This is the author’s version of a work that was submitted/accepted for publication in the following source:


Selected for Construction Industry Council (CIC) ‘Industry Challenge’ Showcase Papers– ranked 1st

This file was downloaded from: https://www.escholar.manchester.ac.uk

© Copyright 2002 The Authors

Reproduced in accordance with the copyright policy of the publisher. Notice: Changes introduced as a result of publishing processes such as copy-editing and formatting may not be reflected in this document. For a definitive version of this work, please refer to the published source.
The Costs of Different Procurement Systems: A Decision Support Model

David Lowe, Manchester Centre for Civil and Construction Engineering, UMIST, UK
Margaret Emsley, Manchester Centre for Civil and Construction Engineering, UMIST, UK
Roy Duff, Manchester Centre for Civil and Construction Engineering, UMIST, UK

ABSTRACT

Existing research which has attempted to determine differences between the costs of the different procurement routes has consistently aimed to determine a single figure for the difference for projects as a whole. No attempt has been made to provide a difference which is project specific (Duff et al., 1998). Furthermore, no previous research has determined the cost to the client using any objective method.

The absence of such a technique is significant. It means that the client’s advisors have no means of providing an objective measure of the cost of following different procurement routes. The client must depend upon the judgement of the advisors, which is based on their own perception of both the project and the different procurement routes, and is hence subject to their opinions and prejudices.

This paper reports on the development of a neural network model (ProCost) which is able to determine the total cost to the client of a project, it functions as a decision support tool by enabling the project specific comparison of alternative procurement routes and other strategic decisions.

Keywords: Cost Modelling, Decision Support, Early Stage Estimating, Neural Networks, Procurement

INTRODUCTION

Masterman (1994) found no evidence of cost differences being considered when choosing a building procurement system. This discovery prompted a proposal to the Engineering and Physical Sciences Research Council (EPSRC), in 1997 for a one-year pilot study into the potential problems of producing a cost model that would predict the cost differences of different procurement systems. The successful completion of this study led to a further two-year main investigation, again funded by EPSRC.

Past research that attempted to determine differences between the costs of different procurement routes has aimed to determine a mean percentage difference for projects as a whole (Building Economic Development committee, 1974; Department of Industry and Department of the Environment, 1982; Brandon et al., 1988, University of Reading, 1991). No attempt has been made to determine a difference, on a project-by-project basis, that allows for the differing effect of other project variables (Duff et al., 1998). Furthermore, no previous research has attempted to determine the total cost to the client, including the client’s own internal and external costs. This means that the client’s advisors have no means of providing an objective measure of the relative total cost of following different procurement strategies. The client must depend upon the subjective judgement of advisors.

The simplest and most effective way of determining this cost difference between procurement routes is to create a model that is able to predict the cost of the project from all the cost significant variables, including procurement route. This model can then be used to evaluate the expected costs of the project under different procurement routes, taking into account the differences in other project variables. In order to accurately evaluate the differences in cost between the different procurement routes for any particular project, the model is required to model the complex and little understood interrelationships that exist between all the cost significant variables whose values can be determined at the early stages of the project. Because the relationships between the different variables are so difficult to identify and quantify, a neural network was considered to be the most appropriate modelling tool to use.
This paper presents the findings of research that has been funded through the two EPSRC awards, which total in excess of £200,000.

**Aim, Objectives and Key Advances of the Research**

The aim of the research was to develop reliable models of project cost that would allow clients and their advisors to forecast the total cost of a future project, including both clients’ own administration and consultancy costs and the final cost of construction, as well as enabling the project specific comparison of alternative procurement routes. The specific objectives of the project were to:

1. Collect sufficient data to fulfil initial model requirements;
2. Develop predictive cost models;
3. Validate the models;
4. Produce a user-friendly model interface;
5. Disseminate the results.

These objectives have been achieved. Further, the main advances resulting from the research are:

1. Cost models based on a large set of reliable data;
2. Significant improvements in data representation in the model building;
3. Much greater understanding of the nature of the relationships in the data and of the most appropriate ways to represent the dependent variable, cost;
4. Models of construction cost which appear to be more accurate than current professional practice;
5. Innovative models of clients’ own project costs, in addition to construction.

**METHODOLOGICAL ISSUES**

**Data Collection**

Despite the abilities of neural networks to model any function, the accuracy that they achieve is dependent upon the number of projects used in training the models. This poses a problem for using neural networks for this research, because collecting data from past projects is a difficult and time-consuming process; and, the number of projects ideally required for modelling construction costs is between 400 and 500, though less for clients’ costs due to the smaller number of significant variables.

Projects were defined, initially, in terms of around 40 variables. The variables fall into three main groups: project strategic variables, such as procurement and tender strategy; site related variables, such as the topography and site access; and design related variables, such as the frame type and gross internal floor area (Duff et al., 1998, Harding et al., 1999b, Harding et al., 2000a).

Previous research has generally used the tender price to evaluate the cost of the project, whereas the cost to the client of a building contract is the final contract sum. In addition, the whole cost to the client includes professional consultants fees, planning and building regulations fees and the cost of any internal resources the client has to provide (Harding et al., 2000a). Therefore, the following project costs were collected:

a) The final cost of the construction contract;
b) External and internal client costs.

Data collection was undoubtedly the most difficult problem faced during the research programme. The initial target was 500 sets of project data to be collected by visits from the research assistants. Achievement of this target was initially expected through three collaborating quantity surveying practices, supplemented by a number of construction client
organisations that had indicated their support and a small number of additional QS practices. These sources proved insufficient and, in the case of some clients who had originally expressed support, unhelpful. A number of alternative strategies had to be adopted. First, every opportunity was taken to publicise the research through presentations to over 30 RICS Regional Branches, individual QS practices and client organisations, where demonstrations of some of the early model software were made. Many fruitful contacts resulted in this way. Second, experience in data collection from a particularly helpful client was used to develop a questionnaire style data form and every QS practice in the UK Yellow Pages was telephoned to ask for assistance by completing at least one of these. Over 1200 positive responses were followed up by the postal questionnaire, supported by a letter from the RICS. This process yielded only a few additional projects. Third, the suitability of RICS Building Cost Information Service (BCIS) data was evaluated and a supplementary questionnaire prepared, to cover data not currently collected by BCIS. This was circulated to the quantity surveyors for the 285 suitably recent projects and yielded 29 additional sets of project data.

After all these efforts, the database now contains 288 projects.

Data Representation/Transformation

The input data to the model (independent variables) were of three types, ratio, ordinal and nominal. In the case of ratio data (e.g. number of lifts, duration, height) there was no difficulty in its representation. In the case of ordinal data (e.g. quality of finishes: low; medium; high), the neural network training was assisted, where possible, by representing the data by cost ratios derived from available cost data. The nominal data (e.g. building function, electrical and mechanical installations [heating, air conditioning]) were handled in one of two ways. Where figures for relative cost were available, then the same approach was taken as with ordinal. For variables like building function, where the cost influence can not be estimated, they were treated as a set of dummy, binary variables, one for each option, that take the values yes/no. These problems and the solutions adopted are described fully in Harding, et.al. (1999b).

The research has produced two classes of model, one to predict construction cost and one to predict clients’ costs. In each case, various methods of representing the output (dependent variable) were tested (Harding et al., 2000b). Use of the raw cost figure was found to be the least useful as it is so highly correlated with building size, represented by Gross Internal Floor Area (GIFA), that the effect of virtually every other variable was being masked. This casts some doubt on the validity of some published results of previous attempts to model project costs. This problem is particularly evident in modelling construction cost. Clients’ costs are less highly correlated with building size and, therefore, the problem is less evident. Log of raw cost was tried, to improve the normality of the cost distribution and reduce the influence of small numbers of very large, and therefore expensive, projects. However, ultimately, two other transformations of raw cost, cost/unit of floor area and log of cost/unit of floor area, were found to be the more satisfactory, the choice depending on the particular objective and circumstances of the cost prediction (Emsley, et.al., 2002).

Model Building

Although neural networks had been shown, in the pilot study (Duff et al., 1998), to be a more satisfactory technique than regression modelling, regression was used, for two purposes, in the early stages of modelling. The first was to provide a benchmark of model capability, with which to compare the neural network models. The second was to assist in identifying the relative importance of the input variables, in order to inform the manipulation of neural network input data and speed up the model building process (Harding et al. 2000b, Emsley et.al., 2002). TRAJAN Neural Network Simulator Release 4.0E software was used.

In order to maximise the accuracy of the model, it was necessary to employ techniques that maximise the accuracy of the neural network for a fixed number of projects. This was done by:

a) Analysis of the inputs to the neural network to try to reduce the number of input variables, and so reduce the number of data sets required for training a satisfactory model, by identifying and eliminating the variables having the least significant effects. This was attempted by stepwise regression, factor and cluster analyses and the generalised regression neural network approach (GRNN) (Emsley, et.al., 2002). In fact, all proved
unsuccessful in reducing the input to the model as the best performing models, explaining the greatest amount of cost variation, use all the project variables.

b) Making maximum use of all the data by repetitive training of many models, using a different sample of data for training each time – the ‘voting system’ technique.

The original intention was to use the cross validation technique, rather than $R^2$, to validate the neural network models. However, it was decided to replace this approach with the ‘voting system’ approach, considered to be a most effective way of maximising the value of limited data sets by producing, in effect, a mean of many models to produce the best representation of the available data. It involves the creation of a number of models, each of which uses different training, verification and test sets. The output cost is then taken as the mean of the output of these models. Bias exists within individual models, which use only a proportion of the data set for training (the rest used for verification and testing). However, provided all the projects in the data set are represented in the training sets for the many models developed, individual model bias is compensated.

While determining the best architecture was difficult, because there is little significant difference between the performances of the different architectures, there are significant differences between the performances of different configurations of training, verification and testing sets, a result which validated the use of the voting system approach.

**Reliability and accuracy**

The reliability and accuracy of the models can be summarised as follows. The most accurate model to date (Mean Absolute Percentage Error – 16.6%) has been obtained using all variables and a voting system of 100 networks. This model has been able to explain nearly 80% of the variability in the cost data ($R^2 = 0.789$), showing significant improvement over the best regression model (MAPE = 19.3%; $R^2 = 0.661$).

Past research has shown that traditional methods of early stage cost estimation give accuracies, represented by the MAPE of between 20.8% (Skitmore, et.al., 1990) and 27.9% (Lowe, 1996), significantly less accurate than our neural network models. However, both Lowe and Skitmore found that the same early stage estimators perceived their accuracy of prediction to be less than 10%, suggesting that estimators have an unrealistically high opinion of the quality of their estimating ability. This conclusion was also drawn by Birnie (1993). The discrepancy between the actual and perceived degree of accuracy of forecasts poses a significant problem that needs be overcome before practitioners can accept the model.

**User Interface with Model**

A user-friendly interface with the model has been developed in C++ and connected to the best neural network to provide a decision support model: ProCost. The interface is shown in figure 1 below. Projects can be saved and opened as documents in any normal Windows application, using the Open, Save, and Save As… commands in the File menu. Individual variables can be selected and edited as desired using the edit bar above the data (similar to Excel). Variables are shown in groups, which can be expanded and rolled up by clicking on the group titles. Alternative, project information can be imputed by means of a Data Entry Wizard, which is selected via the Data menu. To avoid imputing every variable for each new project, it is envisaged that ‘standard’ projects are utilised for each project type, which can be opened and amended to suite the project under consideration.

**IMPACT AND BENEFITS OF RESEARCH**

The impact of the research has, so far, been to raise a great deal of interest in the many practitioners who have attended the presentations and demonstrations. It is not possible, at present, to evaluate this, other than subjectively, but the team members who made the presentations are very positive about the genuinely enthusiastic reception that was received. The strong impression remains that interest in the research is increasing as practitioners become familiar with the research objectives and output. Further, our collaborating organisations, E C Harris, Symonds, Faithful and Gould and RICS Building Cost Information Service remain committed and continue to participate in the project.
The potential benefits from using ProCost include:

- Reduced costs from better informed choices of procurement system;
- Early and accurate forecast of the probable cost of a future project;
- Ability to assess options quickly with a minimum of resources;
- Customer Satisfaction

EXPLOITATION

Analysis of the data, to explore the possibility of increasing the accuracy and range of cost models, is continuing. Further, field-testing in QS offices is currently being undertaken and further dissemination events will be mounted as soon as this is completed. Discussions concerning the further expansion of the cost database, possibly in collaboration with BCIS, to expand the cost model and its range of application are continuing. Further, licensing of the software and the development of a professional cost estimating service, probably in collaboration with BCIS, is planned, again, when the field-testing is completed.

CONCLUSIONS

Selection of the most appropriate project strategy, including procurement system, contract strategy, building type etc., is fundamental to a successful project. By collecting and analysing a large database of the costs of adopting different
project strategies, on recent building projects, computer neural network models of clients’ own costs and construction costs have been developed. The models accept data on 41 variables, which define the project, and output estimates of both clients’ own administrative and consultancy costs, and the final cost of construction.

The neural network models are trained on data from almost 300 projects and provide instantaneous comparative costings of various strategic project options. This will enable project managers and quantity surveyors to give UK construction clients quantified cost based advice on a variety of project strategies.

All the objectives of the research project have been met. The accuracy of the model expressed in terms of the mean absolute percentage error is 16.6% and the model explains nearly 80% of the variability in the cost data ($R^2 = 0.789$). The model performs well against traditional methods of early stage cost estimation give accuracies, represented by the MAPE of between 20.8 and 27.9%. However, the discrepancy between the actual and perceived degree of accuracy of forecasts poses a significant problem that needs be overcome before practitioners can accept the model.

A user-friendly software interface in C++ has been developed with a help facility; this will obviate the need for an extensive hardcopy user guide, and attached to the front end of the model. Dissemination has, so far, comprised over 30 seminars and demonstrations in many parts of the UK and eight academic and professional publications.

The main technical advances from this research include the application of a number of approaches to data representation, manipulation and modelling, and the development of a robust early stage estimating decision support tool (ProCost) that predicts the comparative costs of different procurement options. Further, the models provides a ‘What if?’ analysis, enabling user to assess how changing certain characteristics of the building affects the project cost.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the support of the EPSRC and our industrial collaborators: Paul Moore - EC Harris, Chris Powell – Tweeds (now with Faithful and Gould), Alun Williams – Symonds, and Joe Martin – Building Cost Information Service (BCIS); and the contribution made by the research assistants: Mick Gregory, Adam Hickson and Anthony Harding.

REFERENCES

Birnie, J. W., A behavioural study using decision analysis of building cost prediction be Chartered Quantity Surveyors (1993), unpublished PhD thesis,
Building Economic Development Committee (1974), Before you build - what a client needs to know about the construction industry, NEDO, London.
Department of Industry & Department of Enviroment (1982), The United Kingdom Construction Industry - a guide to methods of obtaining a new industrial building in the UK, Invest in Britain Bureau, London.


