. The second section develops project dependency measures, which are extensions of task-level measures. Similarly, the third section develops project modularity measures, where these measures are applied to the classic resource allocation problem.

6.1 Model Development

The causal network method for the corporate portfolio is an extension of the project-level model. The corporate portfolio is comprised of multiple projects; therefore, the model of corporate portfolio risk consists of multiple project-level models, where overall corporate portfolio risk is identified and measured through common corporate resources and associated risk events.

Figure 30 illustrates the general model of corporate portfolio risk. The figure shows cost P of a hypothetical portfolio. In this case, the portfolio is comprised of four projects with costs $\boldsymbol{p}$. Each project is, in turn, comprised of unique project tasks with costs $\boldsymbol{t}$, where $t_{i,j}$ specifies the task cost j for project i. The corporate portfolio consists of corporate resources with costs $\boldsymbol{a}$. Each project task utilizes a subset of the corporate resources for completion, where $a_{i,j}$ specifies the resource cost j for project i and the shaded node A is the subset of corporate resources that are common across multiple tasks. Similarly, the corporate portfolio is conditional on risk events $\boldsymbol{r}$. Risk events may be unique to a project or result in a broad impact across multiple projects. The designation $r_{i,j}$ indicates the risk event j for project i, while the shaded node R_i indicates the subset of risk events that are common across multiple projects. Finally risk events may be conditional on root causes $\boldsymbol{c}$, where $c_{i,j}$ indicates root cause j for project i, and shaded node C indicates the subset of root causes that are common across multiple projects.

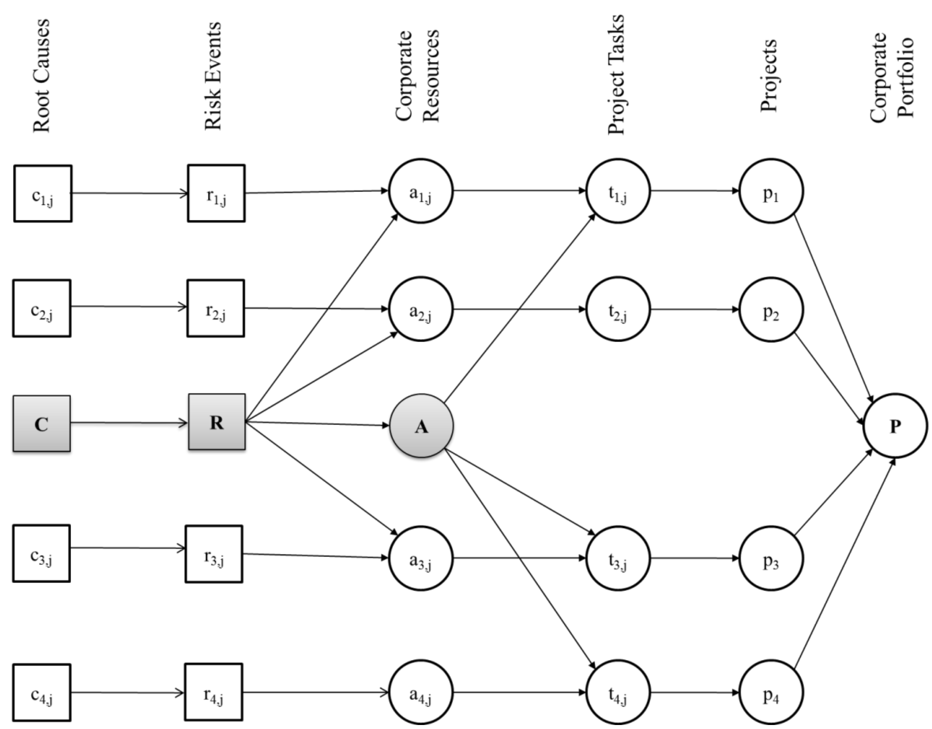


Figure 30: Hypothetical Model of Corporate Portfolio Risk

The model development approach follows from chapter 4. The first step in model development is to specify the directed acyclic graph (DAG) of the corporate portfolio risk network. From equation (15), the DAG may be represented by a lower truncated adjacency matrix ***A***. ***A*** is the square matrix of size n_P, the number of nodes in the corporate portfolio risk network, where the nodes include root causes ***c***, risk events ***r***, corporate resources ***a***, project tasks ***t***, project costs ***p***, and corporate portfolio cost P. ***A*** for the corporate portfolio model is as follows:

$$\boldsymbol{A} = \begin{array}{c|ccccccccccccccccccc}
 & c_1 & \cdots & c_m & r_1 & \cdots & r_l & a_1 & \cdots & a_k & t_1 & \cdots & t_j & p_1 & \cdots & p_i & P \\
\hline
c_1 & & & & & & & & & & & & & & & & \\
\vdots & & & & & & & & & & & & & & & & \\
c_m & & & & & & & & & & & & & & & & \\
r_1 & \bullet & \bullet & \bullet & & & & & & & & & & & & & \\
\vdots & \bullet & \bullet & \bullet & & & & & & & & & & & & & \\
r_l & \bullet & \bullet & \bullet & & & & & & & & & & & & & \\
a_1 & & & & \bullet & \bullet & \bullet & & & & & & & & & & \\
\vdots & & & & \bullet & \bullet & \bullet & & & & & & & & & & \\
a_k & & & & \bullet & \bullet & \bullet & & & & & & & & & & \\
t_1 & & & & \bullet & \bullet & \bullet & & & & & & & & & & \\
\vdots & & & & \bullet & \bullet & \bullet & & & & & & & & & & \\
t_j & & & & \bullet & \bullet & \bullet & & & & & & & & & & \\
p_1 & & & & & & & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & & & & \\
\vdots & & & & & & & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & & & & \\
p_i & & & & & & & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & & & & \\
P & & & & & & & & & & & & & \bullet & \bullet & \bullet &
\end{array} \tag{39}$$

The second step in model development is to specify the conditional probability distributions of the corporate portfolio risk network. The conditional probability distributions of root causes $\boldsymbol{c}$, risk events $\boldsymbol{r}$, and corporate resource costs $\boldsymbol{a}$ follow from equations (16), (17), and (18), respectively. Similarly, the project task costs $\boldsymbol{t}$ follows from equation (21). These model inputs are selected from the project-level models, which are combined for the corporate portfolio. The result of model development is the causal network model of corporate portfolio risk. This model may be used in the same way as the project level model, to analyze corporate risk and evaluate different risk mitigation strategies.

6.2 Project Dependency Measures

The general research methodology involves causal network methods for corporate portfolio risk analysis. However, in practice, data is often not available for model selection and validation. This problem is especially prevalent at the corporate portfolio level, as additional data is required for the complex organization-environment of the corporate portfolio. The motivation for the following network measures is to provide an explanation of the dependency structure of the corporate portfolio in the absence of sufficient data. Network measures provide important information about the corporate portfolio and the interdependencies between projects by analyzing the structure of the corporate portfolio model.

Project network measures follow directly from task network measures. The objective of these measures is to support the management of interdependencies between projects, the coordination of multiple projects, and the management of resources and constraints (Olsson, 2008). Project network measures provide a means to identify and assess interdependencies between projects due to shared risk and resource-related dependencies, and support corporate resource allocation in a way that balances risk and strategic objectives. These measures serve as a framework to guide decision making and overall risk management. If firms then want to move to a formal causal network analysis, the theoretical and methodological framework is in place.

6.2.1 Project Parent Matrix

This section introduces the project parent matrix, which follows directly from the task parent matrix. The project parent matrix is a measure of the *common corporate resource dependencies between projects*. From equation (35), the project parent matrix ***P*** gives the number of corporate resources any pair of projects has in common. The project parent matrix is a square, symmetric matrix of size *i*, the number of projects in the portfolio, where nonzero off-diagonal elements specify the number of corporate resources that a pair of projects have in common or, in other words, the number of common resource dependencies. The diagonal elements indicate the total number of corporate resources required for each project.

Figure 31 presents the project parent matrix for the hypothetical corporate portfolio shown in Figure 30. In this case, the project parent matrix includes the four projects in the portfolio. Projects 1, 3, and 4 each have one common corporate resource; therefore each project is equally connected due to resource dependencies. Project 2 does not have any corporate resources in common and is thus disconnected from the other projects in terms of resource requirements. Similarly, projects 1, 3, and 4 each require 2 resources, while project 2 requires 1 resource.

	Project 1	Project 2	Project 3	Project 4
Project 1	2		1	1
Project 2		1		
Project 3	1		2	1
Project 4	1		1	2

Figure 31: Project Parent Matrix for Hypothetical Corporate Portfolio

The project parent matrix provides a measure of the connectivity of projects due to resources, and thus a measure of resource constraints. This information is important when assessing different management strategies. For example, when assessing contract strategies, a project with a high number of shared corporate resources may be difficult to transfer (i.e. subcontract), while a project with a low number of corporate resource dependencies may be a better candidate. A project with a relatively high number of shared corporate resources indicates a project which requires coordination at the portfolio level. Conversely, a project with a relatively low number of shared resources implies a fairly modular project, one which is relatively disconnected from the other projects in the portfolio.

6.2.2 Project Grandparent Matrix

This section introduces the project grandparent matrix, which follows directly from the task grandparent matrix. The project grandparent matrix is a measure of the *common corporate risk dependencies between projects*. As risk propagates through the corporate resources, firms have common risk dependencies if and only if they have common resource dependencies. Similar to the previous measure, the project grandparent matrix provides an indication of the connectivity of projects due to risk events. A project with a relatively high number of common corporate risk events may indicate a high priority area, one where greater coordination is required for risk management. Conversely, a project with a relatively low number of common risk events may indicate a low priority area, one where less oversight is required.

From equation (36), the project grandparent matrix gives the number of risk events any pair of projects have in common. The project grandparent matrix $\boldsymbol{G}$ is a square, symmetric matrix of size i, the number of projects in the portfolio, where nonzero off-diagonal elements specify the number of risk events that a pair of projects has in common or, in other words, the number of common risk dependencies. The diagonal elements indicate the total number of paths with which risk events may impact each project.

Figure 32 presents the project grandparent matrix for the hypothetical corporate portfolio shown in Figure 30. The project grandparent matrix includes four projects in the portfolio. Projects 1, 2, and 3 each have 1 risk path in common, while project 4 does not have any common risk paths. Therefore, projects 1, 2, and 3 are relatively more connected than project 4 due to risk events. Similarly, projects 1, 2, and 3 each have 2 total risk paths, while project 4 has one.

	Project 1	Project 2	Project 3	Project 4
Project 1	2	1	1	
Project 2	1	2	1	
Project 3	1	1	2	
Project 4				1

Figure 32: Project Grandparent Matrix for Hypothetical Corporate Portfolio

6.2.3 Project Dependency Matrix

The resultant project dependency matrix integrates the project parent and grandparent matrices into one measure of project-related dependencies. The motivation for this measure is to provide a concise explanation of the corporate portfolio dependency structure and to illustrate the overall portfolio design.

The project dependency matrix is a square, symmetric matrix of size *i*, the number of projects in the portfolio, where off-diagonal elements with a value of *Y* indicate a resource-related dependency between a pair of projects, while a value of *Z* indicates a risk-related dependency between a pair of projects.

Figure 33 shows the project dependency matrix for the hypothetical corporate portfolio shown in Figure 33. Of the four projects in the portfolio, projects 1, 2, and 3 each have a common risk path, while project 4 has a common resource. As such, project 1, 2, and 3 seem to be the most connected based on common risk paths, while project 4 is disconnected. Similarly, projects 1, 3, and 4 seem to be the most connected based on common resources, while project 2 seems to be the most disconnected.

	Project 1	Project 2	Project 3	Project 4
Project 1	-	Z	Z	Y
Project 2	Z	-	Z	
Project 3	Z	Z	-	Y
Project 4	Y		Y	-

Figure 33: Project Dependency Matrix for Hypothetical Corporate Portfolio

6.3 Project Modularity Measures

6.3.1 Project Degree Modularity

Following the project dependency measures introduced in the previous section, this section develops measures of project modularity. Project modularity measures follow directly from task modularity measures. Project degree modularity is a measure of the *relative disconnectedness of projects due to the absence of resource dependencies*. A project with relatively high degree modularity shares a low number of resource dependencies and indicates a project that is relatively disconnected from the other projects in terms of resource dependencies. Conversely, a project with relatively low degree modularity shares a high number resource dependencies and indicates a project that is relatively connected with the other projects.

Project degree modularity relates negatively to the number of resources with which a given project shares with other projects. From equation (37), The definition of project degree modularity P_D is as follows:

$$P_D = 1 - \frac{p_{i+}}{k} \tag{40}$$

where $p_{i+} = \sum_{j=1, j \neq i} p_{ij}$. P_D range over [0,1]. The maximum project degree modularity is when a project does not share any corporate resources with any other projects. The minimum project degree modularity occurs when a project shares the maximum resource dependencies k with other projects in the portfolio.

6.3.2 Project Path Modularity

Project path modularity is a measure of the *relative disconnectedness of projects due to the absence of shared risk dependencies*. A project with a relatively high path modularity shares a low number of risk events and indicates a project that is relatively disconnected from the other projects in terms of risk events. Conversely, a project with a relatively low degree modularity shares a high number risk events and indicates a project that is relatively connected with the other projects.

Project path modularity relates negatively to the number of risk paths which a given project has in common with other projects. From equation (38), the definition of project path modularity P_D is as follows:

$$P_P = 1 - \frac{g_{i+}}{l\left(k^2 - 1\right)} \tag{41}$$

where $g_{i+} = \sum_{j \equiv 1, j \neq i} g_{ij}$. P_G range over [0,1]. The maximum project path modularity is when a project does not share any risk paths with any other projects. The minimum project path modularity occurs when a project shares the maximum risk dependencies $l\left(k^2 - 1\right)$ with other projects in the portfolio.

In order to illustrate these concepts, Figure 34 the shows modularity measures for the hypothetical portfolio. By design, project 2 is disconnected from the other projects in terms of common resources. Thus, project 2 has the highest project degree modularity of 1. Projects 1, 3, and 4 each have the maximum number of common resources and thus have the same project degree modularity of 0.000. Conversely, project 4 is disconnected from the other tasks in terms of common risk events. Therefore, project 4 has a project path modularity of 1. Projects 1, 2, and 3 each have the maximum common risk events, and thus have the same project path modularity of 0. From a risk perspective, projects 1, 2, and 3 are all integral, while project 4 is modular. From a resource perspective, projects 1, 3, and 4 are all integral, while project 2 is modular.

	P_D	P_P
Project 1	0.000	0.000
Project 2	1.000	0.000
Project 3	0.000	0.000
Project 4	0.000	1.000

Figure 34: Project Modularity Measures for Hypothetical Corporate Portfolio

6.3.3 An Application to the Project Design of Corporate Resources

Following the project modularity measures introduced in the previous sections, this section illustrates the practical implications of project modularity in an application to corporate resource allocation. The objective is to develop a decision support model of corporate resource allocation for a portfolio of interdependent projects. The decision support model is a linear program model that optimizes corporate resource allocation by balancing risk and resource constraints. The model is flexible and adaptable in order to accommodate changes in the corporate portfolio as well as the environment in which it operates.

The optimization model is a linear program, where the goal is to minimize risk dependencies between projects subject to resource constraints. The model utilizes the project path modularity measure to estimate each project's relative connectivity due to risk dependencies. The model solution offers a method to evaluate alternative mixes of corporate resources, where changing the resource constraints has an immediate effect on the model solution.

In addition to the project path modularity measure, several additional inputs are required. The model inputs include the number of projects in the corporate portfolio N, the number of corporate risk events l, the total number of corporate resources available k, and the project grandparent matrix $\boldsymbol{G}$.

Each cell $g_{i,j}$ in the project grandparent matrix $\boldsymbol{G}$ is subject to a constraint, as some risks cannot be eliminated and some common risk dependencies cannot be broken due to resource constraints. Therefore, the entry $g_{i,j}$ is subject to the following constraint, where $a_{i,j}$ is the minimum number of common risk dependencies for entry $g_{i,j}$:

$$g_{i,j} \geq a_{i,j} \tag{42}$$

The program model makes the simplifying assumption that the cost of changing a common risk dependency is constant for all risks. While this assumption is probably not realistic, a cost function may be added to the model to account for this difference.

As the goal of the model is to minimize the risk dependencies between projects, the model objective function is to maximize the project path modularity. The model objective function is as follows:

$$\text{Maximize } Z = \sum_{i=1}^{N} w_i P_{p,i} = \sum_{i=1}^{N} \left[w_i \left(1 - \frac{g_{i+}}{l\left(k^2 - 1\right)} \right) \right] \tag{43}$$

where w_i is the weighting function for project i and $\sum_{i=1}^{N} w_i = 1$. The model objective function selects the optimal combination of common risk dependencies in order to maximize the total project path modularity subject to specified resource constraints. The model output includes the specific mix of risk dependencies between projects, which may be linked to an alternate combination of corporate resources.

6.4 Summary

This chapter presented the methodology for developing causal networks for project portfolio risk analysis. The chapter extended the causal network methods from the project-level to the portfolio-level in order to develop a project portfolio risk analysis methodology. The following chapter integrates the concepts described in the previous chapters by illustrating the methodology with a case study of a natural gas pipeline project.

CHAPTER VII

CASE STUDY

This chapter presents a case study in order to illustrate the general research methodology. The case study presents causal network methods for project risk analysis of a compressor station project. The chapter is divided into nine sections. The first section describes the case study project. The second section provides the methodology for the case study. The third section presents the model input for the project risk network. The fourth section describes the model development approach for the case study. The fifth section describes the model validation approach. The sixth section illustrates Bayesian inference for project risk analysis. The seventh section illustrates Bayesian learning of the project risk model parameters. The eighth section shows task dependency measures of the project risk model. The ninth section shows task modularity measures of the project risk model.

7.1 Description of the Case Study

The case study project consists of a compressor station and pipeline for a company, hereafter referred to as the Owner, which owns and operates natural gas pipelines among other energy transfer-related assets. The purpose of the project is to increase the minimum delivery pressure to an energy center in order to meet the high demand of the client, while also maintaining and improving the performance of the system. The objectives of the project are to increase the overall revenue for the Owner and to improve the Owner's position for another phase expansion in the future. The overall project scope includes a new compressor station, replacement piping, switch gear, and gas coolers. The total project cost estimate is approximately $67 million and the total project schedule is approximately 27 months.

Figure 35 shows the milestone schedule for the case study. The first phase of the project is feasibility, which involves setting the project objectives and scope of work. The feasibility phase starts in July 2010 and ends in August of the same year. Following a review of the feasibility phase, the second phase is planning, which consists of selecting a preliminary design. The planning phase starts in September 2010 and ends in October of the same year. Following a review of the planning phase, the third phase is engineering, which includes creating the construction plans and specifications. The engineering phase starts in January 2011 and ends in July 2011. The fourth phase, construction, begins in May 2010, during the latter half of the engineering phase, and ends in December 2012. The construction phase lasts over 18 months or approximately 67% of the total project duration. Finally, following the construction phase, the fifth phase is commissioning & start-up, which involves testing the facility for use. The commissioning phase starts in December 2011 and ends in March 2012.

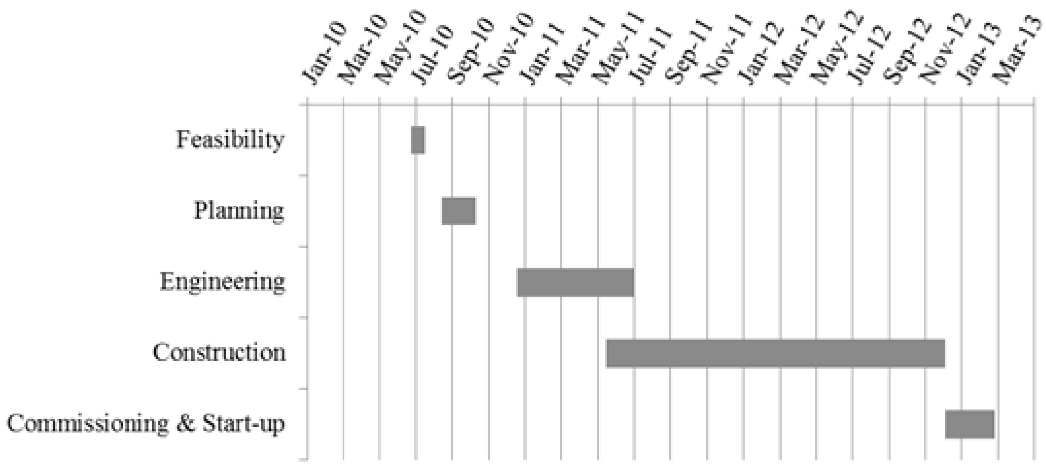


Figure 35: Milestone Schedule for Case Study

Figure 36 shows the general construction plan for the compressor station. The project site currently contains a compressor building (top-center), gas coolers (center-right), and related piping and equipment. The only access to and from the site is an unpaved road located in the top-left of the figure. The proposed compressor station and associated equipment are indicated in the bottom-left of the figure. The construction plan also includes additional equipment located throughout the site such as replacement piping, and the gas coolers (bottom-center).

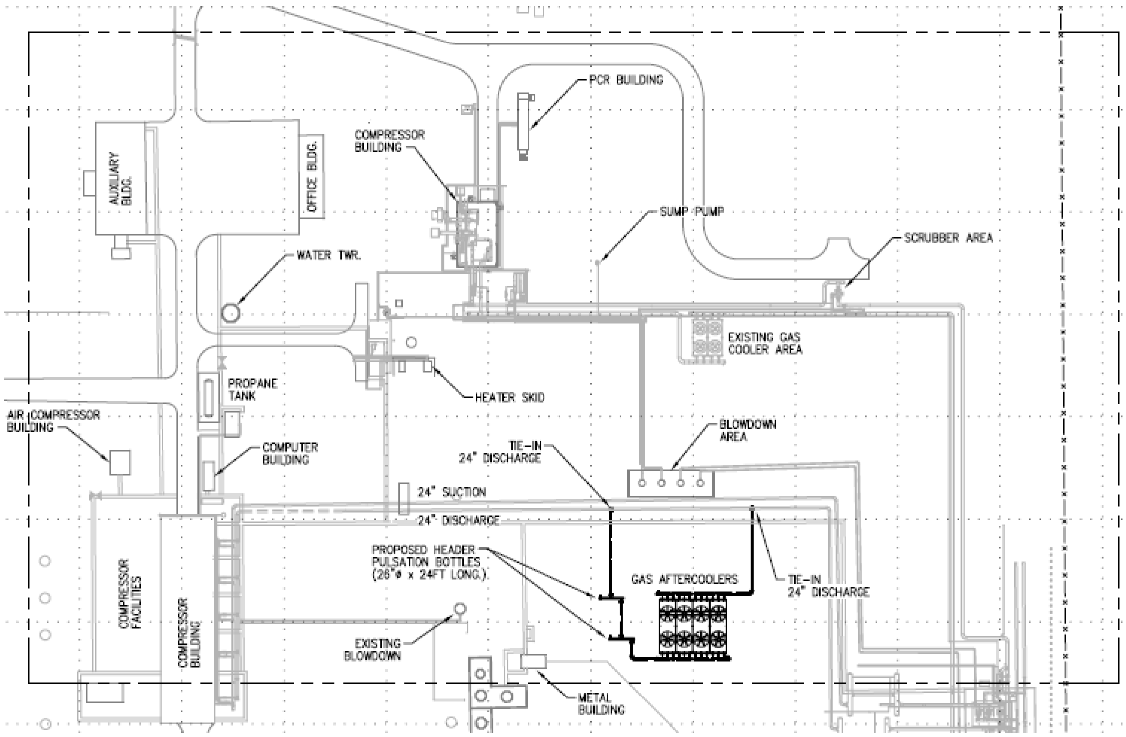


Figure 36: Compressor Station Construction Plan

Figure 37 shows a satellite view of the approximate location of the site, which is indicated by an arrow. This site has a number of identified advantages and disadvantages. As far as advantages, the site is located near a sufficient electrical supply, has few residents in the surrounding area, has relatively less wetlands than the nearby areas, and costs less than competing sites. The disadvantages include insufficient access to the site, a landowner who is unwilling to sell the right of way (ROW), and the likelihood of protected species in the area.

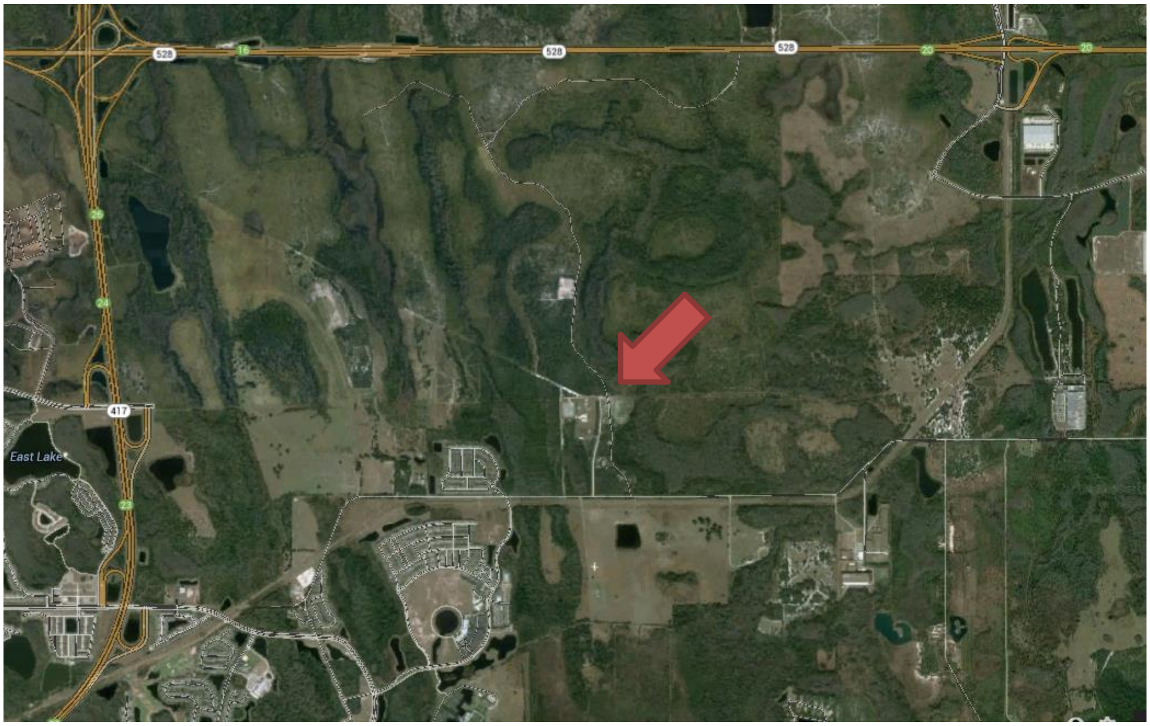


Figure 37: Satellite View of Project Site

7.2 Methodology for the Case Study

The purpose of the case study is threefold: (1) to illustrate the general research methodology, (2) to provide empirical support for the theoretical framework, (3) and to evaluate the methodological framework for practical implications. The case study is a retrospective overview of the planning phase of the compressor station project, with illustrated examples of the causal network methods for project risk analysis. The case study was selected in large part due to the availability of project information required for the project risk network as well as the author's experience with the project.

The overall approach for the case study follows from the general research methodology. First, the model inputs are identified. Second, the model is developed by specifying the model parameters and conditional dependencies. Third, the model is validated using sensitivity analysis. Fourth, the model is analyzed through Bayesian inference. Fifth, the model parameters are updated through Bayesian learning. The model input is collected from project documents, and prior distributions are assumed based on the available project information.

7.3 Model Input

The work breakdown structure (WBS) for the project includes four primary levels. Level 2 includes two main tasks: front-end loading and execution. The front-end loading level 3 tasks are all internal tasks for the Owner, with the exception of the engineering design task, which is to be completed through a contract with an engineering firm. Most of the execution level 3 tasks are also internal tasks for the owner, including the regulatory approvals, procurement, bidding, commissioning & start-up, and service tasks. The engineering design contract continues from the front-end loading phase to the execution phase, and the construction is to be completed through a contract with a construction management firm. The work breakdown structure (WBS) for the project is shown in Figure 38.

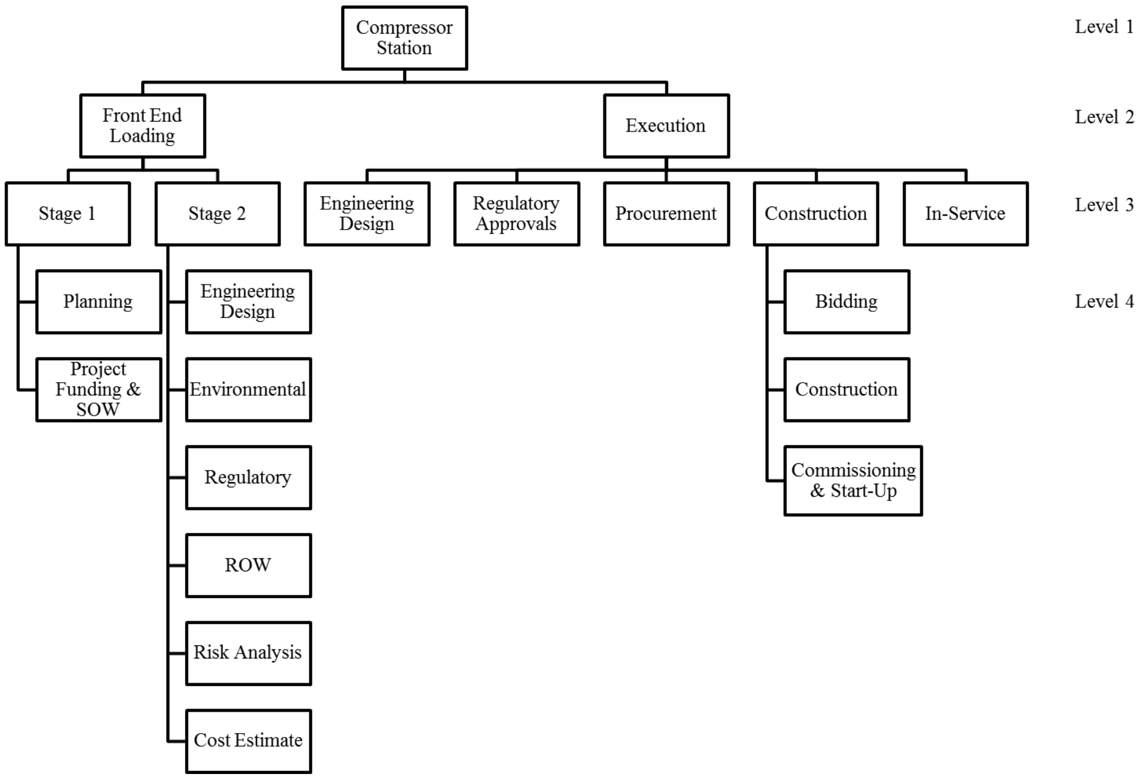


Figure 38: WBS for Case Study

The resource breakdown structure (RBS) for the project includes five primary levels. Level 2 includes three main resource categories: manpower, materials, and monetary resources. Most of the manpower-related resources are internal resources for the Owner, with the exception of the engineering, construction management, and labor categories, which are to be provided through separate contracts. Likewise, most of the material-related resources are internal resources for the Owner, with the exception of certain additional materials, which are to be purchased by the construction manager. The RBS for the project is shown in Figure 39.

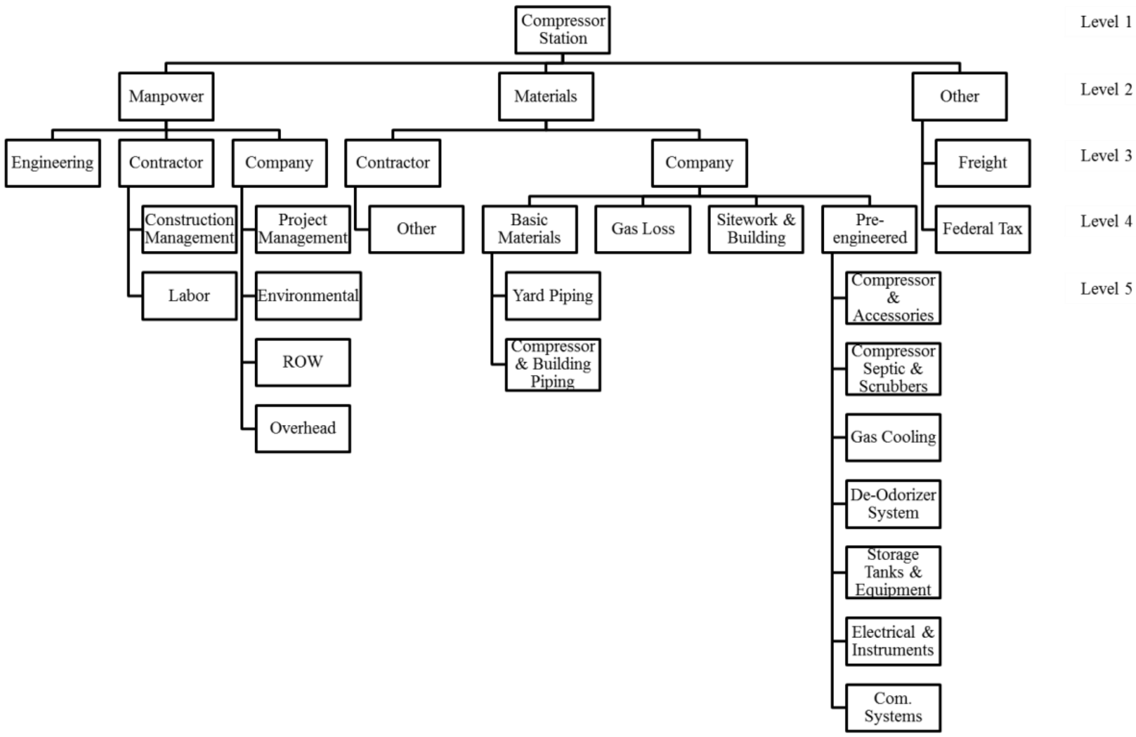


Figure 39: RBS for Case Study

The bottom-up project cost estimate is developed from experience with previous projects, vendor quotes, or other predictive models. Table 1 shows the bottom-up project cost estimate. The project includes 21 general resource categories; 13 lump sum cost ***a*** variables (i.e., subcontracts or supplier quotes), and 8 resource unit cost ***u*** and quantity ***q*** variables. The category with the greatest estimated cost is tax at $15,000,000. As tax is set by the federal government and does not tend to vary, this resource category is not a major risk. The second most costly category, however, is the compressor station and accessories at $11,500,000. One of the identified major risks is, due to a compressed schedule, the originally selected supplier will not be able to deliver the compressor.

Thus, in this situation the Owner would be forced to make a decision to (a) delay the project until the compressor is delivered or (b) select a new supplier, probably at a premium. The third most costly category is labor at $6,300,000. The cost of labor is also at risk due to a delivery delay and most other delays for that matter, as the Owner is required to pay for delays to the project schedule.

Table 1: Bottom-Up Project Cost Estimate for Case Study

Resource	Description	Unit Cost	Unit	Quantity	Total
a_1	Labor	100	hours	63,000	6,300,000
a_2	Freight		lump		2,500,000
a_3	Engineering		lump		1,830,000
a_4	Project Management	275	hours	11,640	3,200,000
a_5	Construction Management		lump		1,500,000
a_6	Environmental	175	hours	4,600	800,000
a_7	ROW	225	hours	20,000	4,500,000
a_8	Construction Materials		lump		1,400,000
a_9	Yard Piping	125	feet	4,400	550,000
a_{10}	Compressor & Building Piping	150	feet	14,000	2,100,000
a_{11}	Gas Loss	15	MCF	9,900	150,000
a_{12}	Compressor & Accessories		lump		11,500,000
a_{13}	Compressor Septic & Scrubbers		lump		650,000
a_{14}	Gas Cooling		lump		1,300,000
a_{15}	De-Odorizer System		lump		75,000
a_{16}	Storage Tanks & Equipment		lump		65,000
a_{17}	Sitework & Building		lump		2,300,000
a_{18}	Electrical & Instruments		lump		3,800,000
a_{19}	Communications		lump		20,000
a_{20}	Overhead	200	hours	20,000	4,000,000
a_{21}	Tax		lump		15,000,000
	Total				66,940,000

The identified major risks are documented in a risk register, which is shown in Table 2. The risk register includes 11 major risks as well as other relevant information such as a description of the risk, root cause, likelihood, impact, strategy, and risk category, where the risk category is linked to the resource categories shown in Table 1. The identified risk with the greatest impact is r_{10}, the risk that the landowner refuses to sell the ROW for the project. Currently, the landowner refuses to sell the ROW, which indicates that the Owner may be forced to pay significantly more than originally estimated. The most likely risk, on the other hand, is r_{11}, the risk of inadequate access to the site. Poor access to the site could result in project delays, which, in turn, could cause additional costs. The current risk mitigation strategy is to reduce the risk by expanding the current road. This strategy, of course, results in additional cost, but the Owner is willing to accept the additional cost in order to lower the risk of schedule delays.

Table 2: Risk Register for Case Study

Risk	Title	Description	Root Cause	Likelihood	Impact	Strategy	Category
r_1	Agency certificate	Delayed agency certificate	High demand for certificates	Possible	Significant	Accept or reduce	Construction management & labor
r_2	Environmental Inspector	Required environmental inspector		Negligible	Significant	Accept	Environmental
r_3	Piles	Required foundation piles	Low soil capacity	Negligible	Major	Accept or reduce	Sitework & building
r_4	Vendor Prices	Higher vendor prices	Equipment shortages	Possible	Significant	Accept	Compressor, scrubbers, & cooling
r_5	Endangered Species	Found endangered species		Negligible	Significant	Accept	Environmental
r_6	Cultural Resources	Found cultural resources		Negligible	Major	Avoid	Environmental

Table 2 Continued

Risk	Title	Description	Root Cause	Likelihood	Impact	Strategy	Category
r_7	Gas Price	Higher gas price	Gas shortage	Possible	Low	Accept	Gas loss
r_8	Weather	Inclement weather	Rain	Possible	Low	Accept	Construction management & labor
r_9	Compressor Station	Deliver delay of compressor station	High demand	Possible	Significant	Reduce	Construction management & labor
r_{10}	ROW	Resisted sale or ROW	Poor communication with landowner	Likely	Major	Accept	ROW
r_{11}	Access	Insufficient access to site	High traffic volume	Likely	Low	Reduce	Construction management & labor

7.4 Model Development

Now that the model inputs are specified, the next step in the research methodology is model development. The first step in model development is to specify the directed acyclic graph (DAG) of the project risk network. From equation (15), the DAG may be represented by the lower truncated adjacency matrix $\boldsymbol{A}$. Figure 40 shows the adjacency matrix $\boldsymbol{A}$ for the case study. $\boldsymbol{A}$ is a square matrix of size 56, which represents the number of nodes in the project risk network. The network contains 8 root causes $\boldsymbol{c}$, 11 risk events $\boldsymbol{r}$, 21 resource costs $\boldsymbol{a}$, 15 task costs $\boldsymbol{t}$ and project cost p. The conditional dependencies between $\boldsymbol{r}$ and $\boldsymbol{c}$ and $\boldsymbol{a}$ and $\boldsymbol{r}$ are both indicated in the risk register.

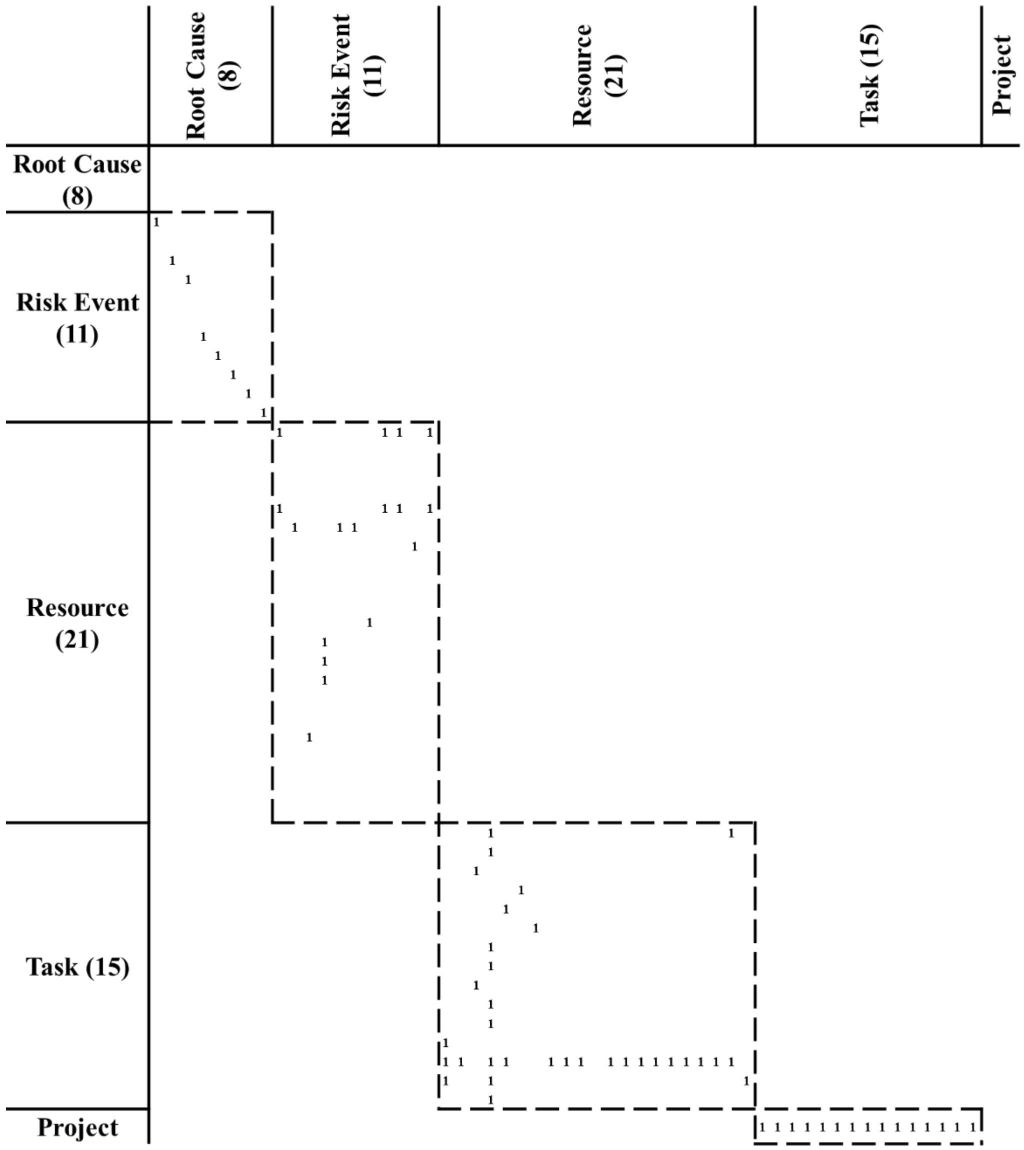


Figure 40: Adjacency Matrix for Case Study

The next step in model development is to specify the conditional probability distributions of the project risk network, which follow from equations (16), (17), (18), (19), (20), and (21). Causal networks represent conditional probability distributions as a list of objects. The size and type of each variable must be specified before specifying the conditional probability distributions. If a variable is discrete, its size is the number of possible values each variable can take, while if a variable is continuous, its size is the length of the vector of parameters. Continuous variables are assumed normally distributed with mean vector $\boldsymbol{\mu}$ and variance vector $\boldsymbol{\sigma}^2$. $\boldsymbol{\sigma}^2$ is a variance vector as opposed to a covariance matrix due to the property of conditional independence, as the parents of a variable are, by definition of causal networks, conditionally independent. If a continuous variable is conditional on q discrete parents, then the size of $\boldsymbol{\mu}$ and $\boldsymbol{\sigma}^2$ are both $q * q$.

First, root causes $\boldsymbol{c}$ are discrete variables of size 2, since a root cause either occurs or does not. Root causes $\boldsymbol{c}$ are selected from the risk register, where the prior of $\boldsymbol{c}$ is assumed based on the qualified probability indicated in the risk register. A negligible likelihood corresponds with probability 0.25, possible corresponds with probability 0.5, and likely corresponds with probability 0.75. Table 3 shows the prior probability of $\boldsymbol{c}$.

Table 3: Prior Probability of $\boldsymbol{c}$

Variable	Conditional Probability
c_1	[0.5 0.5]
c_2	[0.75 0.25]
c_3	[0.5 0.5]
c_4	[0.5 0.5]
c_5	[0.5 0.5]
c_6	[0.5 0.5]

Table 3 Continued

Variable	Conditional Probability
c_7	[0.25 0.75]
c_8	[0.25 0.75]

Second, risk events ***r*** are discrete variables of size 4 if conditionally dependent on some root cause or 2 if conditionally independent. If risk event r_i is conditionally dependent on root cause c_i, then the possible outcomes are $P(r_i = F \mid c_i = F)$, $P(r_i = F \mid c_i = T)$, $P(r_i = T \mid c_i = F)$, and $P(r_i = T \mid c_i = T)$. Risk events ***r*** are assumed based on a combination of the likelihood of the risk and whether or not a root cause is prevalent. Table 4 shows the prior probability of ***r***.

Table 4: Prior Probability of *r*

Variable	Conditional Probability
r_1	[0.25 0.5 0.75 0.5]
r_2	[0.25 0.5 0.75 0.5]
r_3	[0.25 0.5 0.75 0.5]
r_4	[0.5 0.75 0.5 0.25]
r_5	[0.75 0.25]
r_6	[0.75 0.25]
r_7	[0.5 0.75 0.5 0.25]
r_8	[0.5 0.75 0.5 0.25]

Table 4 Continued

Variable	Conditional Probability
r_9	[0.5 0.75 0.5 0.25]
r_{10}	[0.25 0.5 0.75 0.5]
r_{11}	[0.35 0.85 0.65 0.15]

Third, resource costs $\boldsymbol{a}$, unit costs $\boldsymbol{u}$, and quantities $\boldsymbol{q}$ are continuous variables with mean vector $\boldsymbol{\mu}$ and variance vector $\boldsymbol{\sigma}^2$. If a_i is the cost of resource i, and a_i is conditional on vector of risk events $\boldsymbol{r}$ of length j, then a_i has mean vector $\boldsymbol{\mu}$ and variance vector $\boldsymbol{\sigma}^2$ both of length $j * j$. The mean vector $\boldsymbol{\mu}$ and the variance vector $\boldsymbol{\sigma}^2$ are assumed based on the qualified impact and risk category designated in the risk register. For instance, variable q_1 is conditional on r_1, r_8, r_9, and r_{11}, therefore, q_1 has both mean and covariance vector of length 16. Table 5 shows the prior parameters of resource costs.

Table 5: Prior Parameters of Resource Costs

Variable	$\boldsymbol{\mu}$	$\boldsymbol{\sigma}$
$\boldsymbol{u_1}$	0.1	0.01
$\boldsymbol{q_1}$	[63000 78000 67000 78000 67000 82000 93000 82000 82000 71000 82000 97000 86000 87000 86000 101000]	[6300 7800 6700 7800 6700 8200 9300 8200 8200 7100 8200 9700 8600 8700 8600 10100]
$\boldsymbol{a_2}$	2500	250
$\boldsymbol{a_3}$	1830	180

Table 5 Continued

Variable	μ	σ
u_4	0.275	0.0275
q_4	12000	1200
a_5	[1500 2000 1600 2000 1600 2100 2500 2100 2100 1700 2100 2600 2200 2600 2200 2700]	[150 200 160 200 160 210 250 210 210 170 210 260 220 260 220 270]
u_6	0.175	0.0175
q_6	[4500 5500 16000 7500 16000 19000 3000 20000]	[450 550 1600 750 1600 1900 300 2000]
u_7	0.225	0.0225
q_7	[20000 30000]	[200 300]
a_8	1400	140
u_9	0.125	0.0125
q_9	4400	440
u_{10}	0.150	0.0150
q_{10}	14000	1400
u_{11}	[0.015 0.02]	[0.0015 0.0015]
q_{11}	9900	990
a_{12}	[11500 12000]	[1150 1200]
a_{13}	[650 700]	[65 70]

Table 5 Continued

Variable	$\boldsymbol{\mu}$	$\boldsymbol{\sigma}$
a_{14}	[1300 1500]	[130 150]
a_{15}	75	8
a_{16}	65	7
a_{17}	[2300 4500]	[230 450]
a_{18}	3800	380
a_{19}	21	11
u_{20}	0.2	0.02
q_{20}	20000	2000
a_{21}	15000	1500

The result of specifying the DAG and the conditional probability distributions is a fully developed causal network model of project risk. Figure 41 shows the project risk network for the case study.

The project risk model is a multiply connected, hybrid causal network, where discrete variables are indicated with a box and continuous variables are indicated with a circle. For ease of interpretation, Figure 41 does not decompose resource cost nodes into the corresponding unit cost and quantity nodes. However, this is only a graphical simplification, as the unit cost and quantity variable are indicated in Table 5.

The project risk network highlights a number of important issues. First, the network shows which resources are more or less at risk. For instance, resource a_1, labor, and a_5, construction management, are conditional on the highest number of risks. In fact, a_1 and a_5 have the same four risks r_1, r_8, r_9, and r_{11} in common. Construction management and labor are primarily required for the construction phase of the project, which happens to be the longest phase. Consequently, r_1, r_8, r_9, and r_{11} are all related to schedule delays, which cause additional construction management and labor requirements and, in turn, an indirect impact on the required costs of these resources.

The risk of construction schedule delays points to another important issue. The project risk network indicates which tasks require more or less resources. Not surprisingly, t_{13}, construction, requires the highest number of resources. The construction phase also shares a number of resources with the immediately following task t_{14}, commissioning and start-up. If the construction phase is either positively or negatively impacted, then the following commissioning and start-up phase is likely impacted as well. In short, the project risk network specifies which project tasks and, similarly, which project resources are high priorities from a risk perspective. Conversely, the network also indicates which project tasks and resources are low priorities. As previously mentioned, some resources tend to be fairly certain, and thus not a source of major risk. These resources require much less oversight and effort than other more uncertain resources, such as, in this case, labor and construction management.

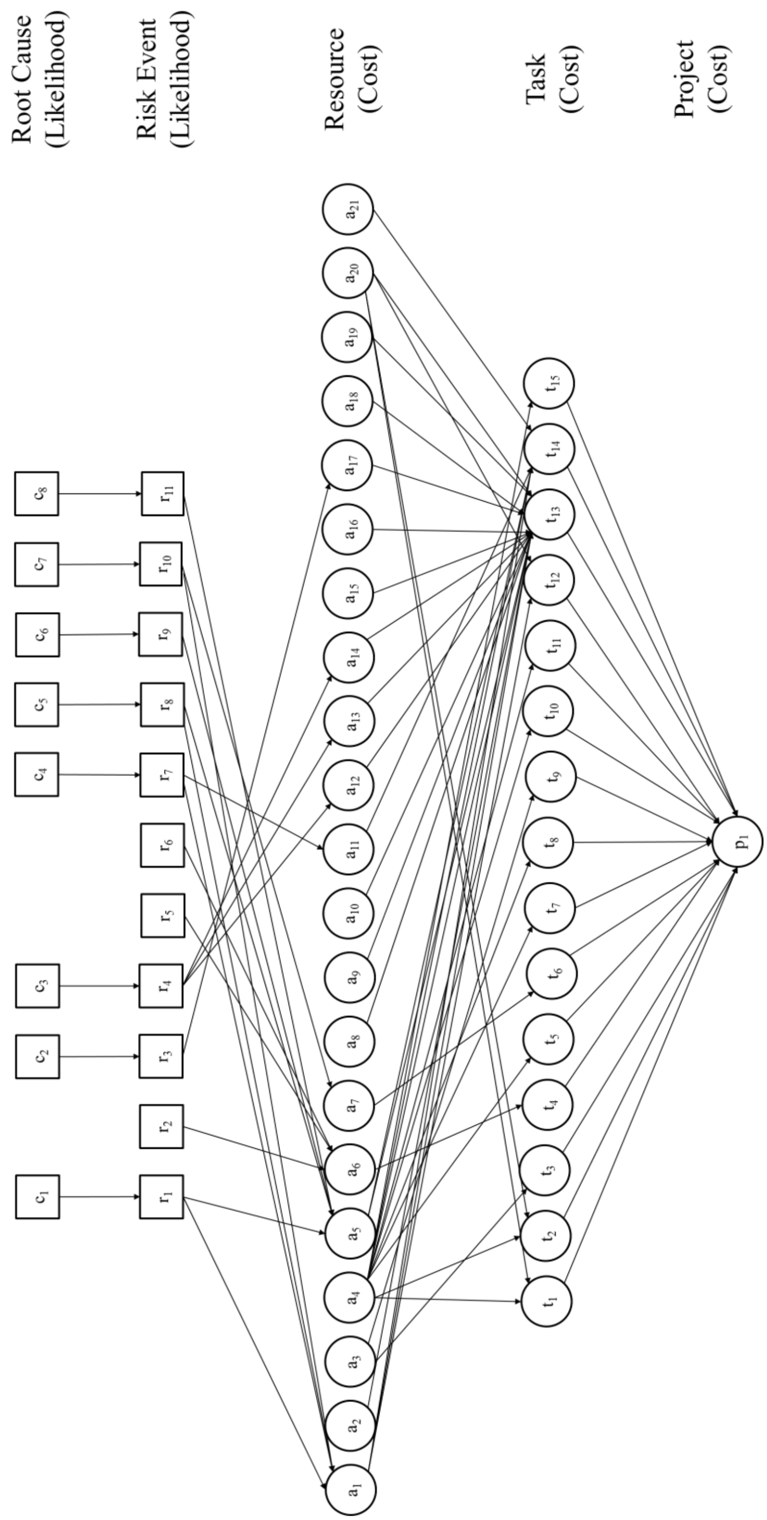


Figure 41: Project Risk Network for Case Study

7.5 Sensitivity Analysis and Validation

Following model development, the next step in the research methodology is sensitivity analysis. While the project risk network may indicate which risk events are more or less important, the actual response of the model to changes in risk events remains uncertain. Sensitivity analysis explains the uncertainty in the project risk model in terms of sources of uncertainty in the model parameters. From a risk management perspective, the primary concern is the sensitivity of overall project cost due to different sources of risk, as project cost is a key objective. In order to quantify the uncertainty of overall project cost due to sources of risk, an expectation-based sensitivity analysis is performed, which follows from equation (24). The motivation for this analysis is to measure the expected amount by which the expectation of project cost will change if one learns the true value of a risk event, where a risk event may either occur or not.

Figure 42 shows the results of the univariant sensitivity analysis. The horizontal axis indicates the expected change in overall project cost, where the expected change is negative if a risk event occurs and positive if it does not. The vertical axis specifies five risk events which cause the greatest expected change in project cost.

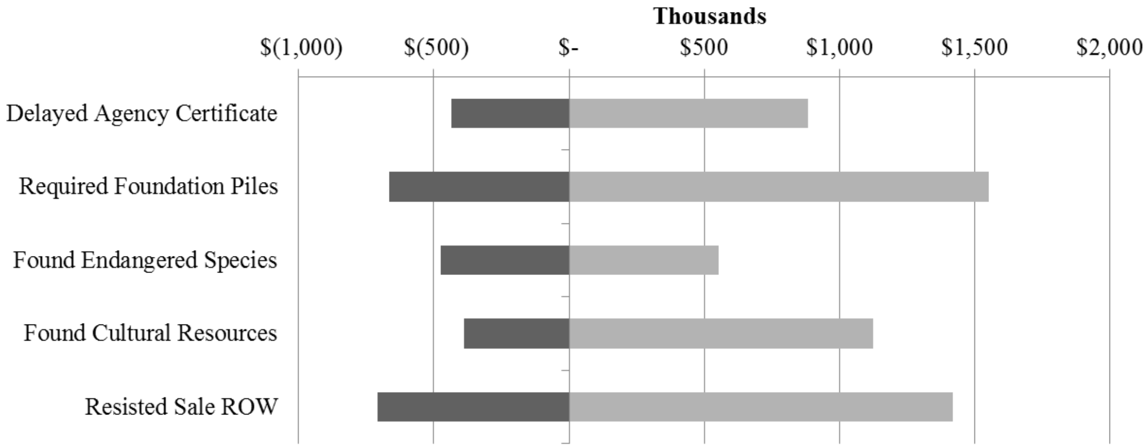


Figure 42: Univariate Sensitivity Analysis of Project Cost (*n* = 1000)

According to the results, risk event r_3, required foundation piles, causes the greatest negative change in expected project cost, followed closely by risk event r_{10}, resisted sale of ROW. Conversely, risk event r_{10}, resisted sale of ROW, causes the greatest positive change in expected project cost, followed closely by r_3, required foundation piles.

In addition to explaining the uncertainty of the project risk model related to model parameters, sensitivity measures serve two important functions. First, as the model response is uncertain, sensitivity measures support model validation of the prior parameters. For instance, the results of the sensitivity analysis are supported by the risk register, where r_3 and r_{10}, which cause the greatest change in expected project cost, are the only risks listed in the *major* impact category. In fact, the other risks shown in Figure 42 are listed in the risk register as having either a *major* or *significant* impact, which is consistent with the sensitivity analysis.

Second, sensitivity analysis supports risk management by indicating issues with current mitigation strategies. For example, the current risk mitigation strategy for r_{10} is to accept the risk of the landowner refusing to sell the ROW, due to the low likelihood of this risk occurring. However, this mitigation strategy is not based on the model response due to r_{10} occurring, and this information may lead to a different strategy, such as reducing the risk by seeking legal counsel. Similarly, the current risk mitigation strategy for r_3 is to reduce the risk through a geotechnical study. The sensitivity analysis may provide support for this strategy or lead to additional strategies that reduce the risk further.

7.6 Bayesian Inference

One of the primary applications of the project risk network is Bayesian inference. Bayesian inference involves updating beliefs based on evidence of events. This application is important for two primary reasons. First, Bayesian inference provides answers to probabilistic questions based on current information. As events are observed over the course of the project, this evidence may be added to the model to produce updated inferences. Second, Bayesian inference supports the evaluation of potential mitigation strategies. In addition to actual information, hypothetical evidence may be added to the model in order to evaluate different mitigation scenarios.

To illustrate Bayesian inference, Figure 43 shows the marginal of a_{17} given different scenarios, which follows from equation (27). The central plot indicates the base case marginal, the leftmost plot indicates the marginal given r_3 did not occur, and the rightmost plot indicates the marginal given r_3 did occur. If r_3 represents actual evidence, then the leftmost plot represents the updated marginal given r_3 did not occur and the rightmost plot is the updated marginal given r_3 did occur. Conversely, if r_3 represents hypothetical evidence, then these two plots represent possible scenarios given the outcome of r_3. Not surprisingly, the expected value of the marginal given r_3 occurs is higher than the base case marginal; while the expected value is lower if r_3 does not occur.

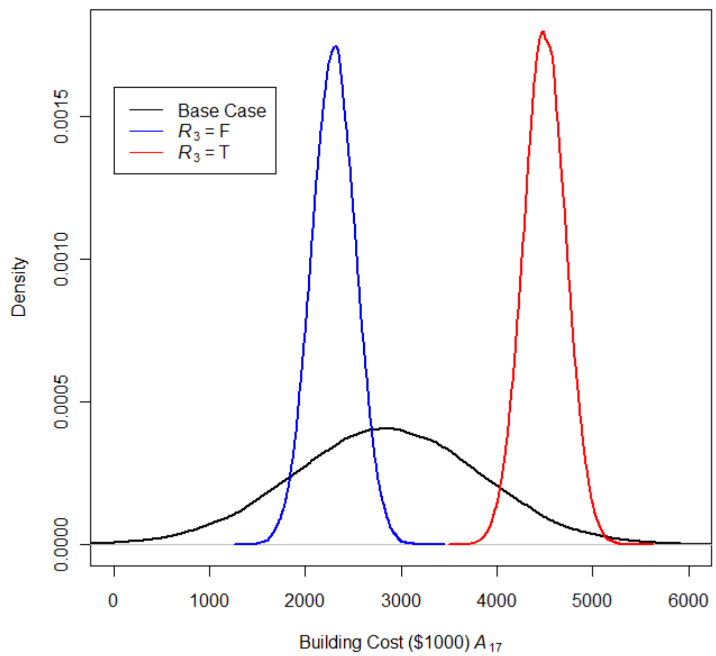


Figure 43: Marginal of a_{17} for Different Scenarios

The overall goal of Bayesian inference is to estimate the probability distribution of project cost. ***Algorithm 3*** provides an MC approach to approximate the distribution of project cost. Figure 44 shows the histogram for a MC simulation of project cost (n = a1000). The expected project cost is $68,478,000 and the variance is 7,415,000.

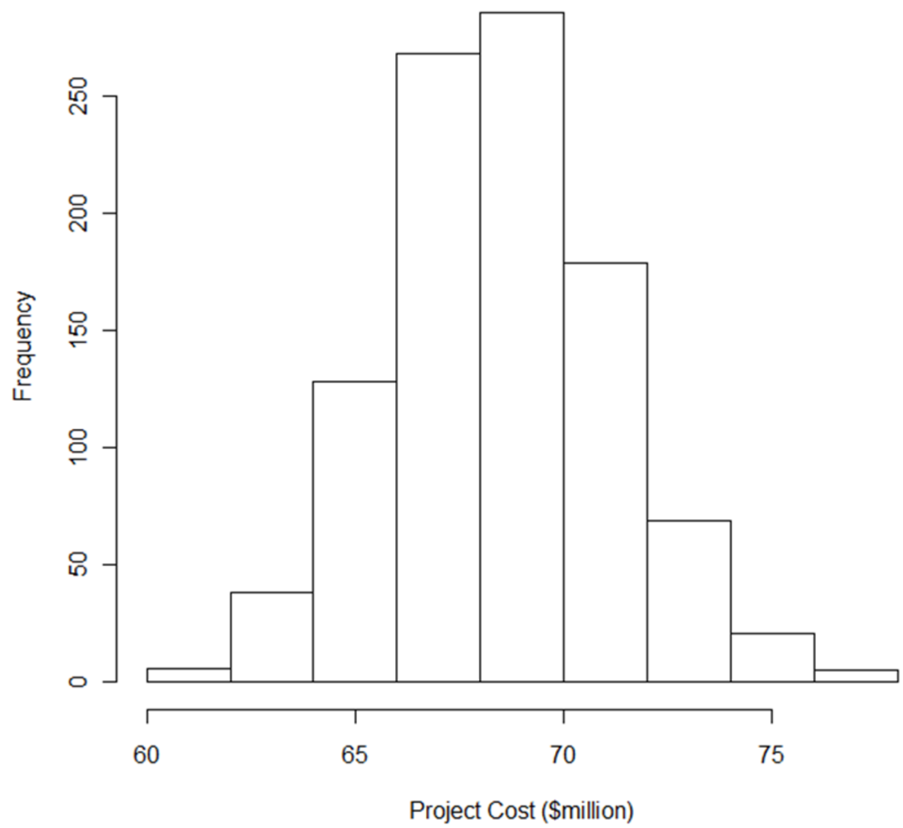


Figure 44: Histogram of Project Cost (*n* = 1000)

The project cost correlation matrix provides a measure of the linear dependency between project tasks. The project cost correlation matrix is a model requirement for a number of risk cost models, such as second moment and MC simulation methods. In order to estimate correlations, these models depend on either historical data or expert opinions. Both of these approaches are difficult in practice, however, as historical information is often unavailable or incomplete, and many experts have difficulty specifying subjective correlations.

The project risk network provides a method to estimate the project cost correlation matrix from the prior parameters of the model. This approach neither depends on data nor expert opinions. Two tasks are correlated if the tasks have common resource requirements and, in turn, common risk dependencies. Therefore, the project cost correlation matrix helps explain the dependency structure of the project.

Following the approach outlined in ***Algorithm 2***, Table 6 shows the upper truncated project cost correlation matrix for the project risk model. The correlation matrix indicates which tasks are more or less dependent. As planning t_1 and project funding t_2 have exactly the same resources in common, and thus the same risk dependencies, the correlation between t_1 and t_2 is exactly 1. Conversely, the correlation matrix also indicates which tasks are relatively independent. Planning t_1 and engineering t_3 do not have any resource requirements in common, and thus no shared risk dependencies. Therefore, the correlation between t_1 and t_3 is close to zero.

Table 6: Project Cost Correlation Matrix (*n* = 1000)

	t_1	t_2	t_3	t_4	t_5	t_6	t_7	t_8	t_9	t_{10}	t_{11}	t_{12}	t_{13}	t_{14}	t_{15}
t_1	1.00	1.00	-0.07	-0.03	0.62	-0.02	0.62	0.62	-0.07	0.62	0.62	1.00	0.20	0.06	0.62
t_2		1.00	-0.07	-0.03	0.62	-0.02	0.62	0.62	-0.07	0.62	0.62	1.00	0.20	0.06	0.62
t_3			1.00	-0.01	-0.03	-0.04	-0.03	-0.03	1.00	-0.03	-0.03	-0.07	0.00	0.03	-0.03
t_4				1.00	-0.01	-0.03	-0.01	-0.01	-0.01	-0.01	-0.01	-0.03	-0.04	-0.02	-0.01
t_5					1.00	-0.05	1.00	1.00	-0.03	1.00	1.00	0.62	0.08	0.08	1.00
t_6						1.00	-0.05	-0.05	-0.04	-0.05	-0.05	-0.02	0.00	-0.02	-0.05
t_7							1.00	1.00	-0.03	1.00	1.00	0.62	0.08	0.08	1.00
t_8								1.00	-0.03	1.00	1.00	0.62	0.08	0.08	1.00

Table 6 Continued

	t_1	t_2	t_3	t_4	t_5	t_6	t_7	t_8	t_9	t_{10}	t_{11}	t_{12}	t_{13}	t_{14}	t_{15}
t_9									1.00	-0.03	-0.03	-0.07	0.00	0.03	-0.03
t_{10}										1.00	1.00	0.62	0.08	0.08	1.00
t_{11}											1.00	0.62	0.08	0.08	1.00
t_{12}												1.00	0.20	0.06	0.62
t_{13}													1.00	0.44	0.08
t_{14}														1.00	0.08
t_{15}															1.00

7.7 Bayesian Learning

Another primary application of the project risk network is Bayesian learning. Bayesian learning is the process of updating priors based on the observed data. As incomplete data is a common problem for many projects, Bayesian learning may only be possible for to a subset of the model parameters. However, a key aspect of Bayesian methods is the ease with which sequential analysis may be performed. As project data become available over the course of the project, this data may be added to the model to update inferences. Furthermore, the project risk model identifies data collection requirements and the relative importance of the data, as the model sensitivity depends on the model parameters.

To illustrate Bayesian learning, consider a_{17} sitework and building. a_{17} is conditionally dependent on r_3, the risk which indicates whether or not a pile foundation is required. As noted in the sensitivity analysis, r_3 causes the greatest change in overall expected project cost. Suppose data is available on a_{17} from historical information on previous projects. Assume a_{17} is bivariate normal with mean vector $\boldsymbol{\theta}$ and covariance vector $\boldsymbol{\Sigma}$. In this case, a_{17} has the density:

$$f\left(Y \mid \theta, \Sigma\right) = \frac{1}{\left(2\pi\right)^{p/2} \left|\Sigma\right|^{1/2}} \exp\left(\frac{1}{2}\left(x-\theta\right)^{T} \Sigma^{-1}\left(x-\theta\right)\right)$$

Further, suppose (y_1, y_2) are measured on a sample of projects of size 10 (n = 10) in constant dollars where:

$$y_1 = r_3 \text{ does not occur}$$

$$y_2 = r_3 \text{ occurs}$$

The sufficient statistics are taken from the sample as follows:

$$\left(\bar{y}_1, \bar{y}_2\right) = \left(2250, 4650\right)$$

$$S^2 = \begin{bmatrix} 25650 & 0 \\ 0 & 28750 \end{bmatrix}$$

A conjugate prior distribution is the normal-inverse-Wishart, which has the following form:

$$p\left(\theta, \Sigma\right) \propto \left|\Sigma\right|^{-\left(v_0+p+2\right)/2} \exp\left(-\frac{1}{2} tr\left(S_0 \Sigma^{-1}\right)\right) \exp\left(-\frac{\tau_0}{2}\left(\theta-\mu_0\right)^{T} \Sigma^{-1}\left(\theta-\mu_0\right)\right)$$

with the parameters v_0, τ_0, μ_0, and S_0, where μ_0 is the prior mean vector $\left[2500 \ \ 4500\right]$ and S_0 is the prior covariance vector $\left[45000 \ \ 49000\right]$ of a_{17}, which are taken from the project risk model. Assume v_0 is 1, and τ_0 is 2. It can be shown that the posterior has the same form but with different parameters. The posterior distribution is given by:

$$\Sigma \mid Y \sim \text{inverse-Wishart}\left(\nu_n, \left(S_n^2\right)^{-1}\right)$$

$$\nu_n = \nu_0 + n,\ \mathrm{S}_n^2 = S_0 + nS^2 + \left(\frac{\tau_0 n}{\tau_0 + n}\right)(\bar{y} - \mu_0)(\bar{y} - \mu_0)^T$$

$$\theta \mid \Sigma, Y \sim N\left(\mu_n, \Sigma/\tau_n\right)$$

$$\tau_n = \tau_0 + n,\ \mu_n = \left(\frac{\tau_0}{\tau_0 + n}\right)\mu_0 + \left(\frac{n}{\tau_0 + n}\right)\bar{y}$$

In this case, the posterior distribution is given by:

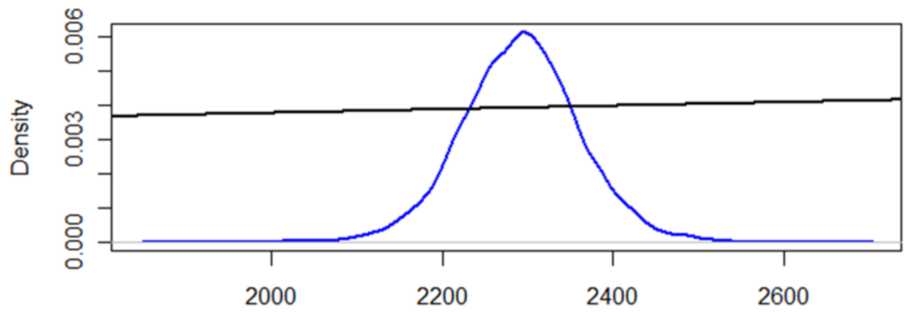
$$\Sigma \mid Y \sim \text{inverse-Wishart}\left(11, \left(S_n^2\right)^{-1}\right)$$

$$\mathrm{S}_n^2 = S_0 + 10S^2 + 1.67(\bar{y} - \mu_0)(\bar{y} - \mu_0)^T$$

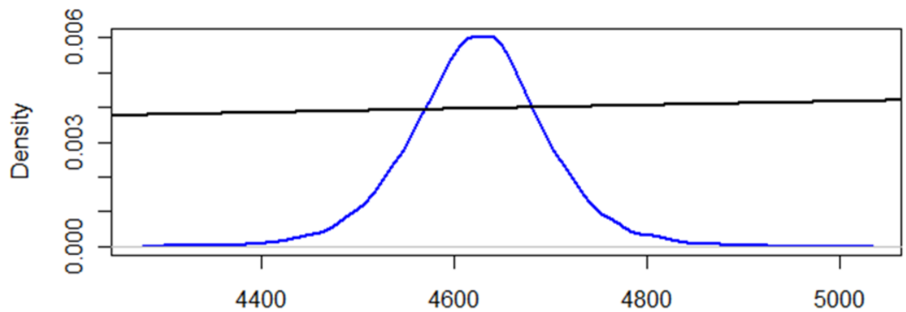
$$\theta \mid \Sigma, Y \sim N\left(\mu_n, \Sigma/12\right)$$

$$\mu_n = 0.17\mu_0 + 0.83\bar{y}$$

Figure 45 shows the prior and posterior of θ_1 (a) and θ_2 (b), where (θ_1, θ_2) is the mean vector of a_{17}. Now by generating a sample from the posterior and taking point estimates, the posterior may be used to update the parameters of a_{17}. For instance, the posterior mean is often taken as the updated parameters of the causal network model. In this case, the posterior mean $(2292,\ 4625)$ serves as the updated mean vector (θ_1, θ_2) of a_{17}.



(a)



(b)

Figure 45: Prior and Posterior Densities of θ_1 (a) and θ_2 (b)

7.8 Task Dependency Measures

7.8.1 Task Parent Matrix

The task parent matrix ***P*** identifies and measures interdependencies between project tasks due to resource constraints. From equation (35), the upper truncated task parent matrix ***P*** for the case study is shown in Figure 46. In this case, ***P*** is a square, symmetric matrix of size 15, where 15 is the number of tasks in the project risk network. Nonzero off-diagonal elements specify the number of resources that a pair of tasks have in common, while nonzero diagonal terms indicate the total number of resources required per task. The total number of common resources between tasks is 51. The two tasks with the highest number of common resources are t_{13} and t_{14}, which have 2 resources in common, and the task with the highest number of common resources is t_{13}, construction, which has 12 total resources in common with the other tasks. Each task has at least one resource in common with other tasks, with the exception of t_6, environmental. Each task requires at least one resource, and the task with the highest number of required resources is t_{13}, which requires 16 total resources.

The task parent matrix indicates that, while no particular task is more or less connected, in general, the project tasks are highly integral in terms of resources. This information has important implications for project management, as no tasks, with the possible exception of t_6, should be managed or controlled independently. As different tasks are often managed by different groups within the organization, the task parent matrix indicates which groups should coordinate in order to control the common resource constraints. Furthermore, the task parent matrix serves as a learning tool. While the matrix indicates expected resource constraints, the matrix may be updated to represent actual resource constraints, which may, in turn, be used on future projects to improve communication. Finally, the task parent matrix indicates which tasks, namely t_{13}, construction, require a disproportionately high number of resources. This information is important for scheduling purposes, as resource allocation involves leveling resources across the project duration. As such, tasks with a high number of resource requirements are difficult to level. t_{13} is also, not surprisingly, the longest task, and this information supports decomposing t_{13} into separate subtasks in order to obtain a better understanding of the resource requirements and constraints.

	Task 1	Task 2	Task 3	Task 4	Task 5	Task 6	Task 7	Task 8	Task 9	Task 10	Task 11	Task 12	Task 13	Task 14	Task 15
Task 1	2	1			1		1	1		1	1		2	1	1
Task 2		1			1		1	1		1	1		1	1	1
Task 3			1	1					1						
Task 4				1											
Task 5					1		1	1		1	1		1	1	1
Task 6						1									
Task 7							1	1		1	1		1	1	1
Task 8								1		1	1		1	1	1
Task 9									1						
Task 10										1	1		1	1	1
Task 11											1		1	1	1
Task 12												1	1	1	
Task 13													16	2	1
Task 14														3	1
Task 15															1

Figure 46: Task Parent Matrix for Case Study

7.8.2 Task Grandparent Matrix

Similarly, the task grandparent matrix ***G*** identifies and measures interdependencies between project tasks due to risks. From equation (36), the upper truncated task grandparent matrix ***G*** for the case study is shown in Figure 47. ***G*** is a square, symmetric matrix of size 15, where 15 is the number of tasks in the project. Nonzero off-diagonal elements specify the number of risks that a pair of tasks have in common, while the diagonal elements indicate the total number of paths with which risks may impact each task. The total number of common risk dependencies is 12, while the only tasks with common risk dependencies the bidding t_{12}, construction t_{13}, and commissioning tasks t_{14}, which each have 4 common risk dependencies. The task with the highest number of risk dependencies is t_{13}, with a value of 6.

The task grandparent matrix indicates that, with the exception of t_{12}, t_{13}, and t_{14}, the project is disconnected in terms of risks. This information has important implications for risk management. First, independent risks may be managed independently, while interdependent risks require coordination for risk management. Second, the task dependency structure supports the evaluation of mitigation strategies. For example, the common risk dependencies occur in subsequent tasks and in the latter half of the project. A probable mitigation strategy is to bundle the closely grouped risks and transfer them to another party. Third, the dependency structure may serve as the basis for prediction models, as observation of a common risk may impact subsequent tasks. Finally, similar to the previous case, the project grandparent matrix serves as a learning tool. The matrix indicates expected common risks, which may be updated to represent actual common risks and may, in turn, serve as the basis of project risk models for future projects.

	Task 1	Task 2	Task 3	Task 4	Task 5	Task 6	Task 7	Task 8	Task 9	Task 10	Task 11	Task 12	Task 13	Task 14	Task 15
Task 1															
Task 2															
Task 3															
Task 4				3											
Task 5															
Task 6						1									
Task 7															
Task 8															
Task 9															
Task 10															
Task 11															
Task 12												4	4	4	
Task 13													6	4	
Task 14														4	
Task 15															

Figure 47: Task Grandparent Matrix for Case Study

7.8.3 Task Dependency Matrix

The resultant task dependency matrix integrates the task parent and grandparent matrices in order to improve the coordination of overall project interdependencies. The resultant task dependency matrix for the case study is shown in Figure 48. The task dependency matrix is a square, symmetric matrix of size 15, where 15 is the number of project tasks. Off-diagonal elements with a value of Y indicate a common resource dependency, while a value of Z indicates a common risk dependency, and diagonal elements indicate a resource requirement Y or a risk dependency Z. The total number of common resource dependencies is 48, and the number of common risk dependencies is

3. t_{13}, construction, has both the highest number of common resources (8) and risks (2) dependencies with other tasks.

		Task 1	Task 2	Task 3	Task 4	Task 5	Task 6	Task 7	Task 8	Task 9	Task 10	Task 11	Task 12	Task 13	Task 14	Task 15
Stage 1 FEL	Task 1	Y	Y			Y		Y	Y		Y	Y		Y	Y	Y
	Task 2		Y			Y		Y	Y		Y	Y		Y	Y	Y
Stage 2 FEL	Task 3			Y	Y					Y						
	Task 4				Z											
	Task 5					Y		Y	Y		Y	Y		Y	Y	Y
	Task 6						Z									
	Task 7							Y	Y		Y	Y		Y	Y	Y
	Task 8								Y		Y	Y		Y	Y	Y
Engineering Design	Task 9									Y						
Regulatory Approvals	Task 10										Y	Y		Y	Y	Y
Procurement	Task 11											Y		Y	Y	Y
Construction	Task 12												Z	Z	Z	
	Task 13													Z	Z	Y
	Task 14														Z	Y
In-Service	Task 15															Y

Figure 48: Task Dependency Matrix for Case Study

7.9 Task Modularity Measures

7.9.1 Task Degree Modularity

Task modularity is a measure of the relative disconnectedness of tasks due to the absence of common dependencies. Task degree modularity T_D relates negatively to the number of resources with which a given task has in common with other tasks, while task path modularity T_P relates negatively to the number of risk paths with which a given task has in common with other tasks. Task modularity measures have similar applications as the dependency measures of the previous section, except that they provide a more concise interpretation of the dependency structure of projects.

	T_D	T_P
Task 1	0.524	1.000
Task 2	0.571	1.000
Task 3	0.905	1.000
Task 4	1.000	1.000
Task 5	0.571	1.000
Task 6	1.000	1.000
Task 7	0.571	1.000
Task 8	0.571	1.000
Task 9	0.952	1.000
Task 10	0.571	1.000
Task 11	0.571	1.000
Task 12	0.904	0.998
Task 13	0.429	0.998
Task 14	0.476	0.998
Task 15	0.571	1.000

Figure 49: Task Modularity Measures for the Case Study

From equations (30) and (31), respectively, the task degree modularity and task path modularity measures for the case study are shown in Figure 49. Low task degree modularity indicates a task that is relatively connected to other tasks due to resource constraints. Conversely high task degree modularity indicates a task that is relatively disconnected from the other tasks. The task with the lowest degree modularity is construction t_{13}, with a value of 0.429, while the tasks with the highest degree modularity are environmental t_4 and ROW t_6, each with a value of 1. To manage risks, one may consider redesigning t_{13} in order to cut the interactions with other tasks.

Low path degree modularity indicates a task that is relatively connected to other tasks due to risk dependencies, while high path modularity indicates a task that is relatively disconnected from the other tasks. The task path modularity measures are all relatively close to one. These results provide further support that the project is relatively disconnected in terms of risk. From the perspective of risk reorganizing t_{12} – t_{14} should be considered.

7.10 Summary

This chapter presented a case study in order to illustrate the general research methodology and validate the general project risk model for practical applications. The following chapter presents a summary of the research findings, conclusions, and recommendations for future research.

CHAPTER VIII

SUMMARY OF MAJOR FINDINGS, LIMITATIONS AND RECOMMENDATIONS FOR FUTURE RESEARCH

This chapter presents a summary of the major findings of this dissertation, a discussion of the limitations, and directions for future research. The chapter is divided into two primary sections. The first section provides a summary of the major findings of this dissertation. The second section discusses the limitations of the methodology and suggestions for future research. The discussion includes possible extensions of the proposed methodology as well as the identification of research needs in the broad area of project risk analysis.

8.1 Summary of Major Findings

The goal of this dissertation is to develop a general corporate portfolio risk analysis methodology that identifies theoretical causal relationships and integrates expert opinions with the observed data. Probabilistic and expert risk analysis methods are often at odds; however, this dissertation integrates the analytical and experiential approaches to risk analysis. The proposed conceptual framework takes a resource-based view, where risk is identified and measured in terms of the uncertainty associated with project resources. The methodological framework utilizes causal networks to model risk and the associated consequences.

The result is a new general corporate portfolio risk analysis methodology. The causal network methodology provides a graphical explanation of risk events and the interdependencies between projects as well as a probabilistic model of project uncertainty. Firms may investigate different risk mitigation and management strategies by simply selecting a different set of causal network model parameters and/or a different structure and measuring the overall effect of these changes on project uncertainty. Furthermore, this methodological framework identifies risk-based dependencies given varying levels of information, and promotes organizational learning by identifying which project information is more or less valuable to the organization.

This research requires a general knowledge of multiple topics, including project risk management, the resource-based view (RBV) theory, Bayesian and causal network theory, and portfolio theory. This background supports a general project risk model, which the firm may use to evaluate risk management and mitigation strategies. From the perspective of the owner, agency, and investor, the objective is to estimate the risk cost or premium that contractors may include in their bids; while from the engineer and contractor's perspective, the objective is to determine the optimal design-construction strategy that minimizes total project cost. The general project risk model supports either objective, regardless of the type, size, or complexity of the firm.

This dissertation included five main research concepts. The first research concept showed that the project or corporate portfolio may be defined from a RBV. This concept is supported by existing literature on the RBV, which has been applied to project management, as well as project management tools, such as the resource breakdown structure (RBS), which is a hierarchical structure of the project resources. The conceptual framework in Chapter 3 introduced this research concept, where the conceptual framework takes a RBV of the firm. Logically, if the firm is defined as a set of resources, and the corporate portfolio is part of the firm, then the corporate portfolio is a subset of firm-related resources.

The second research concept showed that all project-related risk may be identified and measured in terms of the uncertainty associated with project resources. This concept is a logical extension of the RBV of the firm, where if the project is defined as a subset of firm-related resources, then risk events impact the project(s) through project-related resources. The conceptual framework in Chapter 3 also introduced this research concept, which assumes a direct relationship between risk events and project-related resources and an overall impact on the corporate portfolio.

The third research concept was to develop a causal network methodology for project risk assessment. The methodological framework in Chapter 3 introduced this research concept, and chapter 4 developed the formal causal network methodology for project risk analysis. The causal network methodology consists of five primary steps: model input, specification, validation, analysis, learning, and use. The causal network methodology is general in the sense that it may be applied to any project or corporate portfolio regardless of factors such as type, size, or complexity. The methodology provides a graphical display of project risk events and a method to model interdependencies between risk events. The methodology also opens the field of project risk analysis to the broad field of causal network methods. The causal network represents the prior distributions of the project-related parameters, which may be updated based on information contained in data.

The fourth research concept showed that the project risk network topology provides valuable information to project management. In practice, data is often not available for development and validation of the causal network. However, measures of the project risk network provide an explanation of the dependency structure of projects in the absence of sufficient data. Network measures provide important information about the projects and the interdependencies between projects by analyzing the structure of the project risk network.

Chapter 5 introduced network measures for project risk analysis, which support management in risk-related decision analysis. Management may test strategies by changing parameters or breaking dependencies and observing the effect on the overall network. Measures help identify gaps between expected task-related dependencies and actual ones. Firms may use these measures as learning tools to help improve communication concerning task-related risk events and strategies. Measures such as task modularity help break the project into interrelated parts and facilitate coordination among similar tasks.

The fifth research concept extended the causal network methodology for corporate portfolio risk analysis. As the methodological framework is general, the causal network methodology applies at the corporate portfolio level, where risk events may impact multiple projects and resources are often common between projects. This concept is introduced in chapter 6, which developed the corporate portfolio risk analysis methodology as well as network measures for the corporate portfolio. This methodology supports the management of interdependencies between projects, the coordination of multiple projects, and the management of resources and constraints. Project network measures provide a means to identify and assess interdependencies between projects due to shared risk and resource-related dependencies, and support corporate resource allocation in a way that balances risk and strategic objectives. These measures serve as a framework to guide decision making and overall risk management. If firms then want to move to a formal causal network analysis, the theoretical and methodological framework is in place.

Chapter 7 presented a case study of a compressor station project in order to illustrate the general research approach and methodology. The case study showed that the proposed methodology may be implemented by integrating existing project management tools. The case study also provided empirical support for the conceptual framework by providing multiple examples of linking risk events to project resources.

8.2 Limitations and Directions for Future Research

While this dissertation developed a general corporate portfolio risk analysis methodology, risk management remains an important research topic. Many important problems still exist, both in regard to the proposed methodology and in the wide field of project risk management. This research explored multiple risk management and mitigation strategies but numerous more may be evaluated with causal network methods. Some of the identified problems that require further research include the following:

- Develop a dynamic causal network methodology for project risk analysis. This extension is the next logical step in order to model the dynamics of project-related risk over time. This problem is important for multiple reasons: (1) the identification, assessment, and management of project schedule-related risk, and (2) the integration of this methodology with other dynamic models.

- Develop methods for Bayesian structural learning of the project risk network. This research discusses Bayesian parameter learning. However, the underlining model structure may be unknown or only partially known. Firms may require tools to help select arcs between variables and, in turn, validate the conditional dependencies. The Bayesian network literature includes several structural learning algorithms, which could be applied to project risk networks.

- Model intangible resources or capabilities and the associated risk. This research explicitly models tangible resources, such as labor, materials, and equipment. Intangible resources, such as company brand, technology, education, or skill set, and the associated risk are more difficult to identify and measure. This problem requires innovative methods and is a great challenge for future research.

- Validate project risk networks in practice. This topic is of great importance to ensure the successful implementation of the proposed integrated framework for project risk analysis. In spite of the abundance of project-related data available to firms, the data required to validate and learn the causal network is not easily attainable. One reason for this problem is that data is often proprietary, as it directly affects the firm's business. Another reason is that project-related data is not often collected for the variables in the causal network. However, if firms implement the proposed methodology, the method identifies which data is required to validate project risk networks.

REFERENCES

Amit, R., & Schoemaker, P. J. (1993). Strategic Assets and Organizational Rent. *Strategic Management Journal, 14*, 33-46.

Barney, J. (1991). Firm Resources and Sustained Competitive Advantage. *Journal of Management, 17*, 99-120.

Barney, J. B. (2001). Is the Resource-based "View" a Useful Perspective for Strategic Managment Research? Yes. *Academy of Management Review, 26*, 41-56.

Barney, J. B., & Hesterly, W. S. (2010). VRIO Framework: In Strategic Management and Competitive Advantage. 68-86.

Bedford, T., & Cooke, R. (2003). *Probabilistic Risk Analysis: Foundations and Methods.* Cambridge: Cambridge University Press.

Biggs, N., Lloyd, E., & Wilson, R. (1968). *Graph Theory, 1736-1936.* Oxford University Press.

Caron, F. (2007). Engineering and Contracting Projects: a Value at Risk based Approach to Portfolio Balancing. *International Journal of Project Management, 25*, 569-578.

Conner, K. R. (1991). A Historical Comparison of Resource-based Theory and Five Schools of Thought Within Industrial Organization Economics: Do We have a New Theory of the Firm? *Journal of Management, 17*, 121-154.

Conner, K. R., & Prahalad, C. K. (1996). A Resource-based Theory of the Firm: Knowledge Versus Opportunism. *Organizational Science, 7*, 477-501.

Covello, V. T., & Mumpower, J. (1985). Risk Analysis and Risk Management: an Historical Perspective. *Risk Analysis, 5*, 103-120.

Darwiche, A. (2009). *Modeling and Reasoning with Bayesian Networks.* Cambridge: Cambridge University Press.

Davidson, L. B., & Cooper, D. O. (1976). A Simple Way of Developing a Probability Distribution of Present Value. *Journal of Petroleum Technology, 28,* 1069-1078.

Department of Defense. (1968). *Work Breakdown Structures for Defense Material Items.* Washington D.C.

Department of Defense. (2006). *Risk Management Guide for DoD Acquisition.* Washington D.C.

Department of Defense and National Aeronautical Space Administration. (1962). *PERT/COST System Design Guide.* Washington D.C.

Dickinson, M. W., Thornton, A. C., & Graves, S. (2001). Technology Portfolio Management: Optimizing Interdependent Projects Over Multiple Time Periods. *IEEE Transactions on Engineering Management, 48*, 518-527.

Dierickx, I., & Cool, K. (1989). Asset Stock Accumulation and Sustainability of Competitive Advantage. *Management Science, 35*, 1504-1511.

Eckes, G. (2001). *The Six Sigma Revolution: How General Electric and Others Turned Process Into Profits.* New York: Wiley.

Fang, C., Marle, F., Zio, E., & Bocquet, J.-C. (2012). Network-Theory-based Analysis of Risk Interactions in Large Engineering Projects. *Reliability Engineering and System Safety, 106*, 1-10.

Fenton, N., & Neil, M. (2011). The Use of Bayes and Causal Modelling in Decision Making, Uncertainty and Risk. *CEPIS Upgrade, 12*, 10.

Garvey, P. R. (2009). *Analytical Methods for Risk Management.* Boca Raton, Florida: CRC Press.

Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. (2004). *Bayasian Data Analysis.* Boca Raton: Chapman & Hall/CRC.

Groth, K., Wang, C., & Mosleh, A. (2010). Hybrid Causal Methodology and Software Platform for Probabilistic Risk Assessment and Safety Monitoring of Socio-Technical Systems. *Reliability Engineering and Systems Safety, 95*, 1276-1285.

Han, S. H., Diekmann, J. E., Lee, Y., & Ock, J. H. (2004). Multicriteria Financial Portfolio Risk Management for International Projects. *Journal of Construction Engineering and Management, 130*, 346-356.

Hillier, F. S. (1963). The Derivation of Probabilistic Information for the Evaluation of Risky Investments. *Management Science, 9*, 444-457.

Hoff, P. D. (2009). *A First Course in Bayesian Statistical Methods.* New York: Springer.

Hubbard, D. (2009). *The Failure of Risk Management.* Hoboken: John Wiley & Sons.

Huchzermeier, A., & Loch, C. H. (2001). Project Management Under Risk: Using the Real Options Approach to Evaluate Flexibility in R&D. *Management Science, 47*, 85-101.

Ibe, O. C. (2011). *Fundamentals of Stochastic Networks.* Hoboken: John Wiley & Sons, Inc.

ISO. (2009). *Risk Management-Principles and Guidelines.* Geneva: International Organization for Standardization.

J.P. Morgan. (1996). *Risk Metrics.* New York: Morgan Guaranty Trust Company.

Jugdev, K. (2004). Through the Looking Glass: Examining Theory Development in Project Management with the Resource-Based View Lens. *Project Management Journal, 35*, 15-26.

Keefer, D. L., & Bodily, S. L. (1983). Three-Point Approximations for Continuous Random Variables. *Management Science, 29*, 595-609.

Kim, H.-J., & Reinschmidt, K. F. (2011). Association of Risk Attitude with Market Diversification in the Construction Business. *Journal of Management in Engineering, 27*, 66-74.

Kraaijenbrink, J., Spender, J. C., & Groen, A. J. (2009). The Resource-Based View: A Review and Assessment of Its Critiques. *Journal of Management, 36*, 349-372.

Laplace, P. S. (1986). Memoir on the Probability of the Causes of Events. *Statistical Science, 1*, 364-378.

Lavie, D. (2006). The Competitive Advantage of Interconnected Firms: An Extension of the Resource-based View. *Academy of Management Review, 31*, 638-658.

Mahoney, J. T., & Pandian, J. R. (1992). The Resource-based View Within the Conversation of Strategic Management. *Strategic Management Journal, 13*, 363-380.

Makadok, R. (2001). Toward a Synthesis of the Resource-based and Dynamic-Capability Views of Rent Creation. *Strategic Management Journal, 22*, 387-401.

March, J. G., & Shapira, Z. (1987). Managerial Perspectives on Risk and Risk Taking. *Management Science, 33,* 1404-1418.

Markowitz, H. (1952). Portfolio Selection. *The Journal of Finance*, *7*, 77-91.

Markowitz, H. (1959). *Portfolio Selectio: Efficient Diversification of Investments.* New York: John Wiley & Sons.

Marle, F., Vidal, L.-A., & Bocquet, J.-C. (2013). Interactions-based Risk Clustering Methodologies and Algorithms for Complex Project Management. *International Journal of Production Economics*, *142*, 225-234.

Minato, T. (1994). *A Methodology for Project Risk Control: a Work Package-based Approach using Historical Cost Control Data.* Berkeley: University of California at Berkeley.

Moder, J. J., & Rodgers, E. G. (1968). Judgment Estimates of the Moments of PERT Type Distributions. *Managment Science, 15*, 86-83.

National Aeronautics & Space Administration. (2011). *NASA Risk Management Handbook.* Washington D.C.

National Institute of Standards and Technology. (2010). *Guide for Applying the Risk Management Framework to Federal Information Systems.* Gaithersburg, Maryland.

National Institute of Standards and Technology. (2012). *Guide for Conducting Risk Assessments.* Gaithersburg, Maryland.

Neapolitan, R. E. (1989). *Probabilistic Reasoning in Expert Systems: Theory and Algorithms.* Hoboken, New Jersey: Wiley.

Neapolitan, R. E. (2003). *Learning Bayesian Networks.* Upper Saddle River, New Jersey: Prentice Hall.

Neil, M., Fenton, N., & Tailor, M. (2005). Using Bayesian Networks to Model Expected and Unexpected Operational Losses. *Risk Analysis, 25*, 963-972.

Nicholas, J. M., & Steyn, H. (2008). *Project Managment for Business, Engineering, and Technology 3rd Edition.* Oxford: Elsevier Ltd.

Oakley, J. E., & O'Hagan, A. (2004). Sensitivity Analysis of Complex Models: A Bayesian Approach. *Journal of the Royal Statistical Society. Series B (Statistical Methodology), 66*, 751-769.

Olsson, R. (2008). Risk Management in a Multi-Project Environment: an Approach to Manage Portfolio Risks. *International Journal of Quality & Reliability Management, 25*, 60-71.

Pearl, J. (1985). *Bayesian Networks: A Model of Self-Activated Memory for Evidential Reasoning.* Irvine: UCLA.

Pearl, J. (1988). *Probabilistic Reasoning in Intelligent Systems.* San Francisco: Morgan Kaufmann.

Pellegrinelli, S. (1997). Programme Management: Organizing Project-based Change. *International Journal of Project Management, 15*, 141-149.

Penrose, E. G. (1959). *The Theory of the Growth of the Firm.* New York City, New York: Wiley.

Perry, C., & Greig, I. D. (1975). Estimating the Mean and Variance of Subjective Distributions in PERT and Decision Analysis. *Management Science, 21*, 1477-1480.

Peteraf, M. A. (1993). The Cornerstones of Competitive Advantage: A Resource-based View. *Strategic Management Journal, 14*, 179-191.

Porter, M. E. (1980). *Competitive Strategy.* New York: Free Press.

Pourret, O., Naim, P., & Marcot, B. (2008). *Bayesian Networks: A Practical Guide to Applications.* West Sussex: John Wiley & Sons Ltd.

Priem, R. L., & Butler, J. E. (2001). Is the Resource-based "View" a Useful Perspective for Strategic Management Research?". *Academy of Management Review, 26*, 22-40.

Priem, R. L., & Butler, J. E. (2001). Tautology in the Resource-based View and the Implications of Externally Determined Resource Value: Further Comments. *Academy of Management Review, 26*, 57-66.

Project Management Institute. (2008). *Standard for Portfolio Management.* Newtown Square: Project Management Institute.

Project Management Institute. (2010). *A Guide to the Project Management Body of Knowledge.* Newtown Square: Project Management Institute.

Rad, P. F., & Cioffi, D. F. (2000). Work and Resource Breakdown Structures for Formalized Bottom-Up Estimating. *Journal of Cost Engineering, 46*, 31-37.

Roed, W., Mosleh, A., Vinnem, J. E., & Aven, T. (2009). On the Use of the Hybrid Causal Logic Method in Offshore Risk Analysis. *Reliability Engineering and System Safety*, *94*, 445-455.

Rubin, P. H. (1973). The Expansion of Firms. *Journal of Political Economy, 81*, 936-949.

Rumelt, R. P. (1984). Towards a Strategic Theory of the Firm. *Competitive Strategic Management*, *26*, 556-570.

Sanchez, H., Robert, B., & Pellerin, R. (2008). A Project Portfolio Risk-Opportunity Identification Framework. *Project Management Journal, 39*, 97-109.

Shenoy, C., & Shenoy, P. P. (1999). Bayesian Network Models of Portfolio Risk and Return. *Computational Finance, 99,* 87-109.

Sosa, M. E., Eppinger, S. D., & Rowles, C. M. (2003). Identifying Modular and Integrative Systems and Their Impact on Design Team Interactions. *Transactions of the ASME*, *125,* 240-252.

Sosa, M. E., Eppinger, S. D., & Rowles, C. M. (2007). A Network Approach to Define Modularity of Components in Complex Products. *Journal of Mechanical Design, 129*, 1118-1129.

Steward, D. (1981). The Design Structure Matrix: a Method for Managing the Design of Complex Systems. *IEEE Transactions on Engineering Management*, 71-74.

Sylvester, J. J. (1878). Chemistry and Algebra. *Nature, 17*, 284.

Teller, J. (2013). Portfolio Risk Management and Its Contribution to Project Portfolio Success: An Investigation of Organization, Process, and Culture. *Project Management Journal, 44*, 36-51.

Teller, J., & Kock, A. (2013). An Empirical Investigation on How Portfolio Risk Management Influences Project Portfolio Success. *International Journal of Project Management, 31*, 817-829.

Trigeorgis, L. (1993). Real Options and Interactions with Financial Flexibility. *Financial Management, 22*, 202-224.

Vergara, A. J. (1977). *Probabilistic Estimating and Applications of Portfolio Theory in Construction.* Urbana: University of Illinois at Urbana-Champaign.

Vergara, A. J., & Boyer, L. T. (1977). Portfolio Theory: Applications in Construction. *Journal of the Construction Division, 103*, 23-38.

Wade, M., & Hulland, J. (2004). Review: The Resource-based View and Information Systems Research: Review, Extension, and Suggestions for Future Research. *MIS Quarterly, 28*, 107-142.

Weber, P., Medina-Oliva, G., Simon, C., & Iung, B. (2012). Overview on Bayesian Networks Applications for Dependability, Risk Analysis and Maintenance Areas. *Engineering Applications of Artificial Intelligence, 25*, 671-682.

Wernerfelt, B. (1984). A Resource-based View of the Firm. *Strategic Management Journal, 5*, 171-180.

Yescombe, E. R. (2002). *Principles of Project Finance.* San Diego, California: Academic Press.