Lowe, D J, Emsley, M W and Harding, A (2007) Relationships between total construction cost and design related variables. *Journal of Financial Management of Property and Construction*, **12**(1), 11-23, eScholarID:1g171 | DOI:10.1108/13664380780001090

Relationships between total construction cost and design related variables

David J Lowe, Senior Lecturer, Centre for Research in the Management of Projects, Manchester Business School, University of Manchester, UK

Margaret W Emsley, Lecturer, Management of Projects Research Group, School of Mechanical, Aerospace and Civil Engineering, University of Manchester, UK

Anthony Harding, Halcrow Group Ltd, Hammersmith, London

Correspondence Address:

Dr David J Lowe
Centre for Research in the Management of Projects
Manchester Business School
University of Manchester
Manchester
M15 6PB

Tel: 0161 306 4643
Fax: 0161 306 4646
E-mail: david.lowe@manchester.ac.uk

Lowe, D J, Emsley, M W and Harding, A (2007) Relationships between total construction cost and design related variables. *Journal of Financial Management of Property and Construction*, **12**(1), 11-22. eScholarID:1a171 | DOI:10.1108/13664380780001090

Relationships between total construction cost and design related variables

Abstract

This, the second of two papers, seeks to redress the omission in recent literature on the influence of project strategic, site related and design related variables on the cost of construction. It presents, in part, the results of an investigation into the influence of 41 independent variables on both construction cost and client cost, concentrating on design related variables. Data were collected from 286 construction projects in the United Kingdom and correlation and test for differences were used to determine the relationships that exist between the dependent and independent variables. The analysis ascertains the cost ranking of many design related variables and establishes other relationships which exist within the data, confirming many of the relationships that had been anticipated from the literature. It also established the ordinal sequence of several nominal variables. These data, therefore, can be confidently used to develop models of the total cost of construction as verified by the development of both regression analysis and neural network cost models.

Keywords:

Construction cost, client costs, design related variables

INTRODUCTION

Standard texts on building economics and cost planning (for example, Ashworth, 2004; Ferry *et al*., 1999; Jagger *et al*., 2002; Morton and Jagger, 1995; and Seeley, 1995) discuss the

Lowe, D J, Emsley, M W and Harding, A (2007) Relationships between total construction cost and design related variables. *Journal of Financial Management of Property and Construction*, **12**(1), 11-23. eScholarID:1a171 | DOI:10.1108/13664380780001090

implications of various design cost parameters on construction costs. However, these texts do not provide a measure of the relationship between the various design variables and construction cost or their relative importance, except in a superficial way. For example, Morton and Jagger (1995) state that function and size (floor area) are the ‘two most obvious’ determinants of construction costs.

Equally, empirical studies into the relationship between building characteristics and construction cost generally tend to be dated. Although more recently, Picken and Ilozor (2003) investigated the relationship between height and construction costs of buildings in Hong Kong; Swaffield and Pasquire (1999) analysed the relationship between building form and function, and the cost of mechanical and electrical services; Wing (1999) considered the issue of plan shape complexity and Tan (1999) assessed the influence of height on construction costs.

A previous paper (Lowe *et al*., 2006) examines the relationships associated with the two dependent output variables (construction cost and client cost) and the project strategic, site related and building definition independent input variables. It gave the background to the research, described previous research in this field and discussed data representation. This current paper builds on this work; it examines the relationships between the independent input design related variables (excluding building definition) and, where appropriate, the cost related dependent output variables. It also, summarizes the application of both regression analysis and neural network modelling to these data to produce cost models. The design related variables (see Table 1) are classified into four groups: building definition; structure; finishes and openings; and services installations.

METHODOLOGY

Ultimately, 286 sets of data were collected, (for a detailed account of the data collection issues see Harding et al. 2000), which were divided into dependent input variables and independent output variables (Emsley *et al.,* 2002). Two dependent variables were identified – construction cost (final account) (COST) and client cost (CCOST) (incorporating professional fees and the internal costs incurred by the client) - which may be combined to give the total cost to the client. An extensive literature review identified many output variables and these were finally reduced to 41 variables which it was believed would be known at the early estimating stage (for which the models to be developed were intended). The variables were categorised as shown in Table 1. The variables of time and geographical location were accommodated through the use of the BCIS indices to bring all project to a common location and base date and the cost of activities such as external works and demolition were removed from the final account figure as it was thought that these activities would be extremely difficult to model.

<<<< Insert Table 1 about here >>>>

Data representation

It is necessary to consider the best way of representing each variable in any subsequent modelling, such as regression analyses. Some variables are real numbers, such as number of lifts, and many of these are on a continuous scale, such as cost and GIFA. However, the remaining variables are categorical variables that may be represented in one of three ways. Firstly, there are those variables which are categorised so that there is an obvious relationship between the category and the influence of that variable on the output of the model. For example, site access is categorised as ‘unrestricted’, ‘restricted’ and ‘highly restricted’ and it is reasonable to assume that an increase in restriction will cause an increase in cost. Secondly, there are those variables where no such order is immediately apparent but where the variable may be expressed in terms of cost. For example, for internal wall finishes, the value of input was set to be the standard cost per m^2 of each finish and this provides an order proportional to how much each finish is expected to influence cost. Finally, there are those variables where it is not possible to ascertain a standard cost and where there is a lack of consensus on comparative costs, so a consistent ordering which would apply in all circumstances cannot be obviously obtained. An example is frame type, which is categorised as ‘*in situ*’, ‘masonry’, ‘precast’, ‘steel’ or ‘timber’. This requires the use of binary input coding (yes/no) which is applied to each possible choice, thus treating each categorical variable as a series of binary variables, with the disadvantage of increasing the number of variables in the model and subsequently increasing the requirements for data. However, by analysing the data obtained for these variables (as described in this paper), it is possible that a consistent order may be obtained which will considerably facilitate any subsequent data modelling. Again, a more exhaustive description of the data representation issues can be found in Harding et al. (2000).

Analysis

Data analyses were undertaken using the Statistical Package for the Social Sciences (SPSS for Windows, release 12.0.0). Log transformations were performed for all continuous scale variables to create or approximate a normal distribution in order to satisfy parametric test assumptions. Descriptive statistics and tests for normality were calculated for each independent variable. The continuous scale independent variables were correlated, using Pearson's product moment and Spearman's rank correlation, with the dependent variables. The categorical variables were analysed for differences between subgroups by means of one-way analysis of variance (ANOVA) and its equivalent non-parametric test, the Kruskal-Wallis H test, or by means of the t-test and the Mann-Whitney U/Wilcoxon Rank Sum W test. Relationships between the independent variables were investigated using Spearman's rank correlation. This analysis was exploratory in nature and is, therefore, interpreted by reference to significance levels (two-tailed).

DATA EXPORATION

Independent structure variables

Descriptive statistics for structure related categorical variables are shown in Table 2 and Table 3 gives details of the statistical tests which have been carried out between sub-groups of structure variables

External walls

The variable external walls has a mean cost of £51.86 and median of £39. The distribution is very highly skewed (8.427) with a kurtosis value of 84.775. This is partly because of two outliers. However, it is entirely possible that these projects are representative of the fact that there are a very small proportion of building projects which have very expensive external wall solutions. External walls has no strong correlation with any other variable (see Table 5), suggesting that external walls are comparatively unrelated to other finish type variables, but still influence cost, correlating strongly with construction cost (0.531) and log of client cost (0.517).

<<<< Insert Tables 2 & 3 about here >>>>

Frame

An analysis of variances showed that there are significant differences with cost, GIFA, cost/m^2, client cost, (cost/m^2)/function and function between the frame sub-groups. The significant differences in client cost can be explained by the significant differences in GIFA, as the differences in client cost/m^2 are not significant. Furthermore, the apparent significant differences in the construction costs are significant when the influence of building function is considered ((cost/m^2)/function variable). If the difference in means is used to obtain an ordinal variable for the prediction of the cost/m^2, then the order should be as shown in Table 2.

Piling

The significance of cost differences is assessed by a T test. It might be possible to explain the apparent difference in cost with and without piling by considering the influence of GIFA and function. The (cost/m^2)/function variable takes both of these variables into account, and the results suggests that the cost difference between piled and non-piled buildings is significant.

Roof construction

An analysis of variances was performed to determine whether there might be any significant cost differences between the sub-groups and there are significant differences with cost, GIFA, client cost, function and cost/m^2. Therefore, the significant differences in client cost can be explained by the significant differences in GIFA, as the differences in client cost/m^2 is not significant. Furthermore, the apparent significant differences in the construction cost are not significant when the influence of function is considered. If the difference in means is used to obtain an ordinal variable for the prediction of the cost/m^2, then the order is as shown in Table 2.

Roof profile

An analysis of variances showed that there are significant differences with client cost, cost, GIFA and cost/m^2 between the roof profile sub-groups. The significant differences in client cost can be explained by the significant differences in GIFA, as the differences in client cost/m^2 is not significant. Furthermore, the apparent significant differences in the construction costs are not significant when the influence of function is considered ((cost/m^2)/function variable). However, closer analysis of these results reveals that the average cost/m^2 of a flat roof building is over £100 dearer than a pitched roof building and £34.36 more expensive than a curved roof building. The significance of the difference between these groups was compared using Bonferroni Multiple Comparisons test.

For construction cost/m^2, pitched is significantly lower than flat, however, there are no other significant differences. Likewise for the cost function ((cost/m^2)/function) variable there are no significant differences between the subgroups. However, the average cost values obtained for each of these groups, expressed either as cost/m^2 or as (cost/m^2)/function suggest that the correct order for roof profile is pitched, curved then flat, with pitched being cheapest and flat the most expensive. Traditionally, flat roofs have been considered as the cheapest, followed by pitched and then curved. However, this does not appear to be the case; supporting the assertion of Ferry et al. (1999) that "… for medium to large spans a satisfactory pitched roof is likely to be cheaper than a flat roof of comparable quality, partly because of simplicity of spanning large areas with roof trusses than deep beams". This suggests that other cost significant variables which vary across the groups may also be at play. Perhaps the best suggestion is that the three possible orders for the variable tested for inclusion within a stepwise regression. The order that gives the highest value of T for inclusion would then be the most appropriate order.

Stair types

The existence of significant cost differences between these types is addressed using analysis of variances which shows that there are significant differences with cost, GIFA, function, cost/m^2 and client cost between the stair types sub-groups. The significant differences in client cost can be explained by the significant differences in GIFA, as the differences in client cost/m^2 is not significant. Furthermore, the apparent significant differences in the construction cost are not significant when the influence of function is considered.

The fact that there are no significant differences when function is taken into account could arise from there being no significant differences between the cost of different stair types. Ashworth (2004) considers stairs to "… represent a minor cost element in a building… Because they are largely of a functional nature their structure cost are comparable with each other and the differences in the elemental analysis are therefore determined by the finishes which might be applied". However, it could also arise from the fact that different stair types are associated with different building types. If the significant differences between the groups are assumed and compared using the cost/m^2, then the order is shown in Table 2. These results suggest that to have no stairs is more expensive than to have timber stairs, which seems unlikely. However, the values are close together, so the validity of this particular ordering was assessed using Bonferroni Multiple Comparisons test.

For construction cost/m^2, insitu concrete is significantly higher than none and timber, while precast is significantly higher than timber; however, there are no other significant differences.

Likewise for the cost function ((cost/m^2)/function) variable there are no significant differences between the subgroups. Therefore it is not possible to be confident about the ordering implied by this analysis, and the original order for the variable is retained.

Substructure

Differences in cost among the different groups is assessed using analysis of variances which showed that there are significant differences with cost/m^2, (cost/m^2)/function, GIFA and cost between the substructure sub-groups. Likewise, the differences in the construction costs are significant when the influence of building function is considered and an attempt is made to find an order among the groups. For this, the cost/m^2 and (cost/m^2)/function were ordered by value and the significance of the difference between these groups was compared by using Bonferroni Multiple Comparisons test. For both construction cost/m^2 and cost/m^2/function, raft is significantly higher than strip, pad/strip and pad; however, there are no other significant differences between the subgroups. As pad → pad/strip → strip is a logical progression, this order will be used. It is also possible that the increased costs of ground beams are partly due to the fact that these solutions also include piling. Piling has been shown to yield an increase in cost of £106.02/m^2, and as 80% more ground beam solutions are piled than any other solution this suggests that the true cost of ground beams may be significantly less. This would reduce the true cost of ground beams to be more in line with pad and pad/strip foundation types. Therefore the order should be: pad, pad/strip, ground beams, strip, raft.

Upper floors

Differences in cost between these groups are analysed using an analysis of variance which shows that there are significant differences with cost, GIFA, function, cost/m^2 and client cost between the upper floor sub-groups. The significant differences in client cost can be explained by the significant differences in GIFA, as the differences in client cost/m^2 is not significant. Furthermore, the apparent significant differences in the construction costs are significant when the influence of function is considered ([(cost/m^2)/function] variable) if only at the 5% level. Thus there is a significant difference in construction cost for upper floor solutions, although there appears not to be such a significant difference for client cost. The relative costs of the different floor types are ranked by cost as before.

There is no consistency between these groups whatsoever. What is perhaps most interesting is the fact that no upper floors does not tend to be the cheapest option. This is probably because it represents a building with only one storey, which is not necessarily the cheapest option. The significance of the difference between these groups was compared by using Bonferroni Multiple Comparisons test. For construction cost/m^2, 'none' is significantly lower than composite steel/concrete and insitu concrete; timber is significantly lower than composite steel/concrete and insitu concrete; precast is significantly lower than composite steel/concrete and insitu concrete; composite steel/concrete is significantly higher than timber, 'none' and precast; insitu concrete is significantly higher than timber, 'none', and precast; while there are no significant differences between steel and the other subgroups. For the (cost/m^2)/function variable, insitu concrete is significantly higher than precast; however, there are no other significant differences between the subgroups.

Independent finishes and openings variables

Descriptive statistics for the construction cost of finishes and openings variables are given in Table 4. There are no categorical variables within this group; they are described in terms of standard costs (See Lowe et al., 2006 for explanation of data representation). The Spearman and Pearson correlation coefficients between construction and client cost and finishes and

Lowe, D J, Emsley, M W and Harding, A (2007) Relationships between total construction cost and design related variables. *Journal of Financial Management of Property and Construction*, **12**(1), 11-23. eScholarID:1a171 | DOI:10.1108/13664380780001090

openings variables are given in Tables 5 and 6 respectively and Spearman correlation coefficients (>0.4) between finishes and openings variables and other variables are given in Table 7.

<<<< Insert Tables 4, 5, 6 & 7 about here >>>>

Ceiling finishes

Low skewness and kurtosis of ceiling finishes suggest that the distribution is approximately normal. However, many projects have no ceiling finishes (i.e. value = 0). Ceiling finishes correlates strongly with internal wall finishes suggesting that there might be scope for producing a variable to indicate the level of finishes generally, although correlations with other variables which might be considered for inclusion in this finish type, but do not correlate so well, are: floor finishes, internal doors and roof finishes. It also correlates well with cost/m^2, showing that buildings of higher specification (and hence cost/m^2) do tend to have a higher cost.

External doors

The external doors variable is well distributed but does not correlate strongly with any variables.

Floor finishes

The distribution of floor finishes is highly skewed because many projects have only a very small value of floor finishes. This is likely because industrial buildings tend not to have floor finishes, and many commercial buildings tend to be fitted out separately from the building contract. It correlates strongly with mechanical installations, cost/m^2, internal wall finishes and function, which can be explained in part by its correlation with function.

Internal doors

Although there is wide range in the value of the internal doors, the vast majority of projects are grouped around the mean, suggesting that the specification system is not robust, or that there is typically very little difference between door specifications. Consequently, there are only two variables which correlate strongly with internal doors: client cost and mechanical installations.

Internal walls

Internal walls is also not very well distributed. Over 86% of the projects are between the values of £13 and £15/m^2. As with internal doors, there are no strong correlations associated with this variable, the highest correlation being with windows (0.193).

Internal wall finishes

The skewness and kurtosis of internal wall finishes indicates that a large number of results are grouped round a single value. Nevertheless, some strong correlations are observed with cost/m^2, ceiling finishes, floor finishes and mechanical installations. This suggests that higher specification finishes do tend to go together.

Roof finishes

Although roof finishes is well distributed, strong correlations are not found with any other variables, although it correlates with the roof construction variable.

Windows

The windows variable is negatively skewed as 28 projects have no window solution. The windows variable does not correlate strongly with any other variable. The highest correlation coefficients are obtained with upper floors, building function, stairs and floor finishes.

Independent services installations variables

Descriptive statistics and tests for normality for the services installations variables are presented in Table 8. Spearman and Pearson correlation coefficients between construction and client costs and services installations variables are given in Tables 9 and 10 respectively. Partial correlation coefficients between construction and client costs and services installations controlling for GIFA are provided in Table 11 and Spearman correlation coefficients between services installations variables and other variables are shown in Table 12.

<<<< Insert Tables 8, 9, 10, 11 & 12 about here >>>>

Air conditioning

High skewness and kurtosis is most likely due to the fact that most buildings do not have air conditioning (median = 0). Air conditioning correlates strongly with lifts, and mechanical installation, the two other 'mechanical' variables (see Table 12). This can be explained by considering the fact that air conditioning tends to be associated with buildings of high specification, which would also tend to have more lifts and more expensive mechanical installations solutions. Air conditioning also correlates reasonably well with protective installations and floor finishes, but does not correlate so well with electrical installations (less than 0.3). It also correlates strongly with log of client cost (0.500), however, the strength of this relationship reduces when partial correlation is applied controlling for GIFA.

Electrical installations

The main group of electrical installations appears to have a slightly positively skewed distribution with mean of approximately 40. However, there is also a much smaller peak at around 8. This shows that there are a small number of buildings with only very basic electrical installations, while the spread of the rest of the installations is quite narrow in comparison with the two groups. This could mean that electrical installations are not very useful as a cost significant variable. This can also be inferred from the fact that there are no strong correlations with other cost significant variables (see Table 12).

Number of lifts

The number of lifts varied between 0 and 8, with 65% of all the projects having no lifts. This skewed the distribution significantly. The number of lifts correlates significantly with a number of variables (see Table 12) predominantly related to mechanical services and the height of the building. The number of lifts also correlates strongly with both construction cost and client costs (see Table 10). However, this correlation may arise from its correlation with GIFA. Therefore partial correlation coefficients, controlling for GIFA, are shown in Table 11, confirming strong correlations with construction cost and client cost.

There appear to be outliers in the data, but this arises from the fact that so few buildings have more than one or two lifts. Therefore the outliers are ignored.

Mechanical installations

Mechanical installations varies between 0 and 84. The mean is 40.66 and the median 40.55. Skewness is -0.584 and kurtosis 0.547. The variable is well spread with a large peak of over 100 projects between 37.5 and 42.5. There is also a small peak at 0 of around 20 projects which have a value of less than 5.

Mechanical installations correlates strongly with the following variables: floor finishes, lifts, protective installations, air conditioning, internal wall finishes, function and internal doors.

The variable correlates highly with all the well distributed finishes, as well as air conditioning and protective installations, suggesting that there exists a single "level of specification" factor which tends to correlate with both finishes and those installations commonly referred to as M&E. Mechanical installations also correlates strongly with both construction cost/m^2 and client costs (see Table 10); while partial correlation coefficients, controlling for GIFA, (see Table 11), confirm these relationships.

Protective installations

Protective installations is highly skewed (value: 1.73). The median value is 3.22, whereas the range is from 0 to 19, and the mean 3.66. Nevertheless, despite this high skew, the values are reasonably well distributed. Protective installations correlates strongly with mechanical installations, lifts, construction cost and client costs (see Table 10); while partial correlation coefficients, controlling for GIFA, (see Table 11), confirm the relationship with client costs.

Special installations

Special installations include such items as fume cupboards and medical gases installations. However, only 28 projects had such installations, while 258 had none. Therefore the value was 0 for over 90% of the projects. This lack of spread means that this variable does not correlate strongly with any variables. The highest correlations were with protective installations and client costs. Overall the lack of spread suggests that this will not be a very useful cost predictor variable.

APPLICATION OF DATA

Data described in this two-paper series were used to develop both regression and neural network cost models. These models have been reported in detail elsewhere (Lowe *et al*., 2006a and Emsley *et al*. 2002 respectively), however, the following brief summary is provided to illustrate the relative importance of the foregoing predictor variables.

Regression analysis

A total of 6 regression models were developed. The number of variables in the model varied considerably. The smallest number of variables used was 8, in the forward stepwise log of cost model. The largest was 14, in both the log of cost and log of cost/m^2 backward models. Throughout the models a total of 19 different variables were used. However, there were two variables for which the both the logarithmic form and the unfactored form were used in different models: function and duration.

The variables are shown in Table 13. They are ranked by the number of times they appear in the models, and then by the mean value of T for inclusion in those models. The descriptor variables which occur consistently are: GIFA, function, duration, mechanical installations, piling, internal wall finishes, frame, site access, protective installations and internal walls.

<<<< Insert Table 13 about here >>>>

The best regression model is the log of cost backward model which gives an R^2 of 0.661 and a MAPE (mean absolute percentage error) of 19.3%.

Neural network models

The results of the regression analysis were used to inform the development of neural network models. Models were developed using the five variables which appeared in all six regression models, the nine variables which appeared in five of the models and all the variables. In order to improve performance and eliminate bias, the ‘voting system’ technique was also adopted, which involves the creation of a number of models each using different training, verification and test sets with the output being taken as the average of these models; in this case 100 networks were used.

The best neural network model is one which uses all 41 variables and a voting system using 100 networks; this gives an R^2 value of 0.789 and a MAPE of 16.6%. Although this is better than the regression model, other neural network models showed a comparable or poorer performance when compared to the regression model. For example, the nine variable model (without using the voting system) had an R^2 value of 0.688 and a MAPE of 20.1%, which is very similar to the best regression model.

The models developed using both techniques compare favourably with past research which has shown that traditional methods of cost estimation are less accurate as evidenced by reported values of MAPE between 20.8% (Skitmore et al., 1990) and 27.9% (Lowe, 1996).

CONCLUSIONS

This, the second of two papers, seeks to redress the omission in recent literature on the influence of design related variables on the cost of construction, by presenting the results of

Lowe, D J, Emsley, M W and Harding, A (2007) Relationships between total construction cost and design related variables. *Journal of Financial Management of Property and Construction*, **12**(1), 11-23. eScholarID:1g171 | DOI:10.1108/13664380780001090

an investigation into the influence of 41 independent variables on both construction cost and client on cost. Data were collected from 286 construction projects in the United Kingdom and analysed to determine the relationships that exist between the dependent and independent variables.

In respect of the structure variables, a significant difference between the five frame types identified was confirmed, and the order, commencing with the cheapest, is: steel, masonry, timber, insitu concrete and precast concrete. Similarly, a cost order was also established for roof construction as: steel, timber, mixed timber/steel, insitu concrete, precast concrete and composite steel/insitu. However, the apparent differences are not significant when the influence of building function is considered. A cost order was also ascertained for substructure as: pad, pad/strip, ground beams, strip, raft. The results in respect of roof profile were interesting, as the data suggested that correct order for roof profile is pitched, curved then flat, with pitched being cheapest and flat the most expensive. Traditionally, flat roofs have been considered as the cheapest, followed by pitched and then curved, so more investigation of the influence of other cost significant variables is necessary. However, the finding supports the claim of Ferry et al. (1999) that a pitched roof is likely to be cheaper than a flat roof of comparable quality. Analysis of the data associated with upper floors and stair types gave inconsistent results. The cost difference between piled and non-piled buildings was also confirmed as being significant.

In respect of finishes and openings variables, some variables demonstrate high correlations with other variables while others do not correlate strongly with any other variable. However, factor analysis could be used to investigate if a variable can be found to indicate the level of finishes generally.

Most services installations variables have high skewness and kurtosis values, showing that the level of services provision is extremely variable. Some high correlations do occur, but they are mostly those which could be expected, such as number of lifts and height. Mechanical installations correlates highly with all the well distributed finishes, as well as air conditioning and protective installations, suggesting that there exists a single “level of specification” factor which tends to correlate with both finishes and M&E installations.

Overall the analyses described in this two paper series have established the core relationships which exist within the data and confirmed many of the relationships which would be expected to occur. These data, therefore, can be confidently used to develop models of the total cost of construction as verified by the development of both regression analysis and neural network cost models.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the support of the Engineering and Physical Sciences Research Council (EPSRC), who funded the research through two grants; our industrial collaborators: Paul Moore - EC Harris, the late Chris Powell – formerly with Tweeds (then Faithful and Gould), Alun Williams – Symonds, and Joe Martin – Building Cost Information Service (BCIS); and the contribution made by the research assistants: Mick Gregory and Adam Hickson and by Dr Roy Duff, former senior lecturer, Manchester Centre for Civil and Construction Engineering, UMIST.

REFERENCES

Ashworth, A. (2004) *Cost Studies of Buildings.* 4th Edition, Pearson, Prentice Hall, Harlow, Essex, UK.

Ferry, D. J., Brandon, P. S. and Ferry, J. D. (1999) *Cost Planning of Buildings* 7th Edition, Blackwell Science Ltd, Oxford.

Lowe, D.J. (1996) *Experiential learning in design cost estimating.* PhD thesis. UMIST, UK.

Lowe, D. J., Emsley, M.W. and Harding, A. (2006) The relationship between total construction cost and project strategic, site related and building definition variables. Submitted to the *Journal of Financial Management of Property and Construction*.

Skitmore, M., Stradling, S., Tuohy, A. and Mkwezalamba, H. (1990) *The accuracy of construction price forecasts.* University of Salford, UK.

Table 1: Classification of input variables

Project strategic variables								
Contract form	CFORM	(O)	Procurement strategy	PROC	(O)	Tendering strategy	TEND	(O)
Duration	DUR	(S/weeks)	Purpose	PURP	(O)			
Site related variables								
Site access	ACCS	(O)	Type of location	LOC	(O)			
Topography	TOP	(O)	Nature of site	SITE	(N)			
Design related variables								
Building definition			Structure			Finishes and openings		
Envelope	ENVLP	(S/m^2)	External walls	EXWL	(S/£)	Ceiling finishes	CEILF	(S/£)
Function	FUNC	(S/£)	Frame	FRAME	(O)	External doors	EXDR	(S/£)
GIFA	GIFA	(S/m^2)	Piling	PILING	(N)	Floor finishes	FLRF	(S/£)
Height	HEIGHT	(S/m)	Roof construction	ROOF	(O)	Internal doors	INDR	(S/£)
Quality*	QUAL	(O)	Roof profile	RFPRO	(O)	Internal walls	INWL	(S/£)
Shape complexity	SHAPE	(O)	Stair types	STAIRS	(O)	Internal wall finishes	INWLF	(S/£)
Storeys above ground	STABV	(S/number)	Substructure	SUB	(O)	Roof finishes	ROOFF	(S/£)
Storeys below ground	STBL	(S/number)	Upper floors	UPFL	(O)	Windows	WINS	(S/£)
Structural units	UNITS	(S/number)				Services installations		
Wall-to-floor ratio	WALFL	(S/ratio)				Air conditioning	AC	(S/£)
						Electrical installations	ELEC	(S/£)
						No. lifts	LIFTS	(S/number)
						Mechanical installations	MECH	(S/£)
						Protective installations	PROT	(S/£)
						Special installations	SPEC	(S/£)

Note: N = nominal, (e.g., piling: 0 = no piling, 1 = piling); O = ordinal (e.g., shape complexity: 0 = low, 1 = medium, 2 = high); and S = scale (e.g., external doors: a proportional scale based on relative costs).

- Note that in previous publications *Quality* has been classified as a project strategic variable. Here it has been re-classified as a design related variable.

Table 2: Descriptive statistics for structure variables

		Cost/m²		Client Cost/m²		(Cost/m²)/ Function	Function
Frame	**Steel**	170	£ 360.59	46	£37.82	0.621	£582.29
	Masonry	89	£ 375.15	13	£43.98	0.619	£610.07
	Timber	5	£ 462.39	2	£69.76	0.756	£624.40
	Insitu concrete	21	£ 550.96	6	£44.03	0.800	£728.57
	Precast concrete	1	£ 560.47	-	-	0.729	£769.00
Piling	**Non-Piled**	228	£ 360.08	53	£38.14	0.613	£590.40
	Piled	58	£466.10	14	£49.83	0.729	£652.86
Roof construction	**Steel**	137	£ 346.78	34	£ 33.41	0.629	£ 346.78
	Timber	112	£ 382.98	25	£ 46.80	0.627	£ 618.50
	Mixed timber/steel	26	£ 482.29	4	£ 47.97	0.673	£ 728.46
	Insitu concrete	6	£ 509.42	2	£ 40.39	0.678	£ 754.00
	Precast concrete	2	£ 598.19	1	£ 84.66	0.776	£ 771.00
	Composite steel/insitu	3	£ 645.56	1	£ 55.54	0.804	£ 805.67
Roof profile	**Pitched**	241	£ 366.22	59	£ 39.06	0.623	£ 592.08
	Curved	23	£ 447.04	2	£ 53.81	0.673	£ 671.61
	Flat	22	£ 481.40	6	£ 51.26	0.742	£ 651.77
Stair types	**Timber**	39	£ 299.44	4	£ 45.73	0.610	£ 490.82
	None	78	£ 347.08	27	£ 38.31	0.634	£ 544.54
	Precast	97	£ 393.48	19	£ 40.43	0.618	£ 651.00
	Steel	34	£ 401.40	8	£ 49.37	0.634	£ 620.97
	Steel/Insitu	9	£ 473.41	2	£ 41.01	0.679	£ 682.56
	Insitu	29	£ 493.29	7	£ 36.66	0.728	£ 705.45
Sub-structure	**Pad**	26	£ 332.66	6	£ 41.58	0.584	£ 563.50
	Pad/strip	102	£ 353.45	26	£ 33.32	0.607	£ 582.05
	Strip	92	£ 371.15	16	£ 51.48	0.615	£ 607.60
	Ground beams	27	£ 447.04	9	£ 49.60	0.700	£ 651.67
	Raft	39	£ 467.04	10	£ 38.50	0.754	£ 640.08
Upper floors	**Timber**	34	£ 308.59	4	£ 45.73	0.635	£ 480.56
	None	75	£ 342.14	26	£ 37.39	0.630	£ 543.19
	Precast concrete	118	£ 367.40	24	£ 39.26	0.606	£ 616.67
	Steel	3	£ 427.85	-	-	0.793	£ 511.67
	Composite steel/concrete	34	£ 486.93	6	£ 34.79	0.654	£ 752.53
	Insitu concrete	22	£ 535.76	7	£ 58.99	0.775	£ 705.05

Table 3: Tests for differences between sub-groups of structure variables

	FRAME		PILING		ROOF		RFPRO	
	F	Chi-S	t	z	F	Chi-S	F	Chi-S
COST	20.270***	52.855***	-3.641***	-4.977***	20.930***	54.167***	23.693***	29.968***
***LN*COST**	18.312***		-5.507***		14.579***		19.768***	
COSTM^2	6.691***	28.172***	-3.924***	-3.997***	6.098***	34.207***	6.556**	11.871**
***LN*COSTM^2**	7.014***		-3.972***		7.143***		6.180**	
CCOST	3.365*	12.679**	-1.834	-2.699**	10.225***	15.068**	39.370***	13.559***
***LN*CCOST**	4.075**		-2.829**		5.465***		14.914***	
CCOSTM^2	1.084	4.218	-1.488	-1.157	1.523	8.405	0.839	2.098
***LN*CCOSTM^2**	0.958		-1.218		1.674		1.074	
COSTFN	3.378**	12.723*	-2.729**	-3.058**	0.679	8.361	2.968	8.678*
FUNC	3.183*	14.951**	-2.271*	-2.532*	6.786***	30.633***	2.700	6.936*
***LN*FUNC**	3.855**		-2.361*		7.906***		2.870	
GIFA	14.486***	54.220***	-3.534***	-4.073***	12.237***	56.416***	16.333***	23.171***
***LN*GIFA**	16.700***		-4.127***		14.167***		13.066***	

	STAIRS		SUB		UPFL	
	F	Chi-S	F	Chi-S	F	Chi-S
COST	16.048***	109.831***	3.103*	18.079***	21.942***	116.610***
***LN*COST**	32.975***		4.679**		47.067***	
COSTM^2	6.017***	30.665***	4.917***	9.550*	9.510***	35.035***
***LN*COSTM^2**	5.521***		4.347**		8.754***	
CCOST	2.778*	23.542***	1.339	5.471	3.983**	26.294***
***LN*CCOST**	6.436***		1.776		8.027***	
CCOSTM^2	0.264	0.269	1.856	2.915	1.077	4.380
***LN*CCOSTM^2**	0.095		1.148		1.100	
COSTFN	1.189	4.744	4.135**	6.087	2.359*	16.682**
FUNC	8.650***	47.467***	1.452	2.201	12.243***	47.160***
***LN*FUNC**	7.945***		1.744		11.193***	
GIFA	13.493***	83.412***	3.738**	22.900***	18.493***	86.995***
***LN*GIFA**	22.926***		6.171***		28.731***	

Table 4: Descriptive statistics for the finishes and openings variables (Expressed as a cost/m^2)

	CEILF	*LN*CEILF	EXDR	*LN*EXDR	EXWL	*LN*EXWL	FLRF	*LN*FLRF	INDR
Mean	8.169	2.035	169.239	4.967	51.856	3.750	10.153	2.112	74.342
Median	10.000	2.303	170.000	5.136	39.000	3.664	6.740	1.946	70.000
Std. deviation	5.641	0.772	102.315	0.755	69.526	0.487	10.475	0.968	35.862
Skewness	.747	-1.816	.195	-0.998	8.427	2.477	1.508	-0.570	5.033
Kurtosis	1.544	3.606	-.035	-0.125	84.775	10.838	3.076	0.169	56.827
Min.	.000	-1.208	.000	3.296	.000	2.795	.000	-0.515	.000
Max.	31.000	3.434	600.600	6.398	875.660	6.775	67.460	4.212	476.990

	*LN*INDR	INWL	*LN*INWL	INWLF	*LN*INWLF	ROOFF	*LN*ROOFF	WINS	*LN*WINS
Mean	4.339	15.025	2.692	6.141	1.731	18.992	2.855	123.101	4.902
Median	4.248	15.000	2.708	6.000	1.792	22.000	3.091	150.000	5.011
Std. deviation	0.226	11.793	0.217	4.0389	0.488	8.094	0.430	45.764	0.171
Skewness	3.358	13.744	7.920	4.831	-0.442	1.703	-0.199	-1.758	-0.572
Kurtosis	18.523	214.389	82.252	41.238	3.813	11.714	-0.865	2.385	-1.087
Min.	4.154	.000	2.485	.000	-0.512	9.000	2.197	.000	4.554
Max.	6.167	200.00	5.298	47.160	3.854	82.160	4.409	175.000	5.165

Table 5: Spearman correlation coefficients between construction and client costs and finishes and openings variables

		COST	COSTM^2	CCOST	CCOSTM^2
CEILF	**286**	0.101	0.494***	0.294*	0.387***
EXDR	**286**	0.268***	0.095	0.103	-0.052
EXWL	**286**	0.346***	0.368***	0.412***	0.227
FLRF	**286**	0.373***[3]	0.572***	0.432***	0.272*
INDR	**286**	0.390***	0.347***	0.506***	0.119
INWL	**286**	0.149*	0.099	-0.018	-0.224
INWLF	**286**	0.268***	0.570***	0.462***	0.350**
ROOFF	**286**	0.017	-0.070	-0.263*	-0.056
WINS	**286**	0.275***	0.240***	0.285*	-0.087

Table 6: Pearson correlation coefficients between construction and client costs and finishes and openings variables

		COST	*LN* COST	COSTM^2	*LN* COSTM^2	CCOST	*LN* CCOST	CCOSTM^2	*LN* COSTM^2
CEILF	**286**	0.236***	0.202***	0.468***	0.521***	0.361**	0.363**	0.335**	0.364**
***LN*CEILF**	**249**	0.093	-0.030	0.420***	0.505***	0.194	0.130	0.371**	0.470***
EXTDR	**286**	0.142*	0.229***	0.068	0.041	0.133	0.118	-0.050	-0.116
***LN*EXDR**	**271**	0.202***	0.278**	0.078	0.044	0.146	0.189	-0.178	-0.199
EXWL	**286**	0.531***	0.332***	0.304***	0.271***	0.464***	0.517***	0.151	0.2621
***LN*EXWL**	**285**	0.549***	0.426***	0.418***	0.423***	0.435***	0.493***	0.197	0.321**
FLRF	**286**	0.336***	0.400***	0.504***	0.508***	0.364**	0.438***	0.276*	0.276*
***LN*FLRF**	**238**	0.291***	0.335***	0.512***	0.575***	0.332*	0.408**	0.333*	0.364**
INDR	**286**	0.142*	0.227***	0.250***	0.292***	0.421***	0.509***	0.103	0.096
***LN*INDR**	**268**	0.251***	0.337***	0.243***	0.249***	0.496***	0.526***	0.002	0.014
INWL	**286**	0.133*	0.166**	0.186**	0.183**	0.028	0.118	-0.049	-0.118
***LN*INWL**	**277**	0.139*	0.172**	0.123*	0.102	-0.094	-0.067	-0.209	-0.272*
INWLF	**286**	0.332***	0.258***	0.500***	0.472***	0.250*	0.333**	0.345**	0.327**
***LN*INWLF**	**276**	0.265***	0.235***	0.517***	0.540***	0.299*	0.353**	0.326**	0.327**
ROOFF	**286**	0.067	0.061	-0.033	-0.045	-0.244*	-0.276*	-0.083	-0.087
***LN*ROOFF**	**286**	0.028	0.029	-0.090	-0.109	-0.270*	-0.297*	-0.108	-0.123
WINS	**286**	-0.002	0.122*	0.202***	0.238***	0.143	0.258*	-0.034	-0.067
***LN*WINS**	**258**	0.253***	0.347***	0.177**	0.132*	0.067	0.167	-0.137	-0.152

Table 7: Spearman correlation coefficients between finishes and openings variables and other variables

CEILF		**EXWL**		**FLRF**	
INWLF	0.563	RFPRO	0.404	MECH	0.585
ROOF	0.532			INWLF	0.562
FLRF	0.476			FUNC	0.560
FRAME	0.457			INDR	0.493
FUNC	0.421			CEILF	0.476
MECH	0.417			LIFTS	0.467
				PROT	0.448
				AC	0.442
INDR		**INWLF**		**ROOFF**	
MECH	0.502	CEILF	0.563	ROOF	0.501
FLRF	0.493	FLRF	0.562		
LIFTS	0.451	MECH	0.541		
INWLF	0.402	FUNC	0.487		
STABV	0.400	STABV	0.468		
		DUR	0.428		
		INDR	0.402		

Table 8: Descriptive statistics for services installations variables (Expressed as a cost/m^2)

	AC	*LN*AC	ELEC	*LN*ELEC	LIFTS	*LN*LIFTS
N	286	66	286	283	286	99
Mean	13.249	3.604	37.034	3.566	.636	0.410
Median	.000	4.094	39.150	3.667	.000	0.000
Std. deviation	31.112	1.131	9.736	0.407	1.211	0.577
Skewness	2.359	-0.785	-2.104	-3.075	2.863	1.158
Kurtosis	4.360	-0.715	4.702	8.789	9.900	0.196
Minimum	.00	1.386	.00	1.946	.00	0.000
Maximum	145.00	4.977	55.60	4.018	8.00	2.079
	MECH	***LN*MECH**	**PROT**	***LN*PROT**	**SPEC**	***LN*SPEC**
N	286	284	286	267	286	28
Mean	40.660	3.454	3.664	0.846	.223	0.721
Median	40.550	3.703	3.220	1.215	.000	0.737
Std. deviation	17.372	1.059	3.509	1.223	.733	0.519
Skewness	-.584	-2.896	1.730	-0.866	3.516	-1.978
Kurtosis	.547	7.978	4.185	-0.227	12.706	6.529
Minimum	.00	-1.050	.00	-1.897	.00	-1.204
Maximum	84.05	4.431	18.91	2.940	4.75	1.558

Table 9: Spearman correlation coefficients between construction and client costs and services installations variables

		COST	COSTM^2	CCOST	CCOSTM^2
AC	**286**	0.470***	0.409***	0.534***	0.162
ELEC	**286**	0.188***	0.105	0.099	0.050
LIFTS	**286**	0.631***	0.394***	0.610***	0.180
MECH	**286**	0.473***	0.592***	0.544***	0.306*
PROT	**286**	0.579***	0.416***	0.623***	0.113
SPEC	**286**	0.243***	0.195***	0.263*	-0.008

Table 10: Pearson correlation coefficients between construction and client costs and services installations variables

		COST	***LN*COST**	**COSTM^2**	***LN*COSTM^2**	**CCOST**	***LN*CCOST**	**CCOSTM^2**	***LN*CCOSTM^2**
		286	**286**	**286**	**286**	**67**	**67**	**67**	**67**
AC	**286**	0.344***	0.407***	0.371***	0.359***	0.446***	0.500***	0.135	0.163
***LN*AC**	**66**	0.065	0.115	0.285*	0.360**	0.095	0.211	-0.027	0.078
ELEC	**286**	0.149*	0.140*	0.212***	0.214***	0.231	0.226	0.143	0.122
***LN*ELEC**	**283**	0.085	0.067	0.209***	0.219***	0.167	0.207	0.158	0.140
LIFTS	**286**	0.726***	0.630***	0.395***	0.373***	0.746***	0.630***	0.213	0.222
***LN*LIFTS**	**99**	0.673***	0.660***	0.437***	0.401***	0.556**	0.439	0.264	0.288
MECH	**286**	0.407***	0.448***	0.550***	0.579***	0.507***	0.557***	0.246*	0.258*
***LN*MECH**	**284**	0.171**	0.203***	0.363***	0.391***	0.277*	0.330**	0.243*	0.237
PROT	**286**	0.600***	0.587***	0.355***	0.313***	0.678***	0.685***	0.189	0.175
***LN*PROT**	**267**	0.395***	0.475***	0.391***	0.349***	0.433***	0.511***	0.111	0.095
SPEC	**286**	0.236***	0.240***	0.244***	0.198***	0.275*	0.283*	-0.016	0.015
***LN*SPEC**	**28**	0.020	0.010	0.172	0.120	0.896*	0.895*	0.971***	0.985***

Table 11: Partial correlation coefficients between construction and client costs and services Installations controlling for GIFA

		COST	***LN*COST**	**COSTM^2**	***LN*COSTM^2**	**CCOST**	***LN*CCOST**	**CCOSTM^2**	***LN*CCOSTM^2**
		286	**286**	**286**	**286**	**67**	**67**	**67**	**67**
AC	**286**	0.296***	0.376***	0.361***	0.358***	0.372**	0.446***	0.225	0.292*
***LN*AC**	**66**	0.325**	0.332**	0.287*	0.356**	0.074	0.239	-0.025	0.074
ELEC	**286**	0.130*	0.103	0.205***	0.211***	0.180	0.174	0.193	0.191
***LN*ELEC**	**283**	0.156**	0.100	0.209***	0.219***	0.120	0.173	0.199	0.199
LIFTS	**286**	0.560***	0.352***	0.420***	0.422***	0.665***	0.493***	0.395***	0.471***
***LN*LIFTS**	**99**	0.519***	0.482***	0.465***	0.484***	0.429	0.223	0.306	0.295
MECH	**286**	0.405***	0.429***	0.547***	0.585***	0.395***	0.467***	0.377**	0.441***
***LN*MECH**	**284**	0.231***	0.249***	0.359***	0.389***	0.209	0.280*	0.311*	0.334**
PROT	**286**	0.237***	0.256***	0.372***	0.350***	0.498***	0.512***	0.427***	0.492***
***LN*PROT**	**267**	0.177**	0.332***	0.388***	0.362***	0.230	0.346**	0.274*	0.306*
SPEC	**286**	0.098	0.113	0.230***	0.192***	0.209	0.220	0.037	0.090
***LN*SPEC**	**28**	0.014	-0.001	0.176	0.124	0.842	0.942*	0.954*	0.973**

Table 12 Spearman correlation coefficients between services installations variables and other variables

AC		**ELEC**		**LIFTS**	
LIFTS	0.592	PROT	0.261	AC	0.592
MECH	0.572	MECH	0.245	UPFL	0.581
PROT	0.451	LIFTS	0.194	MECH	0.579
FLRF	0.442	INDR	0.183	HEIGHT	0.577
UPFL	0.435			STAIRS	0.577
STAIRS	0.403			GIFA	0.537
HEIGHT	0.416			PROT	0.534
				STABV	0.530
				FLRF	0.467
				INDR	0.451
				INWLF	-0.445
				ENVLP	0.443
				FUNC	0.439
				STBL	0.426
				DUR	0.413
MECH		**PROT**		**SPEC**	
FLRF	0.585	MECH	0.579	PROT	0.298
LIFTS	0.579	LIFTS	0.534	MECH	0.238
PROT	0.579	UPFL	0.479	WINS	0.201
AC	0.572	GIFA	0.474		
INWLF	0.541	STAIRS	0.469		
FUNC	0.514	AC	0.451		
INDR	0.502	FLRF	0.448		
DUR	0.460	HEIGHT	0.432		
STABV	0.454	INWLF	-0.421		
QUAL	0.428				
STAIRS	0.426				
CEILF	0.417				
UPFL	0.401				

Table 13: Variables included in the regression models

	Cost/m^2		Log Cost/m^2		Log Cost			
	For.	**Back.**	**For.**	**Back.**	**For.**	**Back.**	**Total**	**Mean T**
*LN*GIFA	•	•	•	•	•	•	6	17.215
*LN*FUNC			•		•		2	9.686
FUNC	•	•		•		•	4	9.078
*LN*DUR		•	•	•		•	4	5.188
DUR	•				•		2	4.535
MECH	•	•	•	•	•	•	6	4.072
PILING	•	•	•	•	•	•	6	3.492
INWLF	•		•	•	•	•	5	3.885
FRAME	•		•	•	•	•	5	2.943
ACCS	•		•	•	•	•	5	2.809
PROT	•	•	•	•		•	5	2.466
INWL		•		•		•	3	2.386
SUB		•					1	4.038
WALFL				•		•	2	2.919
SPEC		•					1	2.884
EXWL		•					1	2.573
FLRF		•					1	2.572
*LN*HEIGHT				•		•	2	2.553
UNITS				•		•	2	2.163
ELEC				•		•	2	2.056
Total	**9**	**11**	**9**	**14**	**8**	**14**		

Lowe, D J, Emsley, M W and Harding, A (2007) Relationships between total construction cost and design related variables. *Journal of Financial Management of Property and Construction*, **12**(1), 11-23. eScholarID:1a171 | DOI:10.1108/13664380780001090