



Journal of Financial Management of Property and Construction

Emerald Article: Regression modelling of risk impacts on construction cost flow forecast

Henry Odeyinka, John Lowe, Ammar Kaka

Article information:

To cite this document: Henry Odeyinka, John Lowe, Ammar Kaka, (2012), "Regression modelling of risk impacts on construction cost flow forecast", Journal of Financial Management of Property and Construction, Vol. 17 Iss: 3 pp. 203 - 221

Permanent link to this document:

<http://dx.doi.org/10.1108/13664381211274335>

Downloaded on: 02-11-2012

References: This document contains references to 40 other documents

To copy this document: permissions@emeraldinsight.com

Access to this document was granted through an Emerald subscription provided by Emerald Author Access

For Authors:

If you would like to write for this, or any other Emerald publication, then please use our Emerald for Authors service. Information about how to choose which publication to write for and submission guidelines are available for all. Please visit www.emeraldinsight.com/authors for more information.

About Emerald www.emeraldinsight.com

With over forty years' experience, Emerald Group Publishing is a leading independent publisher of global research with impact in business, society, public policy and education. In total, Emerald publishes over 275 journals and more than 130 book series, as well as an extensive range of online products and services. Emerald is both COUNTER 3 and TRANSFER compliant. The organization is a partner of the Committee on Publication Ethics (COPE) and also works with Portico and the LOCKSS initiative for digital archive preservation.

*Related content and download information correct at time of download.



Regression modelling of risk impacts on construction cost flow forecast

Regression modelling of risk impacts

203

Henry Odeyinka

School of the Built Environment, University of Ulster – Jordanstown, Belfast, UK

John Lowe

School of Engineering and Built Environment, Glasgow Caledonian University, Glasgow, UK, and

Ammar Kaka

Heriot Watt University, Dubai Campus, Dubai Academic City, Dubai, United Arab Emirates

Abstract

Purpose – Significant risk factors inherent in construction cost flow forecast were identified in this study. The aim of this paper is to develop regression models to assess the impacts of the identified risks on the baseline forecast at the in-progress stage of construction.

Design/methodology/approach – Two stages were involved in data collection. The first was a structured questionnaire survey administered on 370 UK contractors to identify significant risk factors inherent in cost flow forecast. The second stage was the collection of forecast and actual cost flow data from 55 case study projects. Variations between these pair of data sets were measured at 30 per cent, 50 per cent, 70 per cent and 100 per cent completion periods. Respondents were then requested to score on a Likert type scale, the extent of occurrence of the significant risk factors in the case study projects. This pair of data sets were used in regression modelling.

Findings – Significant risk factors were identified from the questionnaire survey analysis as: changes to initial design, variation to works, production target slippage, delay in agreeing variation/dayworks and delay in settling claims among others. Using the identified significant risk factors and the periodic variability measurements, multiple linear regression models were developed. The models were promising in that they helped to establish the fact that the phenomenon under consideration could be modelled. They also provided some insights in explaining the observed variability between the baseline cost flow forecast and actual cost flow based on risk impacts.

Research limitations/implications – The developed models showed a promising level of accuracy but also indicated that the phenomenon under consideration is not strictly linear and may need to explore some other form of modelling.

Practical implications – The developed models provide invaluable information to the construction contractors regarding the likely impacts of significant risk variables on cost flow baseline forecast at different stages of construction so that a pro active risk response can be put in place.

Originality/value – This study makes an original contribution of providing a modelling insight into the phenomenon of how risks inherent in construction could impact the baseline cost flow forecast at different stages of construction. The information is invaluable in making pro active risk response.

Keywords Construction projects, Cost flow, Contractors, Regression modelling, Risk factors, United Kingdom, Risk management, Construction industry, Costs

Paper type Research paper



Introduction

Cash flow forecasting is undoubtedly very important for the survival of any construction company as companies tend to fail due to inadequate attention to proper cash flow forecasting. The detailed approach to cash flow prediction usually involves the breakdown of the bill of quantities (BOQ) in line with the contract programme to produce an estimated expenditure profile (the schedule-based method). This could be expected to be reasonably precise provided that the BOQ is accurate and the contract programme is complied with. However, this method could be cumbersome and time consuming and dependent on the availability of the BOQ and contract programme. Over the past four decades, researchers have engaged in devising short cut approaches that will remove the tedium from cash flow forecasting. The short cut approach, also known as the cost profile method (Kenley, 2003) have been devised to predict the client cash flow as well as the contractor's cash flow. On the construction client's side, we have models developed by Hudson (1978), Berny and Howe (1983), Khosrowshahi (1991) and Khosrowshahi and Kaka (2007) among others. On the contractor's side which is the focus of this study, essentially, three approaches have been used:

- (1) Based on contract value.
- (2) Based on cost commitment (cost flow).
- (3) Based on net cash flow.

Kaka and Price (1991) have demonstrated that the net cash flow curve was difficult and unreliable to model. Evans and Kaka (1998) also demonstrated that with very detailed classification of construction projects, it was very difficult to model the value curve accurately. However, Kaka and Price (1993) demonstrated that cost flow curves are more reliable to model. Using the cost flow short cut method, Kaka and Price (1993) modelled "ideal" average curves based on cost flow data from several projects with similar duration categorisations. Whilst the developed models were statistically proven to have very good predictive ability, it is noteworthy that they are more useful to the construction contractor at the tender stage.

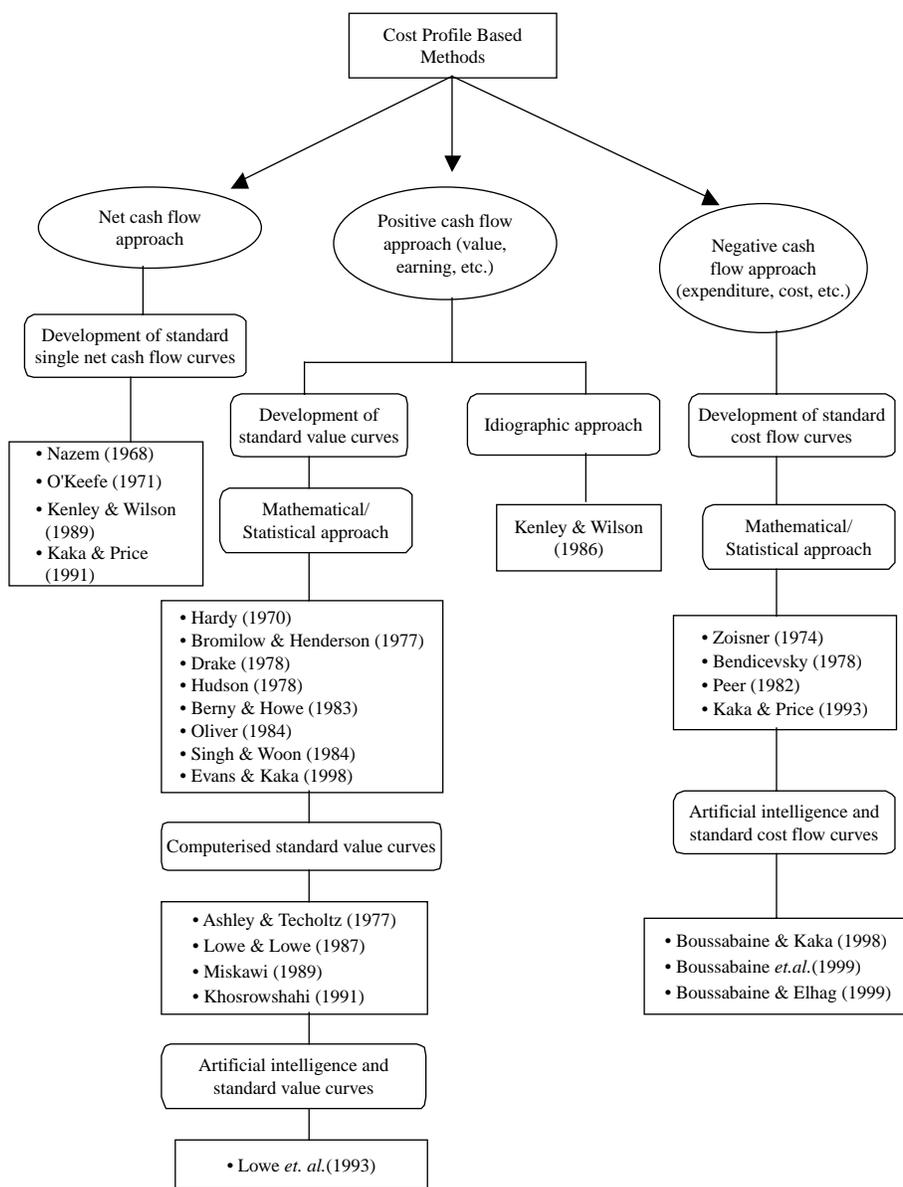
This present study acknowledges the usefulness of the average cost flow curve as a base line forecast at the commencement of a project but contends that the cost flow profile is subject to change during the in-progress construction phase due to risk factors eventuating during the progress of the project. The study therefore attempts to answer two questions:

- (1) What are the significant risk factors impacting the variability between the base line cost flow forecast and the out turn cost flow?
- (2) How can the impacts of the identified significant risk factors on the base line forecast be modelled to assist in predicting risk impacts on cost flow forecast in future projects?

Construction cash flow forecasting perspectives

Previous attempts have been made by researchers to develop short cut approaches to cash flow forecasting to help the construction contractors to take the tedium out of the detailed approach to cash flow forecasting. The majority of the developed models are based on project's financial data and construction duration and are referred to as the cost profile method (Kenley, 2003). Utilising the three sides of the cash flow equation,

the models are based on one of three approaches: net cash flow, contract value data (value flows) and construction cost data (cost flow or cost commitment). The approach was based on developing standard curves which Kenley and Wilson (1986) referred to as the nomothetic approach. Figure 1 details the methods adopted under each approach.



Source: Odeyinka (2003)

Figure 1. Conceptual classification of models developed using cost-profile based methods

Kenley and Wilson (1989) attempted to develop net cash flow model by collecting the cash out and cash in data from 26 commercial and industrial projects. The goodness of fit was reasonably accurate and 26 net cash flow profiles were produced using the technique of logit transformation. Comparison between the results however indicated that there was a wide degree of variation between the profiles of individual projects, suggesting the unreliability of developing standard net cash flows. Kaka and Price (1991) also attempted to develop a net cash flow model to be used by contractors at the tendering stage. The model, based on cost commitment curves instead of the usual value curves exploited the possibility of building an ideal cost curve by building an average curve from the available projects. The average curve was then used to forecast net cash flows for five projects. The study however concluded that the net cash flows analysed tended to vary widely. Comparing these two attempts suggests that an ideal or standard net cash flow was not possible.

A lot of research effort has also gone into the possibility of developing an ideal value curve based on historic data. Results have been obtained by fitting selected functions (mostly polynomial regression or mathematical formulae) to the available data. These approaches although have gained general acceptance; they have not been without criticism. Hardy (1970) using 25 building projects of different types demonstrated that within a company, there is some appropriate form of S-curve geometry for specific types of work when the cumulative value versus time were expressed in percentage terms. Although he also found that there was no close correlation between the values considered even when separating the projects into different categories. Bromilow and Henderson (1977) used four general building projects to develop their value S-curve. Drake (1978) collected projects from regional health authorities and further classified them into different cost categories. He fitted an S-curve into each of these categories. Unfortunately, no figures were published of the number of projects analysed or of the level of accuracy of the fitted functions.

Hudson (1978) developed an expenditure-forecasting model for hospital buildings for the Department of Health and Social Security (DHSS). The model based on mathematical equation forecasts the value S-curves and construction duration and was obtained by using the cumulative monthly values of work executed before the deduction of any retention monies or addition of fluctuations. These data were expressed as a percentage of the contract sum and plotted against percentage of contract period. The collected data were split into different cost categories and in each category, a line of best fit was drawn, and the resulting S-curve graphs of rate of expenditure were issued as guidance. Hudson's (1978) model which was developed for the construction client generated a lot of research interests with other researchers attempting to validate its predictive power (Sidwell and Rumball, 1982) and others attempting to improve on its predictive accuracy (Berny and Howe, 1982). Others have also criticised it as lacking theoretical basis (Lowe, 1987).

Berny and Howe (1983) modified the wholly nomothetic model by adapting the Hudson's (1978) category model to a form which could reflect the specific form of individual projects. They designed methods for calculating the specific curve for a given project, based on their general equation. Kenley and Wilson (1986) argued that by proposing an equation for the general case of an individual project curve as distinct from the curve of the general (standard) function, Berny and Howe's (1983) model moved from a nomothetic to an idiographic approach. Kenley and Wilson (1986) applied the idiographic methodology further by analysing 72 commercial and

industrial building projects in two groups of data using the technique of logit transformation. The technique is the simplest of the sigmoid transformations and allows the S-curve to be represented in linear form (Kenley and Wilson, 1986). Kenley and Wilson (1986) developed a value S-curve for each individual project and an average one for each of the two groups. The error obtained from the two average curves was found to be much higher than that of the individual fits. This, according to them meant that the systematic error involved in group regression was high and the individual curves took a unique shape. They therefore concluded that it was their belief that group models are both functionally as well as conceptually in error. Kaka (1990) however argued that it was important to acknowledge that the conditions required for predictive purposes were not fulfilled by the idiographic methodology. He observed that the conclusion reached by Kenley and Wilson (1986) may have repercussion within the industry as nomothetic models have achieved widespread support, especially where forecasting is concerned. Kaka and Price (1993) also maintained that it is important to acknowledge that idiographic models are only useful for analytical purposes.

Evans and Kaka (1998) used historical data of 20 projects to develop standard value curves. Attempt was made to improve the accuracy of the standard value curves by selecting a more specific type of building, i.e. food retail building projects. The technique of logit transformation was used to fit the value curves. However, results demonstrated that an accurate standard S-curve was not achieved even when projects were further classified into more detailed groups (i.e. different sizes of superstores).

A lot of research efforts has also gone into developing cost flow models to help the construction contractor. Zoisner (1974) employed the cost flow approach by selecting four typical housing projects of different sizes and rates of progress. He estimated the cost flow of these projects by preparing the detailed construction schedules and compiling and itemising the relevant construction cost. Cost flow prediction curves were then obtained using polynomial regression. Berdicevsky (1978) selected three university projects of different sizes to examine the feasibility of developing a general cost flow formula for public construction. Peer (1982) utilised the data obtained by Zoisner (1974) and Berdicevsky (1978) to develop his cost flow model. The cumulative field costs for all these projects were brought to a common denominator by relating cost to time in percentage. A prediction curve was then obtained by the polynomial regression of the fourth degree. Peer (1982) concluded that the accuracy achieved in combining all the projects was still within acceptable limits. Kaka and Price (1993) however contended that the number of projects used and the errors caused by estimating and scheduling the activities of the contractors undermines the reliability of the models of Zoisner (1974), Berdicevsky (1978) and Peer (1982). They argued that the difficulty in obtaining actual data could be the only justification for the attempted estimate.

Kaka and Price (1993) also adopted the cost flow approach to cash flow forecasting. They argued that tender unbalancing and overmeasure, together with estimating errors which distort the shape of value curves have no effect on actual cost commitment curve of a project. They also observed that cost commitment curves are not affected considerably by contractual arrangements. These observations informed their choice of cost commitment curve for their models. Using pre determined categories of "project type", "size of contract", "company" and "type of contract," data were collected from 150 completed projects. These were which were categorised into different groups and S-curves were fitted into each using the logit transformation

technique (Kenley and Wilson, 1986). Errors incurred when fitting the curves were measured and compared with those associated in fitting individual projects. Kaka and Price (1993) concluded that the results of testing the reliability of selecting the cost commitment curve to model instead of value curves confirmed the hypothesis that cost commitment models are more accurate and reliable than value models.

Computer-based and artificial intelligence techniques have also been applied in modelling standard cost flow curves. Efforts in this regard include those of Boussabaine and Kaka (1998), Boussabaine *et al.* (1999), and Boussabaine and Elhag (1999). Boussabaine and Kaka (1998) employed a neural network approach to cost flow forecasting in building projects. Boussabaine *et al.* (1999) employed a similar approach to model cost flow for water pipeline projects. Boussabaine and Elhag (1999) applied the fuzzy technique to cash flow analysis and concluded that defuzzified cash flow curves were found to be a good compromise between different cash flow scenarios. However, the developed models did not address the issue of risk impacting cash flow. Hwee and Tiong (2002) developed a computerised model to forecast in-progress cash flow trend. Whilst they claimed that the model has a good predictive ability, the theoretical basis of the model is unclear and the basis on which they selected five risk variable in performing sensitivity analysis is also unclear. Park *et al.* (2005) have also developed a project level cash flow forecasting model based on the planned earned value and actual incurred cost on a jobsite level whilst Khosrowshahi and Kaka (2007) also developed a decision support model for construction cash flow management based on expenditure patterns.

It is apparent from the foregoing that researchers over the last four decades have employed cash flow forecasting methodologies, which utilise the various elements of the cash flow equation. It is also evident that the cost flow approach promises the best predictive accuracy. Whilst this research embraces the cost flow approach, it contends that cost flow models are most useful at the tender stage. This is because the out turn cost flow profile will be different from the one at the pre construction stage due to risk factors eventuating during the in-progress phase of construction. This study therefore seeks to uncover the risk factors impacting the cost flow profile and develop a model for assessing the impacts on the out turn construction cost flow.

Risks in cash flow forecasting

The UK based Association for Project Management (APM) (2006) defines risk as “an uncertain event or set of circumstances that, should it occur, will have an effect on the achievement of one or more project objectives”. This definition takes into account the fact that the effect on project objectives could be either negative or positive. The US based Project Management Institute (PMI) (2008) defines risk as “an uncertain event or condition that, if it occurs will have either a positive or negative effect on one or more of the project’s objectives”, which are usually cost, time, scope and quality.

Some earlier definitions of risk tend to focus on the negative impacts alone. For instance Moavenzadeh and Rossow (1976) regarded risk as an exposure to loss only. Bufaied (1987) described risk in construction as a variable in the construction process whose variation results in uncertainty as to the final cost, duration and quality of the project. Akintoye and MacLeod (1997) submitted that risk in construction has been the object of attention because of time and cost overrun associated with construction projects.

However, this research embraces the definition of risk provided by APM and PMI which considers both the positive and negative impacts of the occurrence

of uncertain events. However, Winch (2010) provided further illumination to risk definition offered by PMI and APM saying that in common practice risk is only used to refer to the probability of a detrimental effect, with the word reward being used to signify the probability of a beneficial event occurring. As such, in order to connect with the real world, the common practice usage of the word risk has been embraced in this study to refer to the probability of a detrimental effect. This defines the premise of this research and it is our view that by minimising risk occurrence and detrimental impacts, the benefit of positive impacts will also be realised.

The major problem that construction managers encounter in making financial decisions involves both the uncertainty and ambiguity surrounding expected cash flows (Eldin, 1989). In the case of complex projects, the problem of uncertainty and ambiguity assumed even greater proportion because of the difficulty in predicting the impact of unexpected changes on construction progress and consequently, on cash flows. The uncertainty and ambiguity are caused not only by project-related problems but also by the economical and technological factors (Laufer and Coheca, 1990). Lowe (1987) maintained that the factors responsible for variation in project cash flow could be grouped under five main headings of contractual, programming, pricing, valuation and economic factors. Kaka and Price (1993) and Kaka (1996) in developing a model for cash flow forecasting identified other risk factors affecting cash flow profiles to include estimating error, tendering strategies, cost variances and duration overrun. Khosrowshahi (2000) also identified other risk factors that impact on cash flow to include delay payment and difficulty in obtaining the right amount of funds at reasonable interest rates.

Using the five generic groupings of risk factors impacting construction cash flow (Lowe, 1987) as a basis, Odeyinka (2003) after a detailed review of other authors came up with a summary of risk factors impacting construction cost flows which include the following: changes to initial design, inclement weather, variation to works (due to architect's instruction (AI)), labour shortage, production target slippage, delay in agreeing variation/dayworks, delay in settling claims, problems with foundations, underestimating project complexity, estimating error, under valuation, delay in payment from client, shortage of key materials, delays in interim certificates, delay in retention release, inflation, compliance with new regulations, subcontractor's insolvency, changes in interest rates, shortage of key plant items, access to funds at reasonable interest rate, archaeological remains, changes in currency exchange rates, civil disturbances, labour strikes, and client's insolvency. While these risk variables have been identified in literature, their impacts on the baseline cost flow forecast at the in-progress phase of construction is the concern of this study. The identified risk variables have been used in developing the questionnaire survey used in this study.

Data and methods

A two-stage approach was adopted in data collection. The first was a UK wide questionnaire survey of risk factors thought to impact cost flow forecast. The second was a case study approach whereby data were collected from case study projects regarding the variability between forecast and actual cost flow at different time periods. Based on the analysis result of the questionnaire survey at the first stage of data collection, significant risk factors were determined. Additional sets of data were collected on the extent of occurrence of the identified significant risk factors in the case study projects.

The questionnaire survey identified 26 risk factors from literature and from discussion with other researchers in construction cash flow as well as from discussion with construction practitioners. These factors were perceived to have potential impact on construction cost flow forecast. The questionnaire was then administered on a project by project basis to 350 randomly selected small, medium and large-scale UK based contractors. The contractors were asked to score on a Likert type scale of 0-5, the extent of occurrence and perceived impacts of the identified risk factors on a recently completed or an ongoing building project. The 6-point 0-5 Likert-type scale also helps to avoid the centrality problem commonly encountered with respondents. In all, 96 responses fit for analysis were received, which represents a 27.4 per cent response rate which is typical of the norm of 20-30 per cent response rate in most postal questionnaire survey of the construction industry (Akintoye and Fitzgerald, 2000). As shown in Tables I-IV, about 88 per cent of the respondents are in senior management

Table I.
Designation
of respondents

Position	Frequency	Percentage	Cumulative per cent
Managing director	10	10.4	10.4
Directors	54	56.3	66.7
senior managers	20	20.8	87.5
Managers	12	12.5	100.0
Total	96	100.0	

Table II.
Academic qualification
of respondents

Qualification	Frequency	Per cent	Cumulative per cent
B.Sc./B. Eng.	45	46.9	46.9
HND	13	13.5	60.4
HNC	12	12.5	72.9
None	26	27.1	100.0
Total	96	100.0	

Table III.
Professional qualification
of respondents

Qualification	Frequency	Per cent	Cumulative per cent
Fellow membership, e.g. FCIQB, FRICS	33	34.4	34.4
Full membership, e.g. MCIOB, MRICS, C.Eng.	53	55.2	89.6
None	10	10.4	100.0
Total	96	100.0	100.0

Table IV.
Construction experience
of respondents

Years	Frequency	Per cent	Cumulative per cent
1-10	7	7.3	7.3
11-20	15	15.6	22.9
21-30	44	45.8	68.7
Over 30	30	31.3	100.0
Total	96	100.0	

Note: Mean = 26.26 years (SD = 9.63)

position, about 73 per cent have further or higher education and about 90 per cent are professionally qualified. The mean experience of respondents is 26.26 years (SD = 9.63). This background information regarding the respondents shows that respondents are competent, experienced and capable of exercising sound judgement in responding to the questionnaire. As such, responses provided by them could be relied upon for this study.

Responses to the questionnaire survey were analysed using mean response analysis and also the “degree of risk” measure. The means were then ranked in order to determine the significant risk factors to focus on. Table V summarises the result of this analysis. Using PMI (2008) risk matrix technique, the first 11 high-ranking risk factors, which also happen to have high-ranking impacts, were selected in order to assess their impacts on variation between the forecast and actual cost flow.

Risk factors	Risk occurrence mean score (P)	Rank	Risk impact mean score (I)	Rank	Degree of risk (P*I)	Rank
Changes to initial design	3.32	1	3.72	1	12.35	1
Inclement weather	3.00	2	3.72	1	11.16	2
Variation to works	2.95	3	3.55	3	10.47	3
Labour shortage	2.81	4	3.53	4	9.92	4
Production target slippage	2.70	5	3.47	5	9.37	5
Delay in agreeing variation/dayworks	2.62	6	3.28	6	8.59	6
Delay in settling claims	2.59	7	3.20	7	8.29	7
Problems with foundations	2.46	8	3.15	8	7.75	8
Underestimating project complexity	2.41	9	3.10	9	7.47	9
Estimating error	2.24	10	3.01	10	6.74	10
Under valuation	2.24	10	3.00	11	6.72	11
Delay in payment from client	2.08	12	2.35	12	4.89	12
Shortage of key materials	2.08	12	2.30	13	4.78	13
Delays in interim certificates	2.03	14	2.27	14	4.61	14
Delay in retention release	1.97	15	2.16	15	4.26	15
Inflation	1.86	16	2.03	16	3.78	16
Compliance with new regulations	1.78	17	1.95	17	3.47	17
Subcontractor's insolvency	1.70	18	1.81	18	3.08	18
Changes in interest rates	1.68	19	1.76	19	2.96	19
Shortage of key plant items	1.68	19	1.62	20	2.72	20
Access to funds at reasonable interest rate	1.46	21	1.62	20	2.37	21
Archaeological remains	1.46	21	1.54	22	2.25	22
Changes in currency exchange rates	1.35	23	1.41	23	1.90	23
Civil disturbances	1.24	24	1.35	24	1.67	24
Labour strikes	1.19	25	1.32	25	1.57	25
Client's insolvency	0.11	26	0.49	26	0.05	26

Table V.
Perception of risk
occurrence and impacts
on cost flow forecast

The second stage of the data collection involves collecting data from 55 case study projects. The case study projects were obtained from the archives of a construction company that agreed to participate in the research. They are all new build commercial buildings, educational buildings and health and recreational projects, all of comparable project values. Archives data of monthly forecast and actual cost flows were obtained from completed or ongoing case study projects. From these two pair of archives data, the percentage variation was computed and those at 30, 50, 70 and 100 per cent completion stages were selected as dependent variables for use in modelling. The rationale for choosing these time intervals is that previous research (Kaka, 1999) has shown that the majority of variability in cost flow curves occurred between the 30 and 70 per cent portions of the curves. The 100 per cent point was also selected in order to track down possible cost overruns at the end of the contract period. In addition, the project Quantity Surveyors who worked on the case study projects were required to score the extent of occurrence of the identified 11 significant risk variables in their projects' cost flow. Each of the data set was also used as independent variables for modelling purposes. In all, four pairs of data sets were generated from 55 case study projects and were used in model development.

Developing the regression models

In order to develop multiple linear regression (MLR) models to assess risk impacts on construction cost flow forecast, the identified 11 significant risk factors, were used as the independent variables. The estimated percentage variation at 30, 50, 70 and 100 per cent completion stages were used as the dependent variables. Out of the 55 data sets collected, 40 data sets were used to develop the regression models. The Statistical Package for Social Sciences (SPSS) was employed, using the simultaneous multiple regression procedure. Table VI shows the coefficients of the four models developed which are expressed as equations (1)-(4).

Variation at 30 per cent:

$$Y_{30} = 0.07217 + 0.00001X_1 - 0.00794X_2 - 0.01137X_3 + 0.00486X_4 - 0.02368X_5 + 0.01681X_6 + 0.01044X_7 + 0.00438X_8 + 0.03526X_9 - 0.00227X_{10} + 0.00479X_{11} \quad (1)$$

Significant risk variables	Variation at 30%	Variation at 50%	Variation at 70%	Variation at 100%
(Constant)	0.40551	0.74795	0.15041	0.25657
Changes to initial design	0.00375	-0.17339	-0.10450	-0.07801
Inclement weather	0.06126	0.22315	0.10387	0.17641
Variation to works (AI)	-0.02272	-0.00065	0.03629	-0.03725
Labour shortage	-0.09340	0.16787	0.18827	-0.03728
Production target slippage	-0.11753	-0.17312	-0.17004	-0.10525
Delay in agreeing variation/dayworks	0.18141	0.04925	0.06582	-0.02688
Delay in settling claims	-0.09019	-0.10281	-0.02477	0.06710
Problems with foundations	-0.05765	0.04565	0.10675	0.08062
Underestimating project complexity	0.07520	0.12547	0.04823	0.08784
Estimating error	-0.00894	-0.32675	-0.21538	-0.14384
Under valuation	-0.13734	0.02198	-0.00474	-0.03885

Table VI.
MLR coefficients of variation at 30, 50, 70 and 100 per cent completion stages

Variation at 50 per cent:

$$Y_{50} = 0.07213 + 0.02335X_1 + 0.02730X_2 - 0.02441X_3 - 0.00304X_4 - 0.01450X_5 + 0.00929X_6 + 0.01977X_7 - 0.02724X_8 + 0.01410X_9 + 0.05361X_{10} - 0.02688X_{11} \quad (2)$$

Variation at 70 per cent:

$$Y_{70} = 0.00282 + 0.01408X_1 + 0.02646X_2 - 0.00569X_3 + 0.00006X_4 + 0.01434X_5 + 0.01281X_6 + 0.03998X_7 - 0.06530X_8 - 0.00841X_9 + 0.06900X_{10} - 0.04513X_{11} \quad (3)$$

Variation at 100 per cent:

$$Y_{100} = -0.08218 + 0.01790X_1 + 0.01316X_2 + 0.00937X_3 + 0.01203X_4 + 0.02849X_5 - 0.00588X_6 + 0.01152X_7 - 0.02612X_8 - 0.04557X_9 + 0.06334X_{10} - 0.01660X_{11} \quad (4)$$

The coefficient of multiple correlation R, which shows the correlation between the predicted and actual values of the dependent variables, gives fairly good result as shown in Table VII. According to Nurosis (2009), one touchstone of a good model is its predictive power. The R^2 and adjusted R^2 of multiple regression models are means of assessing their predictive power. They pre-measure the proportion of variance explained or error reduced by the model (Nurosis, 2009). The R^2 shown in Table VII are mostly above average, However, the proportion of variance they explained ranging from 42.7 to 74.5 per cent are not satisfactory but promising. The adjusted R^2 -values are mostly below average with the exception of the adjusted R^2 for variation at 50 per cent completion.

In order to test the predictive power of the models, test data from 15 projects were used. The scores for the extent of risk occurrence (independent variables) were used to predict variations at 30, 50, 70 and 100 per cent completion periods using the regression coefficients in Table VI and equations (1)-(4). The percentage error was used to measure the accuracy of the models. Table VIII summarises the results of the models' prediction and the absolute mean percentage error for each of the models. These range from 35.1 to 90.2 per cent. These percentage errors are obviously not satisfactory. However, the results obtained indicate that the possibility exists of developing models to assess risk impacts on construction cost flow forecast. It is however obvious that the use of multi-linear regression did not produce satisfactory results. Again, further attempts

Measure	Variation at 30%	Variation at 50%	Variation at 70%	Variation at 100%
R^2	0.427	0.745	0.510	0.584
Adjusted R^2	0.202	0.644	0.317	0.420
Coefficient of multiple correlation R	0.654	0.863	0.714	0.764

Table VII.
Accuracy measurement of regression models of variation at 30, 50, 70 and 100 per cent completion

Table VIII.
Prediction performance
of the regression models

Project no.	30% completion			50% completion			70% completion			100% completion			Relative absolute deviation at 100% MAD
	Actual variation at 30%	Predicted variation at 30%	Relative absolute deviation at 30%	Actual variation at 50%	Predicted variation at 50%	Relative absolute deviation at 50%	Actual variation at 70%	Predicted variation at 70%	Relative absolute deviation at 70%	Actual variation at 100%	Predicted variation at 100%	Relative absolute deviation at 100%	
1	-0.35	0.006	1.018	-0.3	-0.052	0.827	-0.4	-0.013	0.968	-0.25	-0.076	0.696	0.877
2	-0.57	-0.119	0.791	-0.51	-0.185	0.638	-0.24	-0.119	0.504	-0.13	0.021	1.163	0.774
3	-0.43	-0.097	0.773	-0.5	-0.306	0.388	-0.44	-0.282	0.360	-0.27	-0.157	0.418	0.485
4	0.3	0.340	0.135	0.12	0.152	0.266	0.2	0.227	0.134	0.2	0.160	0.199	0.183
5	0.45	0.423	0.060	0.5	0.450	0.099	-0.12	0.116	1.965	0.16	0.178	0.113	0.559
6	0.38	0.125	0.670	0.19	0.072	0.620	0.07	-0.067	1.958	-0.05	-0.057	0.131	0.845
7	0.47	0.348	0.259	0.13	0.116	0.105	0.09	0.066	0.261	0.19	0.103	0.457	0.271
8	0.49	-0.154	1.315	-0.3	-0.379	0.263	-0.25	-0.231	0.075	-0.05	-0.070	0.402	0.514
9	0.08	0.077	0.038	0.31	0.141	0.546	0.57	0.213	0.626	0.4	0.218	0.454	0.416
10	0.2	0.199	0.004	0.26	0.349	0.341	0.19	0.240	0.265	0.19	0.243	0.280	0.223
11	-0.21	-0.208	0.009	-0.17	-0.164	0.038	-0.06	-0.179	1.985	-0.04	-0.175	3.385	1.354
12	0.18	0.311	0.725	0.19	0.178	0.063	-0.1	0.004	1.041	-0.21	-0.100	0.525	0.589
13	0.19	0.180	0.051	0.43	0.431	0.001	0.17	0.112	0.341	0.07	0.086	0.228	0.155
14	0.18	0.041	0.773	-0.26	-0.500	0.921	-0.3	-0.304	0.012	0.1	-0.080	1.800	0.877
15	0.36	0.439	0.220	0.24	0.275	0.145	0.17	-0.027	1.159	0.02	-0.046	3.281	1.201
Absolute mean error of the estimate			0.456			0.351			0.777			0.902	

using stepwise multiple regressions as suggested by Kinnear and Gray (2011) only led to deletion of very significant variables and still did not produce satisfactory results. This led to the conclusion that the relationships between the independent and dependent variables are strictly non linear.

Despite the unsatisfactory predictive powers of the regression models, their major achievements seem to be the provision of indication of which variables to pay attention to at different completion periods. For instance, looking at the coefficients of the developed models (Table VI), it is evident that “changes to the initial design” risk variable generated a positive variation at 30 per cent completion while the variations at other completion periods are negatively signed. This suggests that at the early stage of construction (30 per cent completion), changes to initial design made by clients would most likely shift the cost flow profile up and thereby bring about an increase to the initial cost flow forecast. In the case of the “inclement weather” risk variable, positive coefficients of variation were generated at all stages of construction. This is not unexpected, as inclement weather is known to result in loss of productivity at any stage of construction and consequently increase cost.

Moreover, from Table VI, it is evident that variation to works occasioned by AIs’ resulted in negative variations apart from the one at 70 per cent completion which is positively signed. This is not a surprise because at 70 per cent completion, most of the finishes works would be taking place and in most cases, Architects do change their minds about previously specified finishes. Many times, these changes lead to increase in cost. Furthermore, it is obvious from Table VI that “labour shortage” risk variable generates positive variations at 50 and 70 per cent completion stages, while the variations are negatively signed at other stages. Again, the positive variations are not surprising because the labour resource requirement is at its peak between 50 and 70 per cent completion duration. Contractors may at such peak demand periods have to pay more to attract the right calibre of labour force, hence a positive variation between the forecast and actual cost flow. The converse may also be true of negative variations at 30 and 100 per cent completion periods.

Furthermore, it is obvious from Table VI that “delay in agreeing variation/dayworks” generated positive variations at 30, 50 and 70 per cent completion periods. This is not a surprise because such a delay will result in capital lock up which may also increase the cost of borrowing to improve the cash flow. This invariably contributes to a positive shift in the cash flow profile. The “delay in settling claims” risk variable generates negative coefficients of variation throughout the project life apart from at 100 per cent completion. This however is a surprise as delay in settling claims naturally results in capital lock up which also increases the cost of borrowing. However, the fact that positive coefficient of variation was generated at 100 per cent completion may be explained by the fact that the industry is accustomed to getting claims settled towards the end of the contract period. Again it is surprising that the “problems with foundations” risk variable generated negative coefficient of variation at 30 per cent completion period during which time foundation work is expected to be carried out. However, it generated a positive coefficient of variation at 50, 70 and 100 per cent completion period during which time foundation works would have long been completed. While it is difficult to provide a rational explanation for this observation, it may be one of the results of the non-conformity to the linear assumption of the regression models, which resulted in unacceptable prediction errors.

Moreover, from Table VI, it is evident that “underestimating project complexity” generated positive coefficient of variation throughout the project life. This is not surprising because if project complexity is underestimated, cost flow forecast would also tend to be underestimated, making the actual cost flow more than the forecast. On the other hand, “estimating error” generated a negative coefficient of variation throughout the project life. While this is not a surprise, it suggests that estimators initially over-estimated the forecast, thus contributing to making the actual cost flow lower than the forecast. This practice of overestimating by estimators is not uncommon in the construction industry.

Finally, the “under valuation” risk variable generated negative coefficient of variation at 30, 70 and 100 per cent completion periods while it generated positive variation at 50 per cent completion. This is not a surprise because in most cases, contractors would try to guide against under valuation in order to improve their cash flow. The fact that it generated positive coefficient of variation at 50 per cent completion may be due to delay in agreeing variation or settling claims.

Fitting the MLR model curves

In order to have the true graphical picture, it is useful to juxtapose the curves of the actual and forecast cost flows with that predicted by the model due to risk impacts. Kenley and Wilson (1986) and Kaka and Price (1993) have successfully employed the technique of logit transformation to fit cash flow curves. This technique was also employed in this study. The logit transformation according to Kenley and Wilson (1986) is the simplest of the sigmoid transformations and allows the S-curve to be represented in linear form. Kaka (1999) summarised the steps to the logit transformation technique as follows.

The linear equation is found by a logit transformation (equation (5)) of both the independent and dependent variables (in this case, time and cost, respectively):

$$\text{Logit} = \ln \frac{z}{1-z} \quad (5)$$

Where z is the variable to be transformed and Logit is the transformation.

Once the two variables are transformed, a linear equation can be fitted into the transformed data. This linear equation is expressed in equation (6) as follows:

$$Y = \alpha + \beta X, \quad (6)$$

where:

$$Y = \ln \frac{c}{1-c} \quad X = \ln \frac{t}{1-t}$$

In order to transform data for a particular project, X and Y must be calculated for each value of t (time) and c (cost), respectively. Deriving the constants α and β is thus a simple linear regression of the transformed data.

Once α and β are derived, the fitted S-curve can be generated using equation (7):

$$c = \frac{F}{1+F} \quad (7)$$

where:

$$F = e^{\alpha} \left(\frac{t}{1-t} \right)^{\beta}$$

Equation (7) then forms the equation of the sigmoid curve, which describes the flow of cost on a specific building project.

Using the above given transformation, the forecast and actual cost flow curves were fitted for the data set used for testing the model. Moreover, using the logit transformation, the predicted values of variation at 30, 50, and 70 per cent due to risk impacts were used to fit the predicted regression model curves. Examples of these curves are shown in Figure 2.

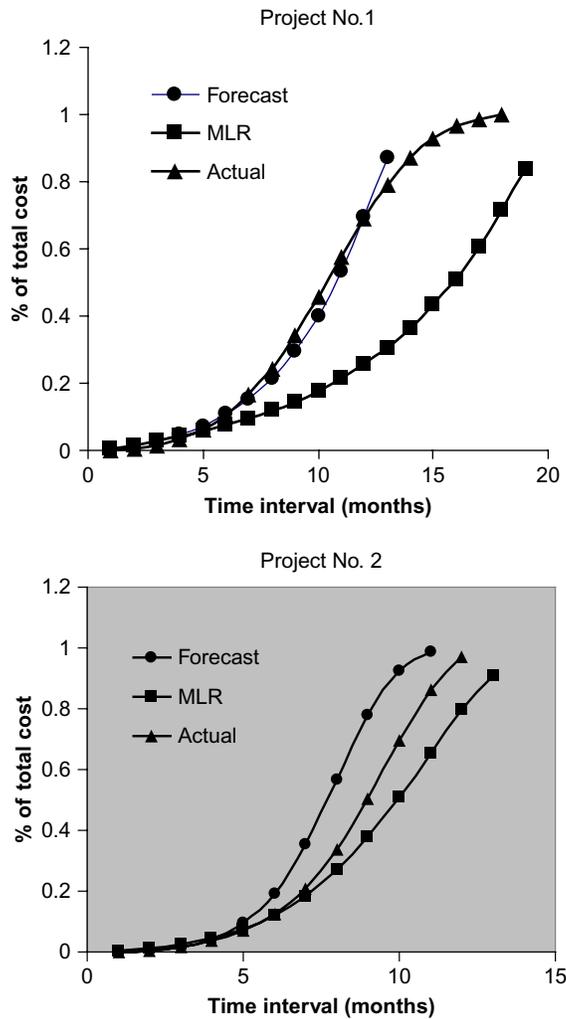


Figure 2. Examples of fitted forecast, MLR prediction and actual curves

Conclusion

This paper offers two main contributions that are important to cost flow forecasting. The first conclusion was the identification of the significant risk factors impacting construction cost flow forecast. The second contribution was the development of a regression model for assessing the impacts of the identified significant risk factors on base line cost flow forecast. Finally, we discuss the limitation of the developed model and the potential for further research.

Starting with the first conclusion, out of the 26 risk factors identified as potentially impacting construction cost flow forecast, the study concluded that 11 of those risk factors are most significant. Using the traffic light signs, these will be classified as “red” risks, indicating that they cannot be ignored. These significant risk factors are: changes to initial design, inclement weather, variation to works, labour shortage, production target slippage, delay in agreeing variation/day works, delay in settling claims, problems with foundations, underestimating, project complexity, estimating error and under valuation. An introspective examination of these factors shows that some of them relate to payment delays, varying the works and natural inhibition. It is interesting to note that these three generic factors also featured prominently in related studies (Ireland, 1983; Kenley and Wilson, 1986; Kaka, 1996; Hwee and Tiong, 2002) as key variables that must be considered in any cash flow or time performance studies.

The second conclusion is based on the development of predictive regression models for assessing the impacts of the identified significant risk factors on construction cost flow base line forecast. The models, developed for the 30, 50, 70 and 100 per cent completion periods show the possibility of modelling the impacts of multiplicity of risk variables on cost flow forecast. Although not with high predictive ability, the developed models showed that some relationships exist between the independent and dependent variables, which may be strictly non-linear. They also demonstrated the possibility of modelling the phenomenon and offer means of possible explanation for the observed variation between the forecast and actual cost flow at different stages of construction. In addition, they provided some insights into the impacts of some identified significant risk factors at different stages of construction. Such information are invaluable to the construction contractor.

Lastly, it is noteworthy that due to the small sample size of case study projects available, it was not possible to segregate the data set into different procurement methods or project types. Kaka and Dawood (2000) found that procurement method has impacts on cost flow profile. It is therefore suggested that future research should attempt to develop models along procurement routes. It is also suggested that future research should consider modelling for different contract values.

References

- Akintoye, A. and Fitzgerald, E. (2000), “A survey of current cost estimating practices in the UK”, *Construction Management and Economics*, Vol. 18 No. 2, pp. 161-72.
- Akintoye, A.S. and MacLeod, M.J. (1997), “Risk analysis and management in construction”, *International Journal of Project Management*, Vol. 15 No. 1, pp. 31-8.
- Association for Project Management (2006), *APM Body of Knowledge*, 5th ed., Association for Project Management, High Wycombe.

- Berdicevsky, S. (1978), "Erection cost flow analysis in public projects", MSc thesis, Technion-Israel Institute of Technology, Haifa.
- Berny, J. and Howe, R. (1983), "Project management control using real time budgeting and forecasting models", *Construction Papers*, Vol. 2, pp. 19-40.
- Boussabaine, A.H. and Elhag, T. (1999), "Applying fuzzy techniques to cash flow analysis", *Construction Management and Economics*, Vol. 17, pp. 745-55.
- Boussabaine, A.H. and Kaka, A.P. (1998), "A neural networks approach for cost flow forecasting", *Construction Management and Economics*, Vol. 16, pp. 471-9.
- Boussabaine, A.H., Thomas, R. and Elhag, T.M.S. (1999), "Modelling cost flow forecasting for water pipeline projects using neural networks", *Engineering, Construction and Architectural Management*, Vol. 6 No. 3, pp. 213-24.
- Bromilow, F.J. and Henderson, J.A. (1977), *Procedures for Reckoning the Performance of Building Contracts*, 2nd ed., CSIRO, Division of Building Research, Highett, Special Report.
- Bufaied, A.S. (1987), "Risks in the construction industry: their causes and their effects at the project level", PhD thesis, University of Manchester, UMIST, Manchester.
- Drake, B.E. (1978), "A mathematical model for expenditure forecasting post contract", *Proceedings of the Second International Symposium on Organisation and Management of Construction, Technion Israel Institute of Technology, Haifa*, pp. 163-83.
- Eldin, N. (1989), "Cost control systems for PMT use", *Transactions of the AACE*, pp. F3.1-F3.5.
- Evans, R.C. and Kaka, A.P. (1998), "Analysis of the accuracy of standard/average value curves using food retail building projects as case studies", *Engineering, Construction and Architectural Management*, Vol. 5 No. 1, pp. 58-67.
- Hardy, J.V. (1970), "Cash flow forecasting for the construction industry", MSc Report, Dept. of Civil Engineering, Loughborough University of Technology, Loughborough.
- Hudson, K.W. (1978), "DHSS expenditure forecasting method", *Chartered Surveyor – Building and Quantity Surveying Quarterly*, Vol. 5, pp. 42-5.
- Hwee, N.G. and Tiong, R.L. (2002), "Model on cash flow forecasting and risk analysis for contracting firms", *International Journal of Project Management*, Vol. 20, pp. 351-63.
- Ireland, V. (1983), "The role of managerial actions in the cost, time and quality performance of high rise commercial building projects", PhD thesis, University of Sydney.
- Kaka, A.P. (1990), "Corporate financial model for construction contractors", PhD thesis, Department of Civil Engineering, Loughborough University of Technology, Loughborough.
- Kaka, A.P. (1996), "Towards more flexible and accurate cash flow forecasting", *Construction Management and Economics*, Vol. 14 No. 1, pp. 35-44.
- Kaka, A.P. (1999), "The development of a benchmark model that uses historical data for monitoring the progress of current construction projects", *Engineering, Construction and Architectural Management*, Vol. 6 No. 3, pp. 256-66.
- Kaka, A.P. and Price, A.D.F. (1991), "Net cash flow models: are they reliable?", *Construction Management and Economics*, Vol. 9, pp. 291-308.
- Kaka, A.P. and Price, A.D.F. (1993), "Modelling standard cost commitment curves for contractors' cash flow forecasting", *Construction Management and Economics*, Vol. 11, pp. 271-83.

- Kenley, R. (2003), *Financing Construction: Cash Flows and Cash Farming*, Spon Press, London, pp. 3-8.
- Kenley, R. and Wilson, O.D. (1986), "A construction project cash flow model – an idiographic approach", *Construction Management and Economics*, Vol. 4, pp. 213-32.
- Kenley, R. and Wilson, O.D. (1989), "A construction project net cash flow model", *Construction Management and Economics*, Vol. 7, pp. 3-18.
- Khosrowshahi, F. (1991), "Simulation of expenditure patterns of construction projects", *Construction Management and Economics*, Vol. 9, pp. 113-32.
- Khosrowshahi, F. (2000), "A radical approach to risk in project financial management", *Proceedings of the 16th Annual ARCOM Conference, Glasgow Caledonian University, 6-8 September*, pp. 547-56.
- Khosrowshahi, F. and Kaka, A. (2007), "A decision support model for construction cash flow management", *Computer-Aided Civil and Infrastructure Engineering*, Vol. 22, pp. 527-39.
- Kinney, P.R. and Gray, C.D. (2011), *IBM SPSS 18 Statistics Made Simple*, Psychology Press, New York, NY.
- Laufer, A. and Coheca, D. (1990), "Factors affecting construction planning outcomes", *Journal of Construction Engineering and Management*, Vol. 116 No. 6, pp. 135-56.
- Lowe, J.G. (1987), "Cash flow and the construction client – a theoretical approach", in Lansley, P.R. and Harlow, P.A. (Eds), *Managing Construction Worldwide*, Vol. 1, E & FN Spon, London, pp. 327-36.
- Moavenzadeh, F. and Rossow, J. (1976), "Risks and risk analysis in construction management", *Proceeding of the CIB W65, Symposium on Organisation and Management of Construction, 19-20 May*, US National Academy of Science, Washington, DC.
- Nurosis, M.J. (2009), *PASW Statistics 18 Statistical Procedures Companion*, Prentice-Hall, Upper Saddle River, NJ.
- Odeyinka, H.A. (2003), "The development and validation of models for assessing risk impacts on cash flow forecast", PhD thesis, School of the Built and Natural Environment, Glasgow, Caledonian University, Glasgow.
- Park, H.K., Han, S.H. and Russell, J. (2005), "Cash flow forecasting model for general contractors using moving weights of cost categories", *Journal of Management in Engineering*, Vol. 21 No. 4, pp. 164-75.
- Peer, S. (1982), "Application of cost flow forecasting models", *Journal of the Construction Division ASCE*, Vol. 108, CO2, pp. 226-32.
- Project Management Institute (2008), *A Guide to the Project Management Body of Knowledge*, 4th ed., Project Management Institute, Atlanta, GA.
- Sidwell, A.C. and Rumball, M.A. (1982), "The prediction of expenditure profiles for building projects", in Brandon, P.S. (Ed.), *Building Cost Techniques: New Directions*, E. & F.N Spon, London, pp. 324-38.
- Winch, G.M. (2010), *Managing Construction Projects*, 2nd ed., Wiley-Blackwell, Chichester.
- Zoisner, J. (1974), *Erection Cost Flow Analysis in Housing Projects as a Function of its Size and Construction Time*, MSc thesis, Technion-Israel Institute of Technology, Haifa.

About the authors

Dr Henry Odeyinka is the Course Director of the MSc Commercial Management in Construction in the School of the Built Environment, University of Ulster, Jordanstown, Belfast. He is also a core member of the Built Environment Research Institute. He is a reviewer for many construction

journals and also a member of the scientific committee of COBRA conference. His research interests are in the areas of risk management in construction, financial management in construction and construction procurement. Henry Odeyinka is the corresponding author and can be contacted at: h.odeyinka@ulster.ac.uk

Dr John Lowe is the Programme Leader for the MSc Construction Economics course in the School of Engineering and Built Environment, Glasgow Caledonian University. He is also a reviewer for many construction journals. His research interests are in the areas of macro economics and the construction industry, cash flow management and construction economics.

Professor Ammar Kaka was formerly a Professor of Construction Management and Economics in the School of the Built Environment, Heriot Watt University, Edinburgh. He is currently the Deputy Head of Campus at Heriot Watt University, Dubai campus. His research interests are in the areas of contractor's cash flow modelling, financial management and construction productivity.

Regression
modelling of
risk impacts

221