# MODELLING LEVELS OF ACCURACY IN DESIGN COST ESTIMATION

**D. J. Blackwood[1], S. Sarkar[1] and A. D. F. Price[2]**
*[1]School of Construction and Environment, University of Abertay Dundee, Dundee, DD1 1HG, UK*
*[2]Dept of Civil and Building Engineering, Loughborough University, Loughborough, LE11 3TU, UK*

The use of fee competition by construction industry clients has increased the need for design consultants to predict more accurately the cost of their professional services. There is evidence of concern among clients that fee reductions resulting from fee competition have been accompanied in many cases by a reduction in the level of services provided by consultants. However, there is little published work on design cost estimation methods and their accuracy. This paper describes a recently completed programme of research in design cost estimation practice in the United Kingdom. An overview of design cost estimation practice is presented and the development and operation of a model which simulates the design cost estimation process is described. The paper explains and presents the results of a data collection strategy that was adopted to provide input for the model and conclusions are drawn on the validity of the model when compared to the indicative levels of accuracy in design cost estimation suggested in the literature. Finally, recommendations are given for future research work that should lead to a usable cost estimation tool for design consultants.

Keywords: Design costs, design management, cost modelling, simulation.

## INTRODUCTION

Rowdon and Mansfield (1989) stated that design consultants earn less than they should due to a lack of well developed planning, cost estimation and cost control systems. The pressures of fee competition further increased the need for more accurate cost estimation by designers and for clients to have a greater understanding of the adequacy of design fee bids. The relationship between clients and their professional advisors has being fundamentally altered by external pressures. The philosophy behind competition amongst consultants on the basis of quality, with cost being determined by fee scales, had been under review by the Monopolies and Mergers Commission since 1977 (Rowdon and Mansfield, 1988) and was vigorously attacked by the Office of Fair Trading (Anon, 1991a). Despite opposition from professional organisations (e.g. Anon, 1991b; Anon 1991c), fee competition became widely established in the construction industry. Concern was being expressed that design quality was being adversely affected (Construction Industry Council, 1992; Latham, 1994). In addition to the more competitive environment amongst external providers of design services, Government proposals to extend compulsory competitive tendering to the provision of professional services (Pigg, 1993) and the recent switch of emphasis to demonstrating value for money in the provision of public services have increased the need to ensure that in-house design costs were comparable with those in the private sector.

This paper outlines a programme of research investigating the effectiveness of design cost estimation practice that was initiated in response to the above developments. The

Blackwood, D J, Sarkar, S and Price, A D F (1998) Modelling levels of accuracy in design cost estimation. *In:* Hughes, W (Ed.), *14th Annual ARCOM Conference*, 9-11 September 1998, University of Reading. Association of Researchers in Construction Management, Vol. 1, 330-9.

research methodology and a more detailed analysis of the results are described fully elsewhere, (Blackwood, Sarkar & Price, 1997; Blackwood 1998) but elements contributing to the selection of the modelling methodology are reviewed and the development and testing of a model which represents design cost estimating accuracy is described and evaluated.

## SURVEY OF DESIGN MANAGEMENT PRACTICE

The principal means of data collection was a postal questionnaire but a number of in-depth interviews with design organisations were performed in order to ensure the appropriateness of the questionnaire design. In total, eleven design offices were selected at random for the interviews to give a representative range of construction industry design organisations. The sample comprised:

- the head office of a large firm of consulting engineers (800 staff);
- five regional offices of a consulting engineering practice (staff numbers within each office varied between 20 and 100);
- the head office of a multi-disciplinary design organisation (100 staff);
- the regional office of a medium sized architectural practice (20 staff);
- the office of a small architectural practice (10 staff);
- a local authority water services department design section (80 staff); and
- the design section of a large organisation which provides, internationally, large design and construct energy related projects (100 staff).

The interviews identified that a surprisingly consistent approach to estimating design cost and planning design work was adopted by the organisations. However, the sample was too small to enable meaningful conclusions on the nature of estimating and planning in the construction industry at large and there was a need for a more comprehensive survey of design organisations to enable a wider evaluation to be made of the extent of usage of the various cost estimation approaches by designers. The interviews also suggested that detailed historic cost data would not be readily available in design organisations and that this would clearly influence the potential for rational cost modelling. The survey was extended to ascertain, with greater certainty, the types of data that would be available within design organisations.

A sample frame was assembled which comprised all organisations involved in the provision of design services for construction industry related projects, including; civil engineering consultants, architectural practices, national and local government departments, water authorities, building services engineering consultants and multi-disciplinary design organisations. In view of the large number of organisations involved, some form of sampling from the overall population was required. In this case, there was a considerable amount of information available on the nature of the sample and, thus, a combination of stratified proportionate and two stage random sampling was appropriate. A total target sample size of 300, evenly split between four groupings, was chosen in order to generate sub-sample for the four categories of greater than 30 responses assuming a response rate of 40%.

The results of the questionnaire survey will be reported fully elsewhere but in the context of cost modelling, the main conclusion was that design cost estimation in the construction industry relied predominately on the application of intuitive approaches and that historic cost data were not available in a form which would enable detailed cost models to be constructed. This conclusion supports the observation by Rowdon and Mansfield (1989) that designers believe that the nature of design work is such that it is not amenable to prescriptive planning and rigorous cost control. Whilst this finding suggested that currently the potential for the development of a rational approach to design cost estimation was limited, no firm conclusions could be drawn until alternative approaches to data collection and cost modelling had been fully investigated.

## DATA COLLECTION AND ANALYSIS

It was apparent from the empirical surveys, that, although design cost data were available in design organisations, this lacked detail. The data had been collected from staff time sheets that had been devised to support the organisations cost control systems. These only required projects to be broken down into a small number of broad work packages. If other cost modelling approaches were to be considered it would be necessary to compile a database of design costs comprising a larger number of smaller work packages for each project. Thus, the data collection phase of the research design served two purposes:

- to provide some detailed cost data for the model development and, more importantly in the context of testing the overall research hypothesis;
- to allow an evaluation to be made of the practicality of the collection in industry of a substantial volume of detailed cost data for estimation purposes.

There were two possible strategies for this stage of the work. One strategy would involve data collection from a number of projects across a number of organisations which would give rise to a large volume of cost data. However, there were two major drawbacks to this approach, namely:

- it would rely heavily on the involvement of a large number of employees in the participating organisations and discussions with possible collaborators during the interviews indicated that they would be unable to commit the necessary resources; and
- the collaborators would be unwilling to provide a large volume of commercially sensitive cost information.

The second strategy was to adopt a case study approach to data collection and this was deemed to be appropriate because:

- the case study provides an opportunity for intensive analysis of a single instance of the topic of study. In this case it allowed the researcher to become involved in the data collection process within each of the collaborating organisation which provided an in depth understanding of the problems associated with the process and thereby enabled meaningful conclusions to be drawn on the practicalities of widespread application of the data collection methodology; and
- a more limited trail of the data collection methodology was attractive to the collaborators.

It was decided that the case study should involve projects with a range of durations in different organisations but it was realised that the number of projects and organisations would have to be restricted to enable the data collection to be effectively managed and monitored. A total of four projects with durations of two to sixteen months were selected from three organisations. Following a comprehensive review of literature on staff cost monitoring systems, a specification was devised for the data collection system which gave due regard to the following aspects:

- the cost estimate and the cost control data should be the "property" of the design team;
- the cost estimate should be derived by breaking the project down to the smallest practicable work packages;
- the cost control data should be collected against the smallest practicable work packages to enable useful data to be produced for future cost estimating;
- a bar chart approach to programming the work should be adopted; and
- a graphical output system which is readily assessable by the design team should be provided.

The components of the input system comprised a project bar chart, spreadsheets for building up the cost estimate and spreadsheets to permit the direct entry of weekly time sheet data and the project managers estimate of actual physical progress on each work package. The output from the system served the two purposes namely: the provision of monitoring information for the project under consideration and the provision of planned and actual cost data for a database of historic costs. The most appropriate output for monitoring a project was deemed to be a record of planned progress (which would also represent planned cost), actual progress and actual cost and this could either be viewed on screen or printed on hard copy in tabular or graphical form. The output system consisted of a master spreadsheet to collate and process the input data and produce tabular and graphical output. The interaction between the various components of the system is shown in Figure 1.

The following data were available from three case study projects for the cost model:

- estimated person-hours and costs broken down to the level 2 work packages, and
- actual person-hour and cost data for level 1 work packages.
- The data could be applied to the cost model in three forms (O'Keefe 1994):
- disaggregated person-hours, i.e. total person-hours grouped by grade of employee;
- aggregated person-hours i.e. total person-hours on the work packages; and
- total cost which will be a cost weighted equivalent of aggregated hours.

An analysis of the accuracy of the estimates of disaggregated person-hours, aggregated person-hours and cost was undertaken and an example of the results are presented in Table 1 below. A scale was devised to represent estimating accuracy in terms of the inaccuracy of the initial estimate when compared to the actual person hour requirements as follows:

$$\text{Estimating Inaccuracy}(\%) = \frac{\text{Actual Value}}{\text{Estimated Value}} \text{ X } 100$$

This scale ranges between 0% and + ∞%, with 100% representing total accuracy. The magnitude of the inaccuracy is represented by the distance from the 100 % point of the scale.

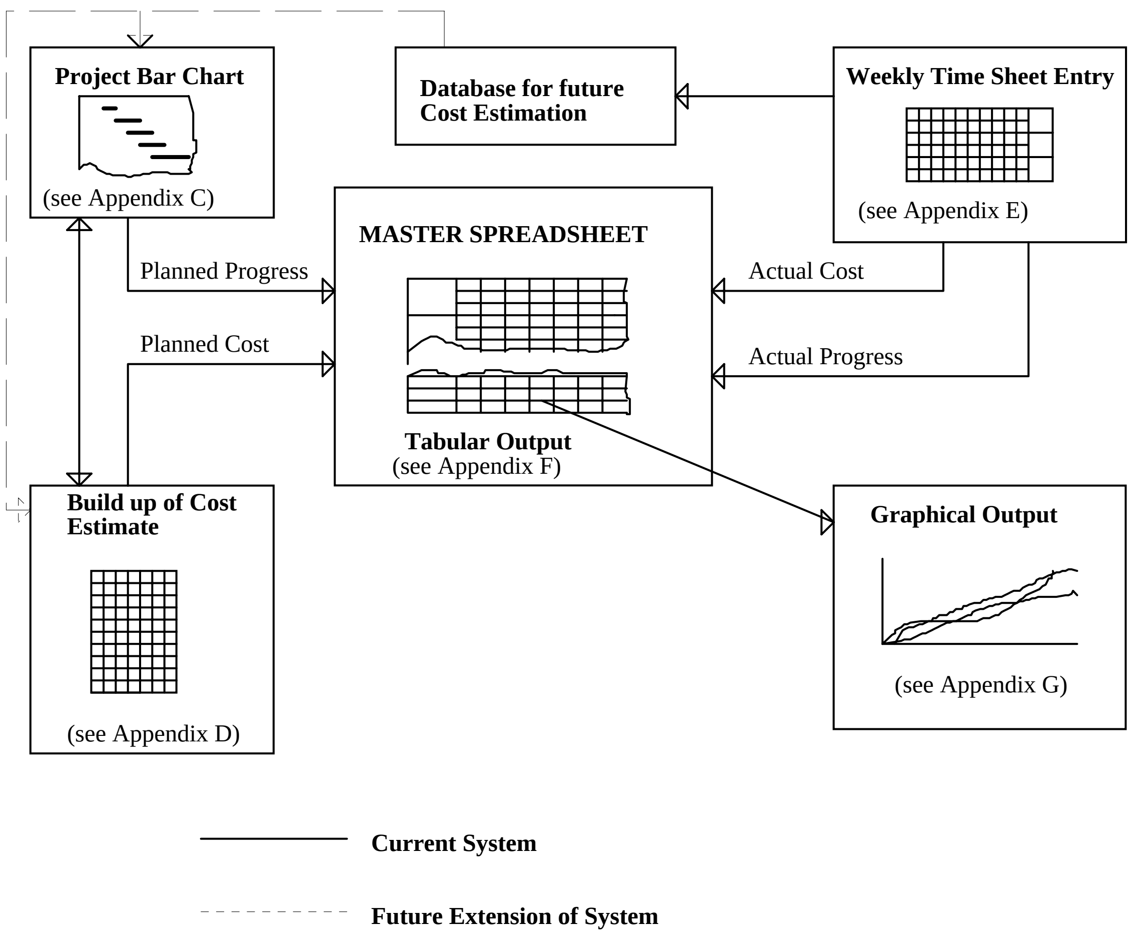


**Figure 1**: Representation of the data collection and cost monitoring spreadsheet system

## DEVELOPMENT AND TESTING OF COST ESTIMATION MODEL

The literature review identified three possible cost modelling approaches and each was critically appraised against the criteria given above. The approaches were:

- parametric cost models (e.g. Boehm, 1981; O'Keefe, 1994);
- simulation models (e.g. Davis and Cochrane, 1987; Cornwell and Modianos 1990), including those used in risk analysis (e.g. Kidd, 1991; Thomson and Perry 1992; Hudson, 1992; Touran and Bolster, 1994; Uher, 1996); and
- fuzzy logic (e.g. Kangari and Bakheet, 1994; Ock 1996).

The full appraisal of the approaches is not given here but a qualitative summary is presented in Table 2. The table demonstrates that the risk analysis approach was the most appropriate form of model.

The risk analysis model developed comprised a WBS representation of the design process and project specific structures were created during the production of the "bottom-up" cost estimate for the projects in the case study. A typical WBS qualitative model for the case study projects is shown in Figure 2.

**Table 1:** Estimated and actual disaggregated person-hour data for level 1 work packages

| | | Estimated person-hours | | | | Actual person-hours | | | | Inaccuracy (%) | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Project | Grade | 1 | 2 | 3 | 4 | 1 | 2 | 3 | 4 | 1 | 2 | 3 | 4 |
| A | WP1 | 33 | | 7 | | 82 | | 60 | | 248 | | 857 | |
| | WP2 | 44 | | 16 | | 71 | | 10 | | 161 | | 63 | |
| | WP3 | 7 | | 10 | | 19 | | 6 | 10 | 271 | | 60 | n/a |
| | WP4 | 13 | | 29 | | 28 | | 140 | | 215 | | 482 | |
| | WP5 | 42 | | 66 | | 47 | | 133 | | 112 | | 202 | |
| | WP6 | 17 | 10 | 10 | | 19 | 30 | 37 | | 112 | 300 | 370 | |
| B | WP1 | 14 | 42 | 7 | | 0 | 1 | 0 | | n/a | 2 | n/a | |
| | WP2 | 42 | 71 | 0 | | 22 | 157 | 1 | | 52 | 221 | n/a | |
| | WP4 | | 84 | 63 | | | 51 | 32 | | | 61 | 51 | |
| | WP5 | | 70 | 35 | | | 55 | 65 | | | 78 | 185 | |
| | WP6 | | 136 | 105 | | | 83 | 65 | | | 61 | 62 | |
| | WP7 | | 56 | 49 | | | 23 | 36 | | | 41 | 73 | |
| | WP8 | | 56 | 49 | | | 10 | 18 | | | 18 | 37 | |
| | WP9 | | 42 | 0 | | | 41 | 0 | | | 98 | n/a | |
| | WP10 | | 54 | 72 | | | 91 | 72 | | | 168 | 100 | |
| C | WP1 | | 0 | 15 | 54 | | 5 | 1 | 62 | | n/a | 7 | 115 |
| | WP2 | | 0 | 0 | 290 | | 20 | 38 | 370 | | n/a | n/a | 127 |
| | WP3 | | 54 | 54 | 145 | | 19 | 38 | 182 | | 35 | 70 | 126 |
| | WP4 | 36 | 36 | 73 | 0 | 22 | 26 | 23 | 58 | 61 | 72 | 31 | n/a |
| | WP5 | | | 40 | | | | 11 | | | | 28 | |

**Table 2**: Evaluation of the alternative modelling approaches

| | Criteria | | | |
|---|---|---|---|---|
| **Modelling approach** | **Availability of appropriate input data** | **Suitability of output data** | **Transparency** | **Potential for practical application.** |
| Parametric cost modelling | Low | Moderate | High | High |
| Simulation | Moderate | High | High | Low |
| Risk analysis | Moderate | High | High | High |
| Fuzzy set theory | Moderate | Moderate | Low | Low |

This qualitative model allowed two forms of estimating model to be devised:

- models which used input probability distributions of the dis-aggregated person-hour estimates at level 2 of the work breakdown structure for Projects A and B; and
- models which used input probability distributions for the aggregated person-hour estimates at level 1 of the work breakdown structure for Projects A, B and C.

The level 2 models were developed and tested initially, prior to the development testing and validation of the level 1 models. The data came from the case studies, described earlier. In addition to the selection of the appropriate range of estimating inaccuracy, a decision had to be made on the most appropriate form of input distributions. Perry and Hayes (1985) suggested that cost distributions should be skewed to represent a greater probability of cost over-run. The use of unsymmetrical distributions was universally adopted by other authors (e.g. Raftery 1994, Kidd 1991, Davis and Cochrane 1995, Cornwell and Modianous 1990) with the use of a range of distributions including: step rectangular (histogram), triangular, beta and log normal being reported. The @Risk package contained the following unsymmetrical distributions: histogram, triangular, beta, beta subjective, beta pert, gamma, log

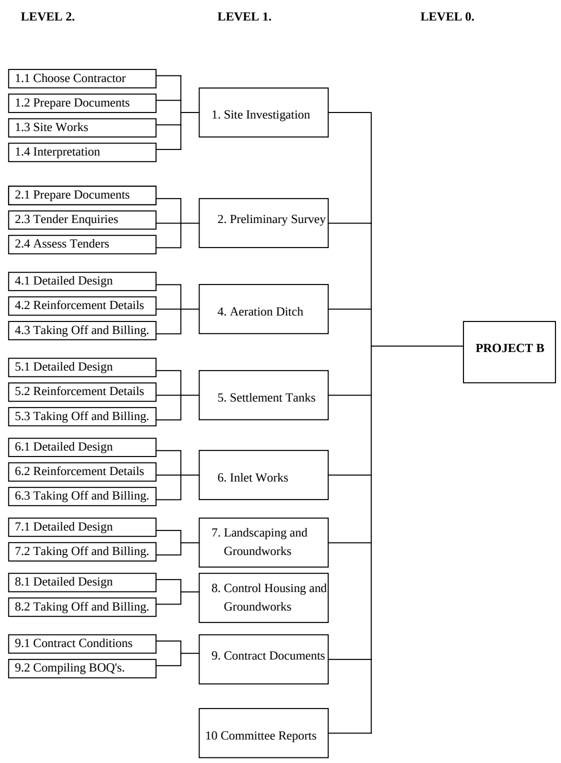


**Figure 2**: Project B work packages

**Table 3**: Comparison of actual and simulated project total hours

| | **Simulated Hours** | | | | | | | **Simulation Inaccuracy** | |
|---|---|---|---|---|---|---|---|---|---|
| | **Min.** | **Mean** | **Max.** | **Mode** | **5%ile** | **95%ile** | **Actual Hours** | **Mean (%)** | **Mode (%)** |
| Proj. A | 220 | 414 | 875 | 351 | 284 | 593 | 688 | 60 | 51 |
| Proj. B | 522 | 1047 | 2234 | 1095 | 713 | 1526 | 864 | 121 | 127 |
| Proj. C | 595 | 868 | 1274 | 874 | 706 | 1055 | 873 | 99 | 100 |

normal, Pearson and Rayliegh, and each were considered using the RiskView sub-package which enables a graphical representation of the probability distribution to be created before their inclusion in the cost model. Following extensive testing of the input distributions, the use of log normal distributions in the estimating models was found to be the most appropriate. The selection of appropriate 99.5 percentile values of the distributions was also shown to be of utmost importance in simulating the effects of the risk in cost estimation with 99.5 percentiles value of 400 per cent of the mode (or estimated value) being appropriate for low risk activities, adjusted to 600% for higher risk activities in model. The results of the simulations are shown in Table 3 and these can usefully be compared by observations on estimating accuracy by other authors as shown in Table 4.

The most useful direct comparison between the simulated results and those from other sources can be made using the data from Nicolson and Popovic’s (1994) survey for “analytical estimating”, where the estimates were also prepared using a work packages breakdown approach. This survey produced fee bids between 60% and 210% of the mode value and in each case the simulated mode values fall within this range.

The performance of the simulation model was further verified through its application to an additional project (Project D) where data was made available on the initial estimates to a level 2 WBS and final costs to level 1 WBS. As a comparison, the predictions of general model was compared with that of a model developed to

**Table 4**: Suggested ranges of estimating inaccuracy

| Source | Minimum estimate (%) | Maximum estimate (%) | Comments |
|---|---|---|---|
| Hudgins and Lavelle (1995) | 50 | 250 | Total project estimate for rough concept design work |
| | 70 | 160 | Total project estimate for general concept design work |
| | 80 | 130 | Total project estimate for detailed design work |
| Nicolson and Popovic (1994) | 60 | 180 | Total project estimate using "scale method" |
| | 60 | 180 | Total Project Costs using "bench mark method" |
| | 50 | 210 | Total project estimate using "analytical method" |
| Kydd (1991) | 50 | 200 | Time for tasks in software development |
| Hudson (1992) | 60 | 150 | Total design costs on a project |

represent the estimator's perception of his estimating accuracy of work package level costs, which the estimator believed to be between ± 5 and 25% of his estimate for each work package. The results of the simulations are shown in Table 5 and Figure 3.

Table 5 showed that the mean of the output distribution for the general model of Project D (which used the standard Log normal input distributions) was virtually identical to the total hours that were required for the project. It could be argued therefore that the standard model has been demonstrated to be valid and that the mean predicted by the general model is a useful indicator of required person hours for a project in a risk seeking environment. However, before concluding on the validity of the model, it is also necessary to consider the significance of cumulative probability distributions for the model's output. This demonstrated that the actual hours correspond with the 57 percentile point of the distribution which suggests that there was a 43% chance that the project person hour requirements could exceed this value. The validity of the general model can only be established if it could reasonably be concluded that the project estimator's personal risk attitude and levels of accuracy were such that the mean prediction of actual hours could be accepted with confidence. In the absence of other data, the estimator's attitude and levels of accuracy can only be measured using the data that was available from this project. The estimator's inaccuracy was between 120 and 173% with an average of 142%.

This demonstrated that the estimator is risk seeking and it could be argued that this risk attitude has contributed 142% of the overall inaccuracy in the estimate. Against this benchmark, the estimator's performance with respect to general inaccuracy within work packages is extremely good, with inaccuracy in the individual work packages being within -23 and +30% of the benchmark figure. The perceived model of estimating accuracy has shown (Figure 3) that work package estimation at this level of accuracy would predict total actual hours that were within a very narrow band of the mean prediction. In the context of prediction of the general model it would be reasonable to accept the mean predicted value when the estimator was known to be risk seeking but was otherwise consistently accurate. Therefore, the general model has been shown to be robust in its application to Project D.

**Table 5**: Summary of analysis of project D

| | Simulated proj total hours.(actual project total hrs=1928) | | | | | | Simulation inaccuracy | |
|---|---|---|---|---|---|---|---|---|
| | **Min.** | **Mean** | **Max.** | **Mode** | **5th percentile** | **95th percentile** | **Mean (%)** | **Mode (%)** |
| General model | 986 | 1916 | 4047 | 1754 | 1260 | 2810 | 99 | 91 |
| Model of perceived accuracy | 1272 | 1301 | 1328 | 1299 | 1287 | 1313 | 67 | 67 |

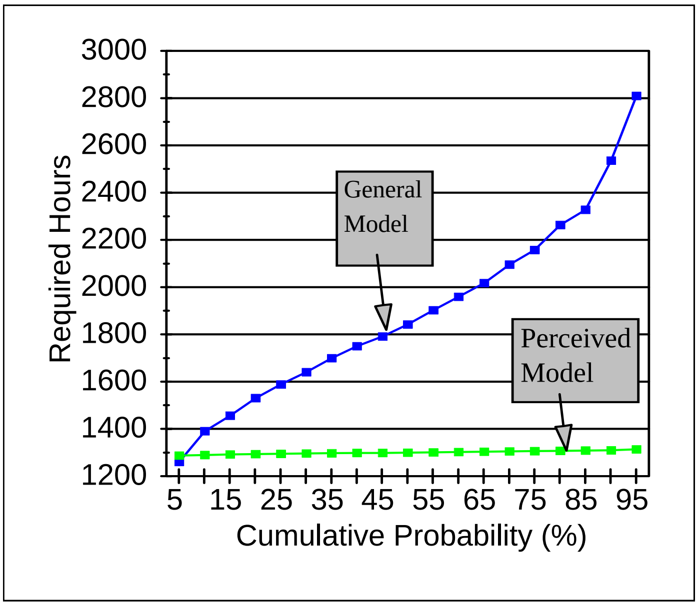


**Figure 3:** Simulated cumulative probability distributions

## CONCLUSIONS

It has been shown that the cost modelling approach described above can provide a realistic representation of design cost estimating accuracy and therefore has the potential to be developed as a tool to assist designers in the preparation of design bids. However, further development and more rigorous testing of the model is required including a full evaluation of the impact of interdependencies between input variables. The key to further development is the establishment of a data base of design cost information related to standard work packages and the research team are currently investigating the use of more efficient means of data capture within design organisations.

## REFERENCES


Anon. (1991a). OFT renews war on professional scales. *New Builder*. 24 January: p8.

Anon. (1991b). Practice. *RIBA Journal*, June: p 25.

Anon. (1991c). ACE gathers fee bid evidence. *New Civil Engineer*, 17 October: p 6.

Blackwood, D. J. (1998) *Rationalisation of design cost estimation in the construction industry*. Unpublished PhD theses, School of Construction and Environment, University of Abertay Dundee.

Blackwood, D. J., S. Sarkar and A.D.F. Price (1997) A research methodology for modelling construction design service costs, *Journal of Construction Procurement,* **3**(3)**,** 47-67.

Boehm, B. (1981) *Software engineering economics*, Englewood Cliffs, NJ: Prentice Hall.

Construction Industry Council (1993) *The procurement of professional services*. London: Thomas Telford.

Cornwell, L.W. and D.T. Modianos (1990) Management tool: using spreadsheets for simulation models. *Production and Inventory Management Journal*, **31**(3), 8-17.

Davis, C. and S.R. Cochrane (1987) Optimisation of the future manpower requirements in a multi-discipline consultancy. *Construction management and Economics.* **5**(1)**,** 45-56.

Hudgins, D. W. and J.P. Lavelle (1995) Estimating engineering design costs. *Engineering Management Journal*, **7**(3), 17-23.

Hudson, C. (1992) A technique for assessing cost uncertainty. *American Association of Cost Engineering Transactions*, F.2.1-F.2.5.

Kangari, R. and M. Bakheet (1994) Linguistic analysis of risk. *Proceedings of the First Congress on Computing in Civil Engineering*, **1**, 216-219.

Kydd, J. B. (1991) Do today's projects need powerful network planning tools? *International Journal of Production Research*, **29**(10), 1969-1978.

Latham, M. (1994) *Constructing the team*, HMSO. London.

Nicolson, M. P. and O. Popovic (1994) Towards developing a methodology for architects' fee bids. *The International Journal of Architectural Management, Practice and Research*, **4**, 65-77.

Ock J. H. (1996) Activity duration quantification under uncertainty: fuzzy set theory application. *Cost Engineering*, **38**(1), 26-30.

O'Keefe, T. D., D. A. Simunic and M. T. Stein, (1994) The production of audit services: evidence from a major public accounting firm. *Journal of Accounting Research*, **32**(2), 241- 265.

Perry, J. G. and R.W. Hayes (1985) Risk and its management in construction projects. *Proceedings of the Institution of Civil Engineers*, **78**(1), 499-501.

Pigg, D.R. (1993) An introduction to compulsory competitive tendering for professional engineering services. *Proceedings of the Institution of Civil Engineers*, **98**, 69-78.

Raftery J. (1994) *Risk analysis in project management.* London; Spon.

Rowdon, I. J. and N.R. Mansfield (1988) What price professionalism? *Proceedings of the Institution of Civil Engineers*, **84**(1), 85-93.

Rowdon, I. J. and N.R. Mansfield (1989) Controlling design practice. *Proceedings of the Institution of Civil Engineers.* **86**,(1), 189-205.

Thomson, P and J. Perry (1992) *Engineering construction risks, a guide to project risk analysis and risk management*. London; Thomas Telford.

Touran, A. and P.J. Bolster (1994) Integration of financial and construction risks: a simulation approach. *Transportation Research Record*, **1450**, 15-22.

Uher, T. E. (1996) A probabilistic cost estimating model. *Cost Engineering*, **38**(4), 33-40.