Available online at www.sciencedirect.com
ScienceDirect
International Journal of Project Management 33 (2015) 1393 – 1404
International Journal of Project Management
www.elsevier.com/locate/ijproman

Performance measurement and the prediction of capital project failure

Hong Long Chen *

Department of Business and Management, National University of Tainan, Tainan 700, Taiwan

Received 26 March 2014; received in revised form 12 February 2015; accepted 22 February 2015
Available online 12 March 2015

Abstract

This paper examines how changes in project-management performance in the execution phase affect project outcomes at completion. While identifying the key determinants of project-management performance is critical, few studies examine the discriminatory power of performance variables for predicting capital project failure at completion. Using 130 capital projects and a longitudinal design, this study develops a performance-measurement model based on changes in project-management performance during the execution phase. Subsequent hierarchical logistic-regression analysis reveals a good explanation of the variation in the failure of capital projects and high classification accuracy. Validating out-of-sample data demonstrates that the optimal model provides a reasonably good overall classification rate of 81.54%. Ultimately, our findings suggest that performance changes in the execution phase explain an important part of project outcomes and, more importantly, are useful predictors for project failure.

Keywords: Project failure; Project-management performance; Performance measurement; Logit analysis; Longitudinal study

1. Introduction

The subject of project management, seen as the basis of a strategic competency in the global economy (Jugdev and Thomas, 2002), has drawn great attention of researchers and practitioners from various disciplines such as management science, operations management, organization theory, and social psychology. A central task in the study of project management is to identify the critical determinants of project management performance. Not surprisingly, researchers and practitioners examine and identify a wide variety of measures to describe project-management performance and the input characteristics that affect project outcomes (e.g., Chen, 2014; El-Sayegh, 2008; Hoegl and Parboteeah, 2007; Oke and Idiagbon-Oke, 2010; Scott-Young and Samson, 2008).

One recent finding, for example, is that management's perception and satisfaction, and project characteristics significantly affect project performance (Lerch and Spieth, 2013). Another finding is that an ambiguous project scope and unclear project goals are the primary risk factors for project performance (Huang and Li, 2012).

Although project-management performance is well-researched and extensively reviewed, most studies are based on the perspective of the overall project life cycle (e.g., Chen, 2014; El-Sayegh, 2008; Oke and Idiagbon-Oke, 2010; Scott-Young and Samson, 2008). Relatively few focus on the perspective of the project execution phase. Some existing studies analyze how project-management performance in the project execution phase affects project outcomes (e.g., Chen, 2014; Hoegl and Parboteeah, 2007; Tabassi and Bakar, 2009), their treatment, however, is contemporaneous in nature. Moreover, little of the focus of these studies' analyses has been on how changes in project-management performance in the execution phase affect project outcomes.

* Tel.: +886 6 260 6123x7741; fax: +886 6 260 6173.
E-mail address: along314@mail.nutn.edu.tw.

http://dx.doi.org/10.1016/j.ijproman.2015.02.009
0263-7863/00/

The first objective of this study, therefore, is to conduct a longitudinal experiment that develops a project-outcome measurement model based on examining the predictive power of changes in project-management performance in the project execution phase. The second objective is to assess, through hierarchical logistic regression analysis, how well the model predicts project failure at completion.

In this study, capital projects served as the basic unit of analysis. The capital project industry includes both the delivery and the maintenance of facilities (e.g., institutional, commercial, and residential buildings; communication, transportation, and energy systems; as well as environmental and industrial facilities). Our focus is on the delivery process of buildings, transportation facilities, environmental facilities, and industrial facilities, e.g., from the initiating to closing phases of projects.[1]

The rest of the paper is organized as follows. Section 2 reviews related studies, Section 3 describes the sample collection and presents the research methodology, and Section 5 depicts the performance-measurement model, and the forecasting-model building and validation. Section 6 discusses the implications of the research results. Section 6 presents the research summary and conclusions.

2. Research background

A paramount question of project-performance management is how to scientifically engineer critical factors and issues to ensure project success. Questions regarding how to manage critical issues systematically and thus enhance project management performance take center stage in the research (Chen, 2011). Naturally, numerous academics and practitioners perform extensive research to develop project management models through examining and identifying the determinants of project-management performance (e.g., Hoang and Rothaermel, 2005; Schwab and Anne, 2008; Scott-Young and Samson, 2008; Wallace et al., 2004).

For example, using the data from 507 software project managers, Wallace et al. (2004) examine the impact of social subsystem risk, technical subsystem risk, and project-management risk on project performance using the structural-equations modeling technique. Their results show that social-subsystem risk significantly affects technical-subsystem risk, which, in turn, influences project-management risk, and ultimately, project performance. Hoang and Rothaermel (2005) use binary logistic analysis to examine the performance of 158 joint research and development (R&D) projects in 43 pharmaceutical firms; they find that the general alliance experience of biotechnology partners, but not of pharmaceutical firms, positively affects joint project performance.

Ho (2006) examines when and how government would initiate a rescue program for a failing project and what the impacts of government's rescue behavior on project procurement and management are. He then develops a game-theory based model to provide theoretic foundations for prescribing effective public–private partnership (PPP) project procurement and management practice. Maytorena et al. (2007) use a combined method of the active information search (AIS) and the cognitive mapping (CM) approach to interview 51 project managers and conclude that information search style, level of education, and risk management training play a significant role in risk identification performance, which in return affects project performance.

Busby and Zhang (2008) examine 13 capital projects using a pathogen approach and concludes that the fundamental causes of project failure are typically decisions, practices, or other basic entities within a project, not external events. Concurrently, Schwab and Anne (2008) examine 239 U.S. movie projects from 1931 to 1940 and determine, using regression analysis, that project performance depends on the perceived relevance of prior performance and on organizational control over project participants.

Subsequent work by Anand et al. (2010) analyzes 98 projects in five companies using hierarchical regression. They show that the inclusion of softer, people-oriented practices for capturing tacit knowledge explains a significant amount of variance in project success. Jani (2011) employs a computer simulation-based experiment to investigate the influence of individual self-efficacy and project risk factors on the perception of risk, using the scenario of a failing IT project. He finds that project managers are more likely to underestimate the risks of a project with endogenous risk factors, and that project managers with higher self-efficacy may underestimate the risks of a troubled IT project as compared to project managers with lower self-efficacy.

Recently, Chou and Yang (2012) use the structural equation modeling (SEM) technique to prioritize the practice of the *PMBOK Guide* in the capital project industry. Based on a sample of 127 project managers and stakeholders, they reveal the interrelationships between the *PMBOK Guide* and project performance. Chen (2013) analyzes 121 capital projects and identifies key variables in the initiation and planning phases of projects that differentiate between healthy and distressed projects at completion. He then demonstrates that it is feasible to discriminate simultaneously between healthy and distressed projects prior to the project execution phase.

Despite the panoply of studies that use a wide variety of measures to describe project-management performance and the input characteristics that affect project outcomes, most studies concentrate on the perspective of the overall project life cycle (e.g., Chen, 2014; El-Sayegh, 2008; Oke and Idiagbon-Oke, 2010; Scott-Young and Samson, 2008; Wallace et al., 2004). For example, Scott-Young and Samson (2008) analyze 56 capital projects in 15 process-industry companies using factor analysis and regression analysis and reveal that organizational context, team leadership, team design, and team process factors

[1] Capital projects have several important characteristics. They (1) involve long-lived assets (e.g., buildings, bridges and roads); (2) typically involve the delivery of a project; (3) usually require long-range planning and extensive financing; and (4) have a project-life focus, rather than a year-to-year focus (Granof and Khumawala, 2012). In addition, one of the major reasons for this study to include a wide range of capital projects is due to the difficulty of data collection. In particular, our research design is longitudinal in nature that considerably increases the difficulty of data collection. To include a wide range of capital projects enables us to maximize our sample size.

are determinants of project outcomes in the overall project life cycle.

Oke and Idiagbon-Oke (2010) examine 93 innovation projects based on the perspective of the overall project life cycle using structural equation modeling and conclude that the richness of communication channels influences project outcomes. Ling et al. (2009) analyze 33 construction capital projects in China and concludes, using Pearson's correlation analysis, that 24 project-management practices (such as effective communication and high customer satisfaction) are significantly correlated with project performance in the overall project life cycle.

Using in-depth data from 29 interviews in 12 companies, Lerch and Spieth (2013) qualitatively examine the impacts of human factors on innovation project performance and conclude that management's perception and satisfaction, and project characteristics significantly affect project performance throughout the project life cycle. Based on the perspective of the overall project life cycle, Chen (2014) analyzes 121 capital projects using hierarchical robust regression analyses. He shows that the relationships among project innovation stimulants, innovation capacity, and project performance are indeed significant.

Some existing studies (e.g., Chen et al., 2010; Hoegl and Parboteeah, 2007; Tabassi and Bakar, 2009) investigate how project-management performance variables in the project execution phase affect project outcomes, these studies, however, are contemporaneously designed in nature. Limited studies examine the potential predictors of project outcomes using longitudinal data.

In particular, while few published studies examine the impact of project-management performance variables in the execution phase on project outcomes using longitudinal data, these studies (e.g., Calamel et al., 2012; Keller, 1992) focus on the longitudinal relationship between team-performance variable and project outcomes. For example, Keller (1986) analyzes 32 project groups in a large R&D organization using a longitudinal design based on hierarchical regression analysis and concludes that group cohesiveness, job satisfaction, and innovative orientation in the execution phase are factors in project outcomes.

Keller (1992) subsequently examines 66 project groups in three industrial R&D organizations using a longitudinal experiment based on regression analysis. He concludes that transformational leadership, the sum of charismatic leadership and intellectual stimulation, significantly affects project outcomes. More recently, Calamel et al. (2012) evaluate two collaborative R&D projects in a large global innovation cluster in France using a longitudinal design. They conclude that team collaboration, which is a product of social construction, is an important factor in project success.

Consequently, there appears to be a lack of research that longitudinally evaluates how changes in various project-management performance variables in the project execution phase affect project outcomes and, thus, predicts the likelihood of project failure at completion. Modeling how performance changes in the project execution phase affect project failure likelihood is important, since it enables management to intervene in the problems of potential failing projects so that corrective actions can be implemented in a timely manner to avert project failure.

3. Research methodology

Now the question is: *How do we build a forecasting model based on changes in project-management performance in the execution phase to predict project failure at completion*? The methods to answer this question are organized as the following.

3.1. Participants and procedures

The survey instrument was designed based on a systematic review of literature in the project-management and organization-theory fields. In particular, prior to the data collection, a panel of experts from the Taiwan's Chinese National Association of General Contractors (CNAGC) critiqued the questionnaire for structure, readability, clarity, and completeness. Based on the feedback from these experts, the survey instrument was then modified to strengthen its validity.

The final version of the survey questionnaire comprises two sections. The first section, composed of open-ended questions, gathers detailed background information such as annual revenue; project type; project cost including contract price, budget, contract price for project changes, and actual cost; and the project schedule including the contract schedule, scheduled time, contract schedule for project changes, and actual schedule.

Section 2 consists of multiple-choice questions in which respondents indicate on a 9-point Likert scale the extent to which certain project-management variables likely affect project performance. (If not otherwise indicated, all measures use a scale in which 1 means "strongly disagree" and 9 means "strongly agree." High scores suggest good performance; low scores indicate poor performance.) Because of space limitations, complete survey questionnaires are not presented here but are available from the authors on request.

Of the 620 members of Taiwan's CNAGC that we randomly selected and invited to participate in this research, 130 companies participated—a 21.00% response rate (CNAGC has over 8000 members). Compared to other similar surveys of the industry, the response rate of 21.00% was considered good.[2] Each of the 130 companies in the sample had assigned a project manager who had just completed the initiation and planning of a capital project scheduled to finish within the next two years.

Data collection occurred in two stages and lasted two years. In the first stage, immediately after the end of a project's initiation and planning stages, participants (project managers) respond to the portion of the questionnaire that excludes questions regarding project actual cost, project actual schedule, contract price and schedule for project changes, and actual cost and schedule for project changes. In the second stage, right before the close of the

[2] Our strategy to promote our survey response rate was based on the data sharing concept (Jalili et al., 2011; Xu, 2007). In other words, we provided the survey participants with the primary source of information and the preliminary research results upon request.

capital project, participants responded to all questions of the questionnaire.

Since the survey questionnaire was given to one respondent per project, single-respondent bias might exist. However, because a project manager is assumed to have a total understanding of his/her project, because we test longitudinal changes in the data, not the raw numbers, and because we examine longitudinal changes in the data comparatively, and not against project outcomes while developing our overall performance-measurement model, the problem of single respondent bias is reduced (though not perfectly resolved).

The 130 capital projects fall into four categories: buildings (54 projects), transportation facilities (20 projects), environmental facilities (19), and industrial facilities (37 projects). Project managers average between one and 30 years of experience; 42 participants had fewer than five years of experience; 45 had between five and 10 years; 28 had between 11 and 20 years; and 15 participants had over 20 years of experience.

3.2. Constructs and measures

We choose project time, cost, and profitability as the criteria for capital project failure. The rationales are straightforward: delays in completion time may turn a promising investment opportunity into an expensive failure (Scott-Young and Samson, 2008), cost overrun directly encroaches on profit (Teerajetgul et al., 2009), and project profitability ensures business growth and development (Chen, 2011). Project time, cost, and profitability are calculated using the following equations based on Anbari (2003) and Hartley and Watt (1981):

$$\text{Time} = \text{Revised Estimated Duration}/\text{Actual Duration} \quad (1)$$

$$\text{Cost} = \text{Revised Estimated Cost}/\text{Actual Cost} \quad (2)$$

$$\text{Profitability} = (\text{Revised Contract Price–Actual Cost})/\text{Revised Contract Price} \quad (3)$$

where the revised estimated duration, revised estimated cost, and revised contract price include the additional estimated duration, additional estimated cost, additional contract price due to changes in project scope, respectively. Thus, our dependent variable, *Project Failure*, is binary, with 1 (failed) indicating that a capital project fails to finish within budget and/or scheduled time frame and/or suffers a loss; otherwise, it is 0 (nonfailed).

Measures of project-management performance variables, including *Communication*, *Team*, *Scope*, *Creativity*, *Technology*, *Risk*, *Quality*, and *Materials* are based on a systematic review of literature, where a systematic review is a literature review concentrated on a research question that tries to identify, appraise, choose and synthesize all relevant research evidence to that question (Adams et al., 2006). Importance to the review process is the use of explicit, reproducible criteria in the selection of articles for review, an appraisal of the quality of the research, and the strength of the findings (Tranfield et al., 2003). We broadly adopt the review methodology detailed by Tranfield et al. (2003) and Colicchia and Strozzi (2012).[3]

Δ *Communication* (Cronbach's Alpha = .766, where Δ = Likert-scale value of performance at the end of the execution phase — Likert-scale value of that performance at the end of the project-initiation and planning phases) is measured according to a nine-item scale based on the representative studies, including Ling et al. (2009), and Oke and Idiagbon-Oke (2010). Sample items are "Co1: The project team identifies all the key stakeholders of the project," and "Co2: The project team meets the communications needs of the stakeholders."

Δ *Team* (Cronbach's Alpha = .904) is measured according to a 10-item scale based on the representative studies, including Anand et al. (2010), Hoegl and Parboteeah (2007), Ling et al. (2009), Scott-Young and Samson (2008), and Tabassi and Bakar (2009). Sample items are "Te1: Sense of motivation for achieving the project's objectives is high," and "Te2: Enthusiasm about project success is high."

Δ *Scope* (Cronbach's Alpha = .770) is measured according to a five-item scale based on the representative studies, including Dumont et al. (1997), Ling et al. (2009), and Roman (1964). Sample items are "Sc1: The scope of project is well defined," and "Sc3: Project owner has verified extent of project scope well."

Δ *Creativity* (Cronbach's Alpha = .735) is measured according to a seven-item scale based on the representative studies, including Amabile and Conti (1999), Keegan and Turner (2002), and Prajogo and Ahmed (2006). Sample items are "Cr1: Top Management support for innovative ideas/solutions is high," and "Cr2: The delegation of autonomy and decision authority to the project manager is high."

Δ *Technology* (Cronbach's Alpha = .925) is measured according to a four-item scale based on the representative studies, including Prajogo and Ahmed (2006), and Urban and von Hippel (1988). Sample items are "Tn1: Our company is on the leading edge of new project practices and technologies in the industry," and "Tn2: We always evaluate the potential of using new project practices and technologies."

Δ *Risk* (Cronbach's Alpha = .793) is measured according to a 13-item scale based on the representative studies, including El-Sayegh (2008), and Luu et al. (2008). Sample items "Ri1: Project team handles inflation and sudden changes in prices well," and "Ri2: Project team handles customer design changes well."

Δ *Quality* (Cronbach's Alpha = .894) is measured according to an 11-item scale based on the representative studies, including Arditi and Gunaydin (1997); Ling et al. (2009), Roman (1964), and Sperpell (1999). Sample items include "Qu1: Every quality metric clearly describes what something is and how to measure it," and "Qu2: Quality baseline is established and well defined."

[3] Our review strategy consists of four steps. The first step forms a review panel composed of domain-relevant academics and experts with an interest in performance measurement and project management. In the second step, we search databases for relevant studies of performance measurement and project management. In the third step, we use the Delphi method, a process composed of a structured design for group communication for resolving complex problems (Linstone and Turoff, 2002), to develop an analytic framework. In the fourth step, we group measures into each construct of the analytic framework.

Table 1
Descriptive statistics, Kolmogorov–Smirnov tests, and Mann–Whitney tests for&& the hypothesis.

Variable	Mean	Minimum	Maximum	Kolmogorov–Smirnov statistic N = 130 [a]	Mann–Whitney statistic N = 130
Δ *Communication*	0.02	−1.33	1.78	0.29**	−0.25
Δ *Team*	0.10	−2.60	0.90	0.30**	1.38
Δ *Scope*	0.04	−3.00	1.40	0.35**	0.34
Δ *Creativity*	0.04	−1.57	2.14	0.38**	2.46*
Δ *Technology*	0.10	−2.00	4.00	0.34**	5.57**
Δ *Risk*	0.08	−1.38	0.85	0.27**	0.12
Δ *Quality*	0.10	−1.18	1.91	0.26**	2.41*
Δ *Materials*	0.10	−3.75	1.00	0.32**	−0.75

*P < .05 and **P < .01.
[a] There are 82 and 48 failed and nonfailed capital projects, respectively.

Δ *Materials* (Cronbach's Alpha = .860) is measured according to a four-item scale based on the representative studies, including El-Sayegh (2008), and Luu et al. (2008). Sample items "Ma1: Material management plan clearly describes how materials will be managed and executed," and "Ma2: Project team effectively keeps a record of material usage and inventory."

Because our research is built on the proposition that there are significant differences in the values of changes in project-management performance during the execution phase between failed and nonfailed projects. We test the following hypothesis:

Hypothesis 1. The mean values of Δ *Communication*, Δ *Team*, Δ *Scope*, Δ *Creativity*, Δ *Technology*, Δ *Risk*, Δ *Quality*, and Δ *Materials* in the execution phase between failed and nonfailed capital projects are all the same.

The 130 capital projects are used to test the hypothesis. Table 1 lists the descriptive statistics and results of Kolmogorov–Smirnov tests, and Mann–Whitney tests for the hypothesis. The use of Mann–Whitney test is justified by the fact that the performance-variable data are not normally distributed based on the result of Kolmogorov–Smirnov tests, where the data are judged abnormally distributed when the probability value is smaller than the threshold value of 0.05.

As the table shows, Mann–Whitney statistics of Δ *Creativity*, Δ *Technology*, and Δ *Quality* are 2.46, 5.57, and 2.41 and with p-values of <0.05, <0.01, and <0.05, respectively. This suggests that significant differences exist in the mean values of Δ *Creativity*, Δ *Technology*, and Δ *Quality* between failed and nonfailed capital projects. We therefore reject the null hypothesis, indicating that the values of changes in project-management performance in the execution phase between failed and nonfailed capital projects are not all the same and, thus, possess a potential discriminatory power for differentiating between failed and nonfailed projects.

3.3. Modeling and analysis

The methodology to develop optimal project-failure prediction models is fourfold. In the first stage, this study performs an isolated model analysis of each dimension of performance construct to evaluate the ability of the set of items to their associated dimension of project-management performance. Using univariate logistic analysis, Δ measure items with an overall correct classification rate lower than 60% are deleted; items with factor loadings smaller than .50 are also deleted. Further deletion of a dimension's item scales for refining the initial measurement instrument is assessed through repeated model fittings based on an examination of standardized loadings, interpretability, and content validity along with a minimum standardized root mean square residual (RMSR) procedure (Frohlich, 2002; Wallace et al., 2004).

In the second stage, this study develops an overall project-outcome measurement model from the refined measurement instrument based on the confirmatory factor analysis (CFA) (Harrington, 2008). Items with factor loadings smaller than .50 are further deleted. In the third stage, this study conducts a hierarchical logistic-regression analysis using a maximum Nagelkerke R-squared improvement procedure to develop optimal project-failure prediction models from the overall project-outcome measurement model.

In the fourth stage, this study evaluates the forecasting accuracy of the prediction models using Type I errors (i.e., a nonfailed project misclassified as a failed project), Type II errors (i.e., a failed project misclassified as a nonfailed project), and overall correct classification rates. Type I errors, Type II errors, and correct classification rates are computed using a rotation estimation (cross-validation) to maximize our holdout sample. Each observation of the 130 capital projects in turn is removed, and the forecasting model is refitted using the remaining observations to predict the removed project.

4. Research results

4.1. Analysis of isolated models

Table 2 lists the item scales ultimately retained for each performance dimension.[4] As seen from the table, the Δ

[4] The first objective of analysis of isolated models is to create a parsimonious performance-measurement model for project-failure forecasting. The second objective is to generate a sufficient condition regarding the sample size for a valid, stable factor solution. There are two general recommendations in terms of minimum sample size in factor analysis. One is the rule of 100, recommending that no sample should be less than 100 (Arrindell and van der Ende, 1985; Gorsuch, 1983). The other is the ratio of sample size to the number of variables. While some researchers suggest a ratio of 2 to 1 (e.g., Kline, 1979), others suggest a more stringent ratio of 5 to 1 (e.g., Gorsuch, 1983; MacCallum et al., 2001) to be a sufficient condition for ensuring the validity and stability of a factor solution. In order to meet the more stringent rule (ratio of 5 to 1), our final performance-measurement model has to contain less than 26 variables.

Table 2
Final survey items included in the performance-measurement model for project-failure forecasting.

Variable	Overall correct [a]	Factor loadings	Measure
Δ *Communication*			
Δ Co1	61.50%	.63	The project team identifies all the key stakeholders of the project
Δ Co2	63.10%	.91	The project team meets the communications needs of the stakeholders
Δ Co3	63.10%	.95	The project team meets the information needs of the stakeholders
Δ Co9	63.10%	.84	Communication within project team members is effective
CR = .79, RMSR = .01, CFI = .98, and TLI = .95			
Δ *Team*			
Δ Te2	63.10%	.83	Enthusiasm about project success is high
Δ Te3	63.10%	.77	Each team member's project role, responsibilities, and rights are clearly defined
Δ Te4	63.10%	.78	Degree of cohesiveness of the project team is high
Δ Te8	63.10%	.88	The interpersonal relationship in the project team is good
CR = .80, RMSR = .01, CFI = .99, and TLI = .98			
Δ *Scope*			
Δ S2	63.10%	.78	Quality of contract documents including project definitions, legal terms, specifications, design instructions, and implementation processes is good
Δ S3	63.10%	.91	Project owner has verified extent of project scope well
Δ S4	62.30%	.88	Work breakdown structure (WBS) of the project is well defined and manageable
CR = .75, RMSR < .01, CFI = .99, and TLI = .99			
Δ *Creativity*			
Δ Cr4	62.30%	.86	Project members work in diversely skilled work groups where free and open communication exists among the group members
Δ Cr5	63.10%	.89	Cognitive conflict among project team members is moderately high
Δ Cr6	63.10%	.62	The project manager adopts a bottom-up problem-solving style that incorporates all team members
CR = .75, RMSR < .01, CFI = .99, and TLI = .99			
Δ *Technology*			
Δ Tn1	79.23%	.75	Our company is on the leading edge of new project practices and technologies in the industry
Δ Tn2	63.10%	.99	We always evaluate the potential of using new project practices and technologies
Δ Tn4	61.50%	.93	We are continuously thinking of the next generation of project technology
CR = .75, RMSR < .01, CFI = .99, and TLI = .99			
Δ *Risk*			
Δ Ri3	63.10%	.88	Project team handles owners' improper intervention after project initiation well
Δ Ri5	61.50%	.76	Project team handles the lack of scope of work definition by owner well
Δ Ri9	63.10%	.76	Project team handles delays in approvals well
CR = .73, RMSR < .01, CFI = .99, and TLI = .99			
Δ *Quality*			
Δ Qu4	62.30%	.59	Quality checklist verifies that a set of required steps is complete
Δ Qu5	60.80%	.80	Project team performs assessment of overall quality performance on a regular basis
Δ Qu6	61.50%	.76	Project team re-evaluates quality standards on a regular basis
CR = .75, RMSR < .01, CFI = .99, and TLI = .99			
Δ *Materials*			
Δ Ma1	63.10%	.71	Material management plan clearly describes how materials will be managed and executed
Δ Ma2	63.10%	.92	Project team effectively keeps a record of material usage and inventory
Δ Ma3	63.10%	.71	The supervision of project team in receiving and delivering materials on sites is well maintained
CR = .75, RMSR < .001, CFI = .99, and TLI = .99			

[a] Δ measures' predictive ability of each performance is assessed by univariate logistic analysis.

measures' classification ability ranges from 60.80% to 79.23%, where Δ Qu5 and Δ Tn1 provide the lowest and highest univariate classification accuracy, respectively. The fit Δ measures for the isolated models based on the refined scales concluded from repeated model fittings using a minimum RMSR procedure show a good fit with the observed data. The RMSRs of the isolated models are all smaller than .10, which indicates an excellent model fit (Harrington, 2008).

The comparative fit indices (CFIs) and the Tucker–Lewis indices (TLIs) of the isolated models range from .98 to .99 and .95 to .99, respectively, and are higher than the threshold value of .90 (Bentler, 1990). The respective composite reliabilities (CR) and standardized factor loadings range from .73 to .80 and from .59 to .99, which are higher than the recommended respective threshold values of .60 and .50 (Fornell and Larcker, 1981; Kline, 2010). We therefore conclude that the item scales

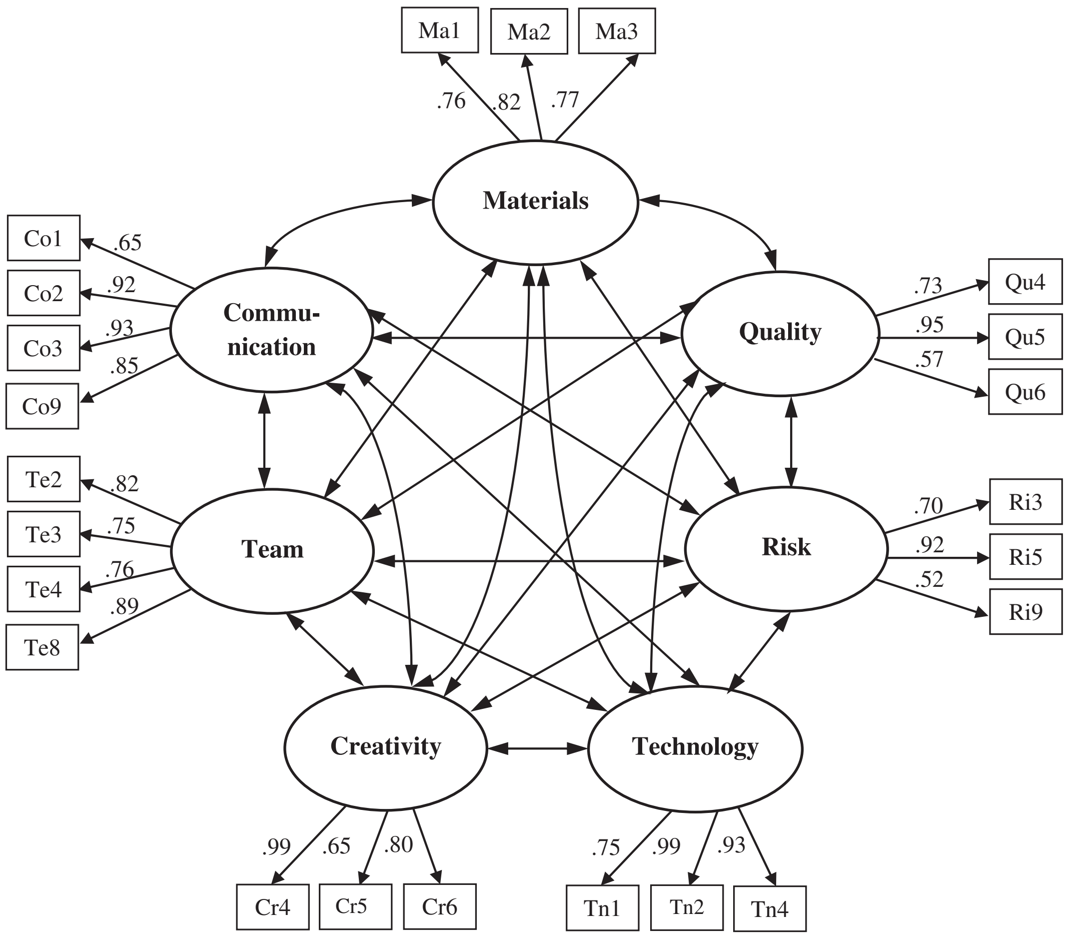

Fig. 1. Performance-measurement model for project-failure forecasting.

provide an adequate and reliable measure of fit for the eight dimensions of the performance-measurement model.

4.2. Analysis of overall performance-measurement model

Fig. 1 depicts the performance-measurement model for project-failure forecasting. The measurement model, which is congeneric, has seven constructs that correlate with all other constructs. We delete Δ *Scope* from the overall performance-measurement model, since it has negative correlations with Δ *Materials* and Δ *Technology*, violating nomological validity for the measurement model.[5] One possible explanation of this phenomenon is that scope performance during project execution is stable, resulting in a non-significant impact of Δ *Scope* on the performance-measurement model. An alternative explanation is that scope performance possesses significant impact on project outcomes mostly during the initiation and planning phases of projects, as is evident from the contents of the measures.[6]

To test for convergent validity, we use the standardized factoring loadings, composite reliability (CR), and average variance extracted (AVE) to evaluate the relative convergence among item measures. High loadings on a factor indicate that they converge on a common point, suggesting high convergent validity (Harrington, 2008). As Fig. 1 shows, all standardized factor loadings range from .52 to .99 and are significant at the $p < .001$ level, suggesting the existence of convergent validity.

[5] Nomological validity is whether the correlations among the constructs in a measurement theory make sense (Hair et al., 2009). Based on prior research (e.g., Dumont et al., 1997; Ling et al., 2009), as *Scope* performance increases, *Materials* and *Technology* performance should not decrease.

[6] While defining a project's scope is a process by which the project is outlined and prepared for execution, a well defined and managed scope is one of the principal tasks in the project-initiation and planning phases (Dumont et al., 1997). A poorly defined scope prior to project execution can experience considerable changes, resulting in cost overruns, delays in the project schedule, rework, disrupted project rhythm, and lower productivity (Dumont et al., 1997). Not surprisingly, several studies (e.g., Chen, 2013; Dumont et al., 1997; Ling et al., 2009; Roman, 1964) suggest that scope performance in the project-initiation and planning phases is one of the leading influences on project outcomes. The contents of the final survey items (measures) retained in *Scope*, including "S2: Quality of contract documents including project definitions, legal terms, specifications, design instructions, and implementation processes is good," "S3: Project owner has verified extent of project scope well," and "S4: Work breakdown structure (WBS) of the project is well defined and manageable," are mostly related to the conditions of scope performance prior to project execution.

Table 3
Squared correlations, average variance extracted, and composite creditability.

Variable	Δ *Communication*	Δ *Team*	Δ *Creativity*	Δ *Technology*	Δ *Risk*	Δ *Quality*	Δ *Materials*
Δ *Communication*	1						
Δ *Team*	.479	1					
Δ *Creativity*	.306	.267	1				
Δ *Technology*	.088	.001	.051	1			
Δ *Risk*	.098	.441	.160	.167	1		
Δ *Quality*	.075	.198	.108	.047	.476	1	
Δ *Materials*	.104	.643	.051	.008	.234	.125	1
Average variance extracted	.714	.651	.681	.803	.545	.587	.614
Composite creditability	.814	.797	.732	.746	.739	.745	.753

The square of a standardized factor loading addresses how much the latent construct explains the variation in an item measure, which is termed the variance extracted of the measure. Thus, an average variance extracted (AVE) of .5 or higher demonstrates adequate convergence. This indicates that, on average, less error remains in the measures than variance explained by the latent construct (Kline, 2010). As the bottom of Table 3 shows, the respective AVE values of Δ *Communication*, Δ *Team*, Δ *Creativity*, Δ *Technology*, Δ *Risk*, Δ *Quality*, and Δ *Materials* are .714, .651, .681, .803, .545, .587, and .614, respectively, revealing an adequate convergence for all the constructs.

Further, composite reliability (CR), computed from the squared sum of factor loadings for a construct divided by the squared sum of factor loadings and the sum of error variance terms for the construct, shows whether the measures consistently represent the same latent construct. While a CR of .6 or higher indicates convergent validity (Kline, 2010; Lee, 2007), the bottom of Table 3 shows that the CR values of Δ *Communication*, Δ *Team*, Δ *Creativity*, Δ *Technology*, Δ *Risk*, Δ *Quality*, and Δ *Materials* are .814, .797, .732, .746, .739, .745 and .753, respectively, confirming an adequate convergence for all the constructs.

To test for discriminant validity, we compare the AVE values for any two constructs with the square of the correlation estimate between the constructs, which is a more rigorous test (Fornell and Larcker, 1981). As seen from Table 3, the AVE value of Δ *Communication* is .714 that is greater than the square of the correlation estimate between Δ *Communication* and any of the other constructs. The AVE value of Δ *Team* is .651 that is greater than the square of the correlation estimate between Δ *Team* and any of the other constructs, and likewise, the AVE values of Δ *Creativity*, Δ *Technology*, Δ *Risk*, Δ *Quality*, and Δ *Materials* are all greater than the square of their respective correlation estimates. This comparison suggests that any of latent constructs in the measurement model explains more of the variance in its item measures than it shares with other latent constructs, providing strong evidence of discriminant validity for the measurement model.

The analysis results of the performance-measurement model suggest an adequate fit with the data. The model chi-square (χ^2)/degrees of freedom = 1.717, which is smaller than the threshold value of 2.000 suggested by Kline (2010); CFI = .960 and TLI = .932 are both higher than the threshold value of .900 suggested by Bentler (1990); and the RMSR = .027 and the root mean square error of approximation (RMSEA) = .075 are both smaller than the respective threshold values of .100 and .080 (Kline, 2010; Lee, 2007).

4.3. The forecasting model

Using hierarchical logistic-regression analyses by a maximum Nagelkerke R-squared improvement procedure, this study further develops the performance-measurement model into optimal project-failure prediction models from the 130 capital projects. Table 4 reports the model-building results.

As the table shows, the optimal project-failure forecasting model at step 1 (Model 1) is the one with the Δ Tn1 variable, where the model deviance between the observed and predicted values of *Project Failure* is 125.70, and 40% of the variation in the *Project Failure* data is explained; the corresponding Type I error, Type II error, and overall correct classification rates are 52.08%, 2.44%, and 79.23%.

At step 2, the optimal project-failure forecasting model (Model 2), composed of the Δ Tn1 and Δ Cr4 variables, explains 42% of the variation in the *Project Failure* data, which is 2% more than that of Model 1; the deviance is reduced from 125.70 to 123.70. The Type I error, Type II error, and overall correct classification rates are 52.08%, 3.66%, and 78.46%, respectively.

The corresponding added variables to Models 3 to 21 are Δ Co1, Δ Co2, Δ Co3, Δ Co9, Δ Te2, Δ Te3, Δ Te4, Δ Te8, Δ Cr5, Δ Cr6, Δ Tec2, Δ Ri3, Δ Ri5, Δ Ri9, Δ Qu4, Δ Qu5, Δ Qu6, Δ Ma2, and Δ Ma3. At step 21, the optimal project-failure forecasting model (Model 21) explains 75% of the variation in the *Project Failure* data, where the model deviance is 68.26. The respective Type I error, Type II error, and overall correct classification rates are 20.83%, 4.88%, and 89.23%.

As the Nagelkerke R-squared values are not improved by adding more variables after Model 21, this study selects Model 21 as the optimal project-failure forecasting model. The final performance-measurement model for project-failure forecasting therefore excludes the Ma1 and Tn4 measures.

The analysis results of this final performance-measurement model (see Fig. 2), excluding Ma1 and Tn4, further indicate an improved fit with the data. The model chi-square (χ^2)/degrees of freedom reduces from 1.717 to 1.705; the respective CFI and

TLI increases from .960 to .964 and from .932 to .934; and the RMSR and RMSEA reduces from .027 to .026 and from .075 to .074, respectively. The general form of our project-failure forecasting model (Model 21) is summarized as follows:

$$P(\mathit{failed}) = \Delta \mathit{Communication} + \Delta \mathit{Team} + \Delta \mathit{Creativity} + \Delta \mathit{Technology} + \Delta \mathit{Risk} + \Delta \mathit{Quality} + \Delta \mathit{Materials} \quad (4)$$

where

P(*failed*) the odds of project failure
Δ *Communication* changes in *Communication* measured by Δ Co1, Δ Co2, Δ Co3, and Δ Co9
Δ *Team* changes in *Team* measured by Δ Te2, Δ Te3, Δ Te4, and Δ Te8
Δ *Creativity* changes in *Creativity* measured by Δ Cr4, Δ Cr5, and Δ Cr6
Δ *Technology* changes in *Technology* measured by Δ Tn1 and Δ Tn2
Δ *Risk* changes in *Risk* measured by Δ Ri3, Δ Ri5, and Δ Ri9
Δ *Quality* changes in *Quality* measured by Δ Qu4, Δ Qu5, and Δ Qu6
Δ *Materials* changes in *Materials* measured by Δ Ma2 and Δ Ma3

Table 4
Project-failure forecasting models created with the hierarchical logistic regression using a maximum Nagelkerke R-squared improvement.

Variables and sources	Model 1		Model 2		…	Model 21	
	B	S.E.	B	S.E	…	B	S.E
Intercept	1.01**	.23	1.09**	.23	…	1.24**	.42
Step 1: Δ Tn1	3.50**	.77	3.54**	.77	…	3.85	1.07
Step 2: Δ Cr4			−.74	.61	…	−1.94	2.70
Step 3: Δ Co1					…	6.05	3.27
Step 4: Δ Co2					…	−3.76	2.67
Step 5: Δ Co3					…	.90	1.35
Step 6: Δ Co9					…	6.24	3.41
Step 7: Δ Te2					…	4.08	2.60
Step 8: Δ Te3					…	2.43	1.71
Step 9: Δ Te4					…	−3.02	1.98
Step 10: Δ Te8					…	.95	1.46
Step 11: Δ Cr5					…	−28.09**	10.84
Step 12: Δ Cr6					…	8.98	5.07
Step 13: Δ Tec2					…	14.89*	6.19
Step 14: Δ Ri3					…	−1.05	1.49
Step 15: Δ Ri5					…	−9.45	4.28
Step 16: Δ Ri9					…	2.16	1.53
Step 17: Δ Qu4					…	1.33	1.61
Step 18: Δ Qu5					…	4.01	2.51
Step 19: Δ Qu6					…	8.33*	3.59
Step 20: Δ Ma2					…	−7.33*	3.43
Step 21: Δ Ma3					…	−3.15	1.66
					…		
−2 Log likelihood	125.70		123.70		…	68.26	
Nagelkerke R^2	.40		.42		…	.75	
Type I Error (%)	52.08		52.08			20.83	
Type II Error (%)	2.44		3.66			4.88	
Overall Correct (%)	79.23		78.46			89.23	

*$P < .05$ and **$P < .01$.

4.4. Validation of the forecasting model

This study utilizes a rotation estimation (cross-validation) to maximize our holdout sample. Each observation of the 130 capital projects in turn is removed, and the forecasting model is refitted using the remaining observations to predict the removed project. This rotation produces a holdout sample that contains all the 130 capital projects. The holdout-sample forecasting accuracy is then calculated for the omitted observation in each case; Type I error, Type II error, and overall correct classification rates are used to summarize the model' predictive power.

Table 5 reports the holdout-sample forecasts from the forecasting model. As seen in the table, the respective average Type I error, Type II error, and correct classifications of the holdout-sample forecast data in the model are 27.08%, 13.41%, and 81.54%, respectively. Those for the within-sample data in the forecasting model (see Table 4) are 20.83%, 4.88%, and 89.23%, respectively. The relatively small difference in the average classification rates for the within-sample data and the holdout-sample forecast data in the model suggests that the performance-measurement model for project-failure forecasting is reasonably stable.

5 . Discussion

Our findings regarding the important effects of *Communication*, *Team*, *Creativity*, *Technology*, *Risk*, *Quality*, and *Materials* on project outcomes are consistent with prior studies (e.g., Amabile and Conti, 1999; Chen, 2014; El-Sayegh, 2008; Hoegl and Parboteeah, 2007; Urban and von Hippel, 1988). The present research extends the state of knowledge concerning the predictive power of Δ *Communication*, Δ *Team*, Δ *Creativity*, Δ *Technology*, Δ *Risk*, Δ *Quality*, and Δ *Materials* for project failure.

The optimal project-failure forecasting model (Model 21) suggests that Δ *Communication*, Δ *Team*, Δ *Creativity*, Δ *Technology*, Δ *Risk*, Δ *Quality*, and Δ *Materials* affect project outcomes. That is, changes in communication performance during the execution phase predict project failure at completion.[7] Likewise, changes in team, creativity, technology, risk, quality, and/or materials performance predict project failure at completion. Above all, changes in *Technology*

[7] The concept behind this application is simple. The inputs to develop the project-failure forecasting model are changes in project-management performance (Δ), where Δ = Likert-scale value of a performance measure at the end of the execution phase — Likert-scale value of that measure at the end of the project-initiation and planning phases. During the phase of project execution, we can, however, compute Δ by using up-to-date Likert-scale value of a performance measure minus Likert-scale value of that measure at the end of the project-initiation and planning phases for estimating the likelihood of project failure at completion. Of course, as the project progresses, and if each of the up-to-date Likert-scale values of performance measures in the project execution phase stays constant until the end of the phase, the likelihood of project failure will stay constant throughout the execution phase.

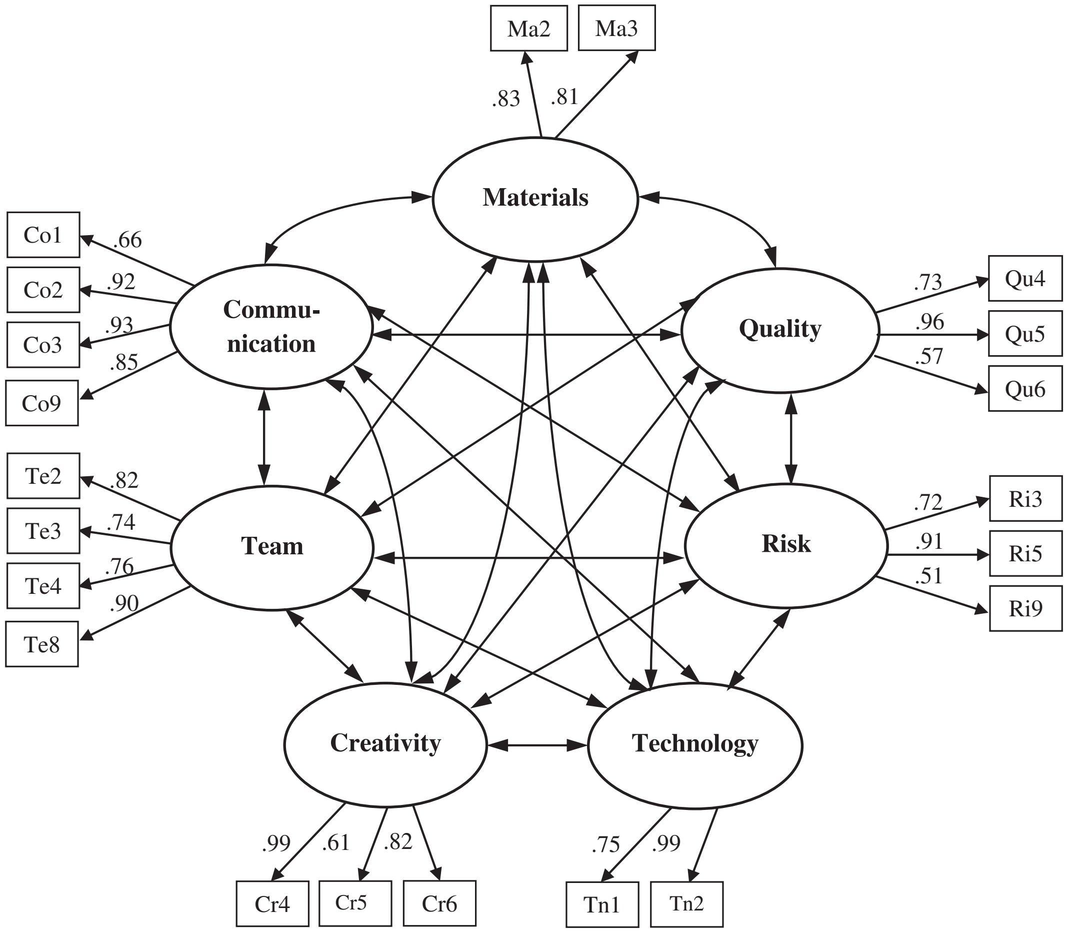

Fig. 2. Final performance-measurement model for project-failure forecasting.

possess the highest discriminatory power for project outcomes. As such, practitioners can employ the forecasting model to monitor and predict project outcomes during the execution phase.

The modeling procedure presented in this paper, combining the performance-measurement theory and hierarchical logistic regression analysis, appears to be robust with straightforward application and, thus, could be extended to a broader analysis of model-building, such as firm bankruptcy and risk forecasting. As such, both academics and practitioners can directly apply the methods in this paper. A more subtle benefit to practitioners is the reminder that even with sophisticated statistical tools and detailed data, project failure forecasting is likely to remain inaccurate. Assessments of the prediction methods and of the measures of a performance construct(s) remain important components for management decision making.

In particular, although the project-failure forecasting model is based on a robust modeling procedure and an extensive review of literature for the choice of performance constructs (e.g., Adams et al., 2006; Anand et al., 2010; Hoegl and Parboteeah, 2007; Scott-Young and Samson, 2008; Tabassi and Bakar, 2009), our model is, however, unable to predict project failure with 100% or nearly 100% accuracy. One possible reason is that the model does not account for project complexity, technological uncertainty, unexpected geological features, negative plurality, etc. in the phase of project development (Flyvbjerg et al., 2003; Miller and Lessard, 2000; Morris and Pinto, 2004).

A more innovative explanation may be that during project development and decision-making, project planners systematically underestimate risks of complexity, scope changes, technological uncertainty, etc. (Flyvbjerg, 2013; Flyvbjerg et

Table 5
Holdout-sample forecasts by the optimal project-failure forecasting model.

Variables	Predicted results			
	Project outcomes		Percentage correct	Error types
	Nonfailed (0)	Failed (1)		
Nonfailed (0)	35	13	72.92	Type I, 27.08
Failed (1)	11	71	86.59	Type II, 13.41
Overall Percentage			81.54	

Note: *Project Failure*, is binary, with 1 indicating that a capital project fails to finish within budget and scheduled time frame and suffers a loss; otherwise, it is 0 (nonfailed).

al., 2009).[8] In other words, optimism bias during project preparation may account for the prediction variance (Flyvbjerg, 2007; Flyvbjerg, 2014).

6. Conclusion

Extensive studies in the project-management field examine and identify a wide variety of measures that describe project-management performance and the inputs that affect that project outcomes. However, few studies investigate how changes in performance in the project execution phase affect project outcomes and, hence, predict the likelihood of project failure at completion. This study therefore develops optimal project-failure forecasting models for capital projects by modeling Δ *Communication*, Δ *Team*, Δ *Scope*, Δ *Creativity*, Δ *Technology*, Δ *Risk*, Δ *Quality*, and Δ *Materials* variables using the combined performance-measurement theory and hierarchical logistic regression analysis.

Analysis of the performance-measurement theory based on CFA reveals that including Δ *Scope* in the performance-measurement model causes a violation of nomological validity. We therefore deleted Δ *Scope* from the measurement model. Subsequent analyses of hierarchical logistic regressions on the longitudinal data demonstrate that a combination of Δ *Communication*, Δ *Team*, Δ *Creativity*, Δ *Technology*, Δ *Risk*, Δ *Quality* and Δ *Materials* provides the highest overall classification accuracy for estimation and holdout samples (89.23% and 81.54%, respectively).

By adding to the benefits of existing project-performance models that provide management with a wide variety of measures that describe project performance (e.g., Anand et al., 2010; El-Sayegh, 2008; Hoegl and Parboteeah, 2007; Urban and von Hippel, 1988), this paper contributes a methodology capable of providing predicted project outcomes prior to project completion. Thus, this paper addresses part of the fundamental improvement of project-performance models.

Extending the research in this paper to study performance measurement of different project types (e.g., R&D, NPD, and capital projects) in response to longitudinal changes in project-management performance not only will improve knowledge of how performance measures behave and why some measures are only adequate for certain types of projects, but will provide management a comprehensive picture of how the predictive power of these variables varies under different types of projects. Future research should also consider collecting data from multiple individuals from a project to avoid possible single-respondent bias.

Conflict of interest

There is no conflict of interest.

[8] One thing we need to point out is that in Flyvbjerg's (2014) break–fix model of megaprojects, he defines success in megaproject management as projects being delivered on budget, on time, and with the promised benefits based on the original estimations, not on the end results. Our definition of project failure/success, on the other hand, is based on revised estimations. Please see Eqs. (1)–(3) for more details about the definition of project failure/success.

References

Adams, R., Bessant, J., Phelps, R., 2006. Innovation management measurement: a review. Int. J. Manag. Rev. 8 (1), 21–47.

Amabile, T.M., Conti, R., 1999. Changes in the work environment for creativity during downsizing. Acad. Manag. J. 42 (6), 630–640.

Anand, G., Ward, P.T., Tatikonda, M.V., 2010. Role of explicit and tacit knowledge in Six Sigma projects: an empirical examination of differential project success. J. Oper. Manag. 28 (4), 303–315.

Anbari, F.T., 2003. Earned value project management method and extensions. Proj. Manag. J. 34 (4), 12–23.

Arditi, D., Gunaydin, H.M., 1997. Total quality management in the construction process. Int. J. Proj. Manag. 15 (4), 235–243.

Arrindell, W.A., van der Ende, J., 1985. Cross-sample invariance of the structure of self-reported distress and difficulty in assertiveness: experiences with the scale for interpersonal behaviour. Adv. Behav. Res. Ther. 7, 205–243.

Bentler, P.M., 1990. Comparative fit indexes in structural models. Psychol. Bull. 107 (2), 238–246.

Busby, J.S., Zhang, H., 2008. The pathogen construct in risk analysis. Proj. Manag. J. 39 (3), 86–96.

Calamel, L., Defélixa, C., Picqd, T., Retour, D., 2012. Inter-organisational projects in French innovation clusters: the construction of collaboration. Int. J. Proj. Manag. 30 (1), 48–59.

Chen, H.L., 2011. An empirical examination of project contractors' supply-chain cash flow performance and owners' payment patterns. Int. J. Proj. Manag. 29 (5), 604–614.

Chen, H.L., 2013. Logit models for early warning of distressed capital projects. J. Bus. Econ. Manag. 14 (sup1), 145–167.

Chen, H.L., 2014. Innovation stimulants, innovation capacity, and the performance of capital projects. J. Bus. Econ. Manag. 15 (2), 212–231.

Chen, W.T., Chang, P.Y., Huang, Y.H., 2010. Assessing the overall performance of value engineering workshops for construction projects. Int. J. Proj. Manag. 28 (5), 514–527.

Chou, J.S., Yang, J.G., 2012. Project management knowledge and effects on construction project outcomes: an empirical study. Proj. Manag. J. 43 (5), 47–67.

Colicchia, C., Strozzi, F., 2012. Supply chain risk management: a new methodology for a systematic literature review. Supply Chain Manag. Int. J. 17 (4), 403–418.

Dumont, P., Gibson, E., Fish, J., 1997. Scope management using project definition rating index. J. Manag. Eng. 13 (5), 54–60.

El-Sayegh, S.M., 2008. Risk assessment and allocation in the UAE construction industry. Int. J. Proj. Manag. 26 (4), 431–438.

Flyvbjerg, B., 2007. Policy and planning for large-infrastructure projects: Problems, causes, cures. Environ. Plan. B Plan. Des. 34 (4), 578–597.

Flyvbjerg, B., 2013. Quality control and due diligence in project management: getting decisions right by taking the outside view. Int. J. Proj. Manag. 31 (5), 760–774.

Flyvbjerg, B., 2014. What you should know about megaprojects and why: an overview. Proj. Manag. J. 45 (2), 6–19.

Flyvbjerg, B., Bruzelius, N., Rothengatter, W., 2003. Megaprojects and Risk: An Anatomy of Ambition. Cambridge University Press, Cambridge, England.

Flyvbjerg, B., Garbuio, B., Lovallo, D., 2009. Delusion and deception in large infrastructure projects: two models for explaining and preventing executive disaster. Calif. Manag. Rev. 51 (2), 170–193.

Fornell, C., Larcker, D.F., 1981. Evaluating structural equation model with unobservable variables and measurement error. J. Mark. Res. 18 (1), 39–50.

Frohlich, M.T., 2002. E-integration in the supply chain: barriers and performance. Decis. Sci. 33 (4), 537–556.

Gorsuch, R.L., 1983. Factor Analysis. Lawrence Erlbaum Associates, New Jersey.

Granof, M.H., Khumawala, S.B., 2012. Government and Not-for-Profit Accounting: Concepts and Practices. John Wiley & Sons, New York.

Hair, J.F., Black, W.C., Babin, B.J., Anderson, R.E., 2009. Multivariate Data Analysis. Prentice Hall, Hoboken, NJ.

Harrington, D., 2008. Confirmatory Factor Analysis. Oxford University Press, New York, NY.

Hartley, K., Watt, P.A., 1981. Profits, regulation and the UK aerospace industry. J. Ind. Econ. 29 (4), 413–428.

Ho, S.P., 2006. Model for financial renegotiation in public–private partnership projects and its policy implications: game theoretic view. J. Constr. Eng. Manag. 132 (7), 678–688.

Hoang, H.T., Rothaermel, F.T., 2005. The effect of general and partner-specific alliance experience on joint R&D project performance. Acad. Manag. J. 48 (2), 332–345.

Hoegl, M., Parboteeah, K.P., 2007. Creativity in innovative projects: how teamwork matters. J. Eng. Technol. Manag. 24 (1–2), 148–166.

Huang, J.W., Li, Y.H., 2012. Slack resources in team learning and project performance. J. Bus. Res. 65 (3), 381–388.

Jalili, H., Amirzadeh, A., Noorafkan, A., Jalili, G., 2011. A new evolutionary algorithm based on data sharing concept for solving optimization problems. Sci. Res. Essays 6 (23), 5082–5090.

Jani, A., 2011. Escalation of commitment in troubled IT projects: influence of project risk factors and self-efficacy on the perception of risk and the commitment to a failing project. Int. J. Proj. Manag. 29 (7), 934–945.

Jugdev, K., Thomas, J., 2002. Project management maturity models: the silver bullets of competitive advantage. Proj. Manag. J. 33 (4), 4–14.

Keegan, A., Turner, J.R., 2002. The management of innovation in project-based firms. J. Long Range Plan. 35 (4), 367–388.

Keller, R.T., 1986. Predictors of the performance of project groups in R&D organizations. Acad. Manag. J. 29 (4), 715–726.

Keller, R.T., 1992. Transformational leadership and the performance of research and development project groups. J. Manag. 18 (3), 489–501.

Kline, P., 1979. Psychometrics and Psychology. Academic Press, London.

Kline, R.B., 2010. Principles and Practice of Structural Equation Modeling. The Guilford Press, New York.

Lee, S.Y., 2007. Structural Equation Modelling: A Bayesian Approach. John Wiley & Sons, West Sussex.

Lerch, M., Spieth, P., 2013. Innovation project portfolio management: a qualitative analysis. IEEE Trans. Eng. Manag. 60 (1), 18–29.

Ling, F.Y.Y., Low, S.P., Wang, S.Q., Lim, H.H., 2009. Key project management practices affecting Singaporean firms' project performance in China. Int. J. Proj. Manag. 27 (1), 59–71.

Linstone, H.A., Turoff, M., 2002. The Delphi method: techniques and applications. http://www.is.njit.edu/pubs/delphibook.

Luu, V.T., Kim, S.Y., Huynh, T.A., 2008. Improving project management performance of large contractors using benchmarking approach. Int. J. Proj. Manag. 26 (7), 758–769.

MacCallum, R.C., Widaman, K.F., Preacher, K.J., Hong, S., 2001. Sample size in factor analysis: the role of model error. Multivar. Behav. Res. 36, 611–637.

Maytorena, E., Winch, G.M., Freeman, J., Kiely, T., 2007. The influence of experience and information search styles on project risk identification performance. IEEE Trans. Eng. Manag. 54 (2), 315–326.

Miller, R., Lessard, D.R., 2000. The strategic management of large engineering projects: shaping institutions, risks, and governance. MIT Press, Cambridge, MA.

Morris, P.W.G., Pinto, J.K., 2004. The Wiley Guide to Managing Projects. Wiley, Hoboken, NJ.

Oke, A., Idiagbon-Oke, M., 2010. Communication channels, innovation tasks and NPD project outcomes in innovation-driven horizontal networks. J. Oper. Manag. 28 (5), 442–453.

Prajogo, D.I., Ahmed, P.K., 2006. Relationships between innovation stimulus, innovation capacity, and innovation performance. R&D Manag. 36 (5), 499–515.

Roman, D., 1964. Project management recognizes R&D performance. Acad. Manag. J. 7 (1), 7–20.

Schwab, A., Anne, S.M., 2008. Learning in hybrid-project systems: the effects of project performance on repeated collaboration. Acad. Manag. J. 51 (6), 1117–1149.

Scott-Young, C., Samson, D., 2008. Project success and project team management: evidence from capital projects in the process industries. J. Oper. Manag. 26 (6), 749–766.

Sperpell, A., 1999. Integrating quality systems in construction projects: the Chilean case. Int. J. Proj. Manag. 17 (5), 317–322.

Tabassi, A.A., Bakar, A.H.A., 2009. Training, motivation, and performance: the case of human resource management in construction projects in Mashhad, Iran. Int. J. Proj. Manag. 27 (5), 471–480.

Teerajetgul, W., Chareonngam, C., Wethyavivorn, P., 2009. Key knowledge factors in Thai construction practice. Int. J. Proj. Manag. 27 (8), 833–839.

Tranfield, D., Young, M., Partington, D., Bessant, J., Sapsed, J., 2003. Knowledge management routines for innovation projects: developing a hierarchical process model. Int. J. Innov. Manag. 7 (1), 27–49.

Urban, G.L., von Hippel, E., 1988. Lead user analyses for the development of new industrial products. Manag. Sci. 34 (5), 569–582.

Wallace, L., Keil, M., Rai, A., 2004. How software project risk affects project performance: an investigation of the dimensions of risk and an exploratory model. Decis. Sci. 35 (2), 289–321.

Xu, G.H., 2007. Open access to scientific data: promoting science and innovation. Data Sci. J. 6 (17), 21–25.