

QUALITY IN STRATEGIC COST ADVICE: THE EFFECT OF ANCHORING AND ADJUSTMENT

Chris Fortune¹ and Mel Lees²

¹*School of the Built Environment, Liverpool John Moores University, Clarence Street, Liverpool L3 5UG, UK*

²*Research Centre for the Built and Human Environment, University of Salford, Bridgewater Building, Salford, M7 9NU, UK*

The formulation of early stage building project cost advice for clients requires the professionals concerned to exercise judgement. The exercise of judgement is a human cognitive process that can be subject to errors, bias and heuristics. One of the biases that affects judgement is termed “anchoring and adjustment”. This study seeks to add to the literature related to judgement in early cost advice by ascertaining whether construction professionals are prone to make judgements that are biased by their reliance on the anchoring and adjustment heuristic.

The paper reports the development of an appropriate measuring instrument and the results of its application to a group of thirty-four subjects. The subjects were a convenience sample drawn from a cohort of final stage part-time students in quantity surveying. Subjects were tested on their propensity to make biased judgements via word problems that were set in their own subject specific domain. The results of the work revealed that the subjects displayed the same level of error in response to the context based word problems as had been displayed in previous studies in which subjects responded to word problems set in non-work related contexts. The paper concludes by setting out the case for further empirical work in this area in order to address the impact of this and other biases on a wider sample of construction professionals.

Keywords: adjustment, anchoring, bias, error, judgement.

INTRODUCTION

Fortune and Lees (1996) reported an empirical study that obtained a practitioner assessment of the relative performance of the building project cost modelling techniques actually used in practice. The extent to which a particular cost model relied upon judgement was found to be a factor that influenced model's incidence-of-use. The results of that study showed that judgement could be considered as being either a positive or a negative influence on the quality of advice provided as perceived by the practitioner. The positive component is the need of the practitioner to adjudicate on the raw outcome of a process or a technique. The negative side is the nature of the human judgement and the propensity of individuals to make errors. Thus it can be seen that research designed to improve the quality of early stage strategic cost advice must also address the development of a better understanding of the role of human judgement in its formulation. Previous work has established that practitioners make errors due to anchoring and adjustment bias in arriving at judgements when confronted with problems unrelated to their industry context - see Fortune and Lees (1998). In addition the potential for humans with particular learning characteristics to make similar cognitive errors has also been investigated - see Fortune and Lees

(1997). This study seeks to make a further contribution to the research in the field of the formulation of strategic cost advice by ascertaining whether construction professionals have a propensity to make errors of judgement, due to anchoring and adjustment bias, when the problems are set in an industry context.

The paper firstly sets out the wider context for the study and then reports on the development and application of an appropriate measuring instrument to thirty-four subjects drawn from a cohort of final stage part-time undergraduate students in surveying. The results of the investigation are then analysed using *Minitab for Windows (v10)* and the paper concludes by outlining a plan for future action that will further contribute to the improvement of quality in decision-making in the field of strategic cost advice.

CONTEXT

Following a review of literature related to early stage building project price forecasting and judgement Raftery (1995) asserted that reliable strategic cost advice required the input of human judgement. However, both Raftery (1995) and Birnie (1995) pointed out that humans make mistakes when making judgements and they stated that more work was needed to understand the behavioral processes involved.

Tversky and Kahneman (1974) and Kahneman, Slovic and Tversky (1985) provided a comprehensive review of the literature related to the existence of systematic biases that affect judgement. They asserted that in making judgements under uncertainty people in general do not appear to follow the calculus of chance or the statistical theory of prediction, instead they rely upon a number of simplifying strategies or heuristics that direct their judgements. Such heuristics can sometimes lead to reasonable judgements and sometimes lead to severe and systematic errors. The heuristics that were acknowledged as being generalisable across the population were (1) the availability heuristic, (2) the representative heuristic and (3) the anchoring and adjustment heuristic. The potential biases attributed to each of these heuristics were listed by Bazerman (1993) as being :- ease of recall, retrievability, presumed associations (the availability heuristic), insensitivity to base rates, insensitivity to sample size, misconceptions of chance, regression to the mean, the conjunction fallacy (the representativeness heuristic), insufficient anchor adjustment, conjunctive and disjunctive events, overconfidence (the anchoring and adjustment heuristic). Mak and Raftery (1992) acknowledged that the majority of the research in cognitive psychology has led to a common understanding and acceptance of the existence of the above listed heuristics and biases in lay people thinking intuitively and making judgements about problems. However, they pointed out that there was as yet no consensus in the literature on bias. In particular they noted that there was little empirical evidence of the propensity for bias in judgements made by experts considering context related problems.

One such study was carried out by Mak and Raftery (1992) in an experiment with quantity surveying students in a simulated real life price forecasting situation. Their conclusions indicated that there was little support for the existence of severe and systematic bias and that the previous research findings on the existence of generalised bias may have been too pessimistic when practitioners were asked to make judgements on matters within their own field. However, Mak and Raftery (1992) went on to point out that their work had methodological limitations and they

suggested that further work with experienced practitioners and subjects from different institutions would be needed to validate their work.

This study seeks to add to the empirical evidence so far collected on practitioners judgements on word problems set in their own subject related domain. The study has centred on the ascertainment of practitioners' propensity to be affected by the anchoring and adjustment bias when making judgements on work related word problems

METHODOLOGY

General

The central problem facing this study was the re-working of the measuring instrument used in the previous study (Fortune and Lees 1997) to introduce an industry context. The instrument takes the form of a series of problems that the subject is required to attempt. In the previous study these were taken from the literature sources set out above and were, therefore, not set in a construction context. Beach *et al.* (1987) criticised the approach of asking practitioners to solve problems that were not set in the context appropriate for their expertise. They argued that this would inevitably lead to evidence of error as the subjects did not apply themselves to the task in hand. The current study focuses on the anchoring and adjustment bias but the previous study had covered the four main sources of error - cognitive, availability, representative and anchoring. The test used in that study had contained twelve questions - three for each of the main types of error. Each question dealt with a particular sub-type of the main error. Originally, the test included 36 questions with each sub-type having three questions randomly spread throughout the test, but piloting suggested that this made the test too long as the subjects became disinterested and the results less valid. Since the current study was limited to one of the main error types - anchoring and adjustment - the number of questions for each sub-type could be increased. The three sub-types of the anchoring and adjustment bias together with the question numbers used in the test are set out in Table 1.

The original test question on the 'conjunctive and disjunctive' sub-type error was based on the selection of coloured marbles from a bag and had its roots in statistical probabilities. This question was set in the context of drawing coloured bricks from a pack containing bricks of two different colours. The statistical background to the questions were not affected by this change.

Table 1: Test questions and sub-types of errors

Anchoring and adjustment bias	Test question number
Sub-type of error	
insufficient anchor adjustment	1, 4, 7
conjunctive and disjunctive	2, 5, 8
overconfidence	3, 6, 9

For the 'insufficient anchor adjustment' questions related to the salaries of various professions. Instead of using salaries the average building prices of three different building types were used. The information was taken from the Building Cost Information Service and was current at the time the test was administered.

'Overconfidence' in the original test involved the subject in guessing the population of various countries and then providing an upper and lower limit that would be their 95% confidence limits for the real value. This question was changed so that the subject had

to provide several rates for items of work contained in a price book that was current in 1992. Again, the 95% confidence limits were required.

By re-working the original test as described above the new test was developed with an industry context. For each of the sub-types of error three versions of the original question were developed and this meant that there were nine questions in total.

Piloting

The new test was piloted on a small group of practitioners to establish whether it could be understood and whether it was of an appropriate length. The results of the pilot indicated that the text of the test was appropriate and that the time required to undertake the test was short enough to retain the interest of the subjects.

Given the limited resources available for the study it was decided to establish a convenience sample of thirty-four subjects drawn from the full cohort of part-time students attending final level surveying degree courses at the University of Salford. Previous involvement in the formulation of early construction cost advice was the criterion set for inclusion in the subject frame and the measuring instrument was applied in the spring of 1998.

RESULTS

The propensity for error test used questions developed in context as described above. In addition to responding to the questions the subjects were asked to indicate on a scale of 1 to 4 how confident they were that their response was correct (1 - not at all sure, 4 - very sure the answer is correct). Therefore, for each question not only was it possible to identify whether an error had been made (ie an incorrect answer), but it was also possible to express the degree of error by using the confidence response.

For example, if a question required the respondent to indicate between two alternatives, A and B, one of the alternatives would be the correct answer. Answering incorrectly would indicate an error. But if the respondent was not sure about their response they could indicate a level of confidence of 1, if they were very confident then they could indicate a higher confidence score. An incorrect answer with low confidence is arguably not an error at all.

The scoring system took this into account and was based on the confidence of the response minus one (ie a confidence level of 1 became 0, 2 became 1 and so on), which if the answer was incorrect, was expressed as a negative number. Therefore, the available scores for any question were +3, +2 and +1 for correct answers, ie no error; 0 were the confidence was low and it could not be assumed that an incorrect answer was indicative of an error; and -1, -2 and -3 for incorrect answers. For each sub-type of error there was a score for each of the three questions and these scores were averaged to produce an overall result for the sub-type for each subject. The results are set out in Table 2.

Table 2: Results of error test (negative figures in brackets, blanks indicate incomplete responses)

Test subject	Insufficient anchor adjustment	Conjunctive and disjunctive	Overconfidence
1	(2.00)	(1.00)	
2	0.00	1.30	0.00
3	(2.50)	(2.67)	(1.00)
4	(2.00)	(1.33)	(0.67)
5	2.00	(1.67)	(1.00)
6	(2.00)	(2.33)	(1.33)
7	(1.00)	(1.33)	(1.00)
8	(0.67)	(1.67)	
9	(1.50)	(0.67)	(1.00)
10	(1.00)	0.00	
11	(1.50)	(1.00)	
12	0.00	0.00	0.00
13	(0.33)	(1.00)	(1.00)
14	(1.00)	(0.33)	0.00
15	(1.33)	(0.67)	(2.67)
16	(1.00)	0.00	0.00
17	(0.33)	(0.67)	
18	(1.00)	(1.67)	
19	(1.00)	(1.30)	
20	(0.67)	(0.67)	
21	(1.50)	0.50	(2.00)
22	(1.00)	0.00	
23	(2.00)	(1.67)	(0.50)
24	(1.00)	(1.00)	
25		0.00	
26	(1.50)	(1.33)	(1.33)
27	(1.00)	0.00	0.00
28	(1.00)	(1.33)	(1.33)
29	(0.50)	(2.00)	0.00
30	(1.50)		
31	(1.33)	(2.00)	(0.67)
32	(1.00)	(1.67)	(0.50)
33	(3.00)	(1.33)	0.00
34	1.00	(1.67)	(0.33)

ANALYSIS

The analysis of the results was carried out using *Minitab for Windows (v10)*. The data shown in Table 2 was used to construct a distribution for each sub-type of error. The distribution shows how the sample as a whole responded in terms of errors of judgement. The scale of +3 to -3 is a continuous scale moving from high certainty of correctness to high incidence of error. There are two possible 'null' scenarios for the expected distribution. The first is that the test is not taken seriously by the subjects and that answers and confidences are generated at random. This would mean that each confidence level had the same chance of being indicated in any given response and would result in a horizontal flat line distribution curve. The second is that the subjects do not commit errors and would indicate a level of confidence of 1 to all questions they suspected they may have got wrong. Since all confidence levels of 1 were re-graded to 0 this would result in a distribution that occurred on the positive side of the x-axis only. The actual results are shown in figs 1, 2 and 3.

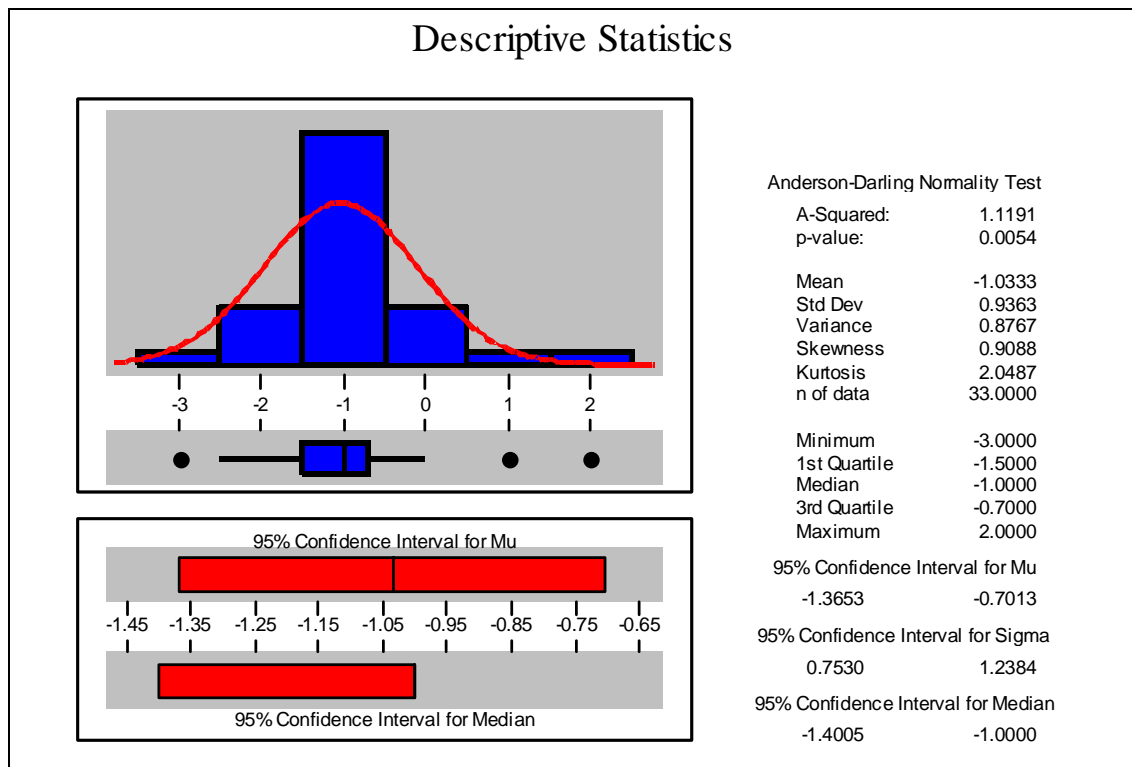


Figure 1: Distribution of responses for 'insufficient anchor adjustment'

The questions about 'insufficient anchor adjustment' required the respondent to estimate the average building prices in £/m² of gross floor area for 'one-off' dwellings, offices and factories. In each case the respondent was given fictitious information about the most recent equivalent project from within their organisation. This information was positioned as an anchor in the highest or lowest 20% of the distribution of costs as taken from BCIS. An answer was deemed correct if it was within +/- 20% of the mean value given by BCIS, an error was made if the response was nearer the anchor than the mean value and a null was recorded if the answer was not correct but farther away from the anchor than the mean.

The distribution is clearly centred around negative values. The mean is given as -1.0333 and the 95% confidence limits for the mean are both negative at -1.3653 and -0.7013. The standard deviation is 0.9363. The distribution indicates that, as a group, the subjects make systematic errors of judgement by failing to make sufficient adjustment to information given to them. This means that surveyors are likely to be influenced by the data that is given to them and work with that data rather than be truly objective. This makes them susceptible to questions being 'framed' and likely, therefore, to making errors.

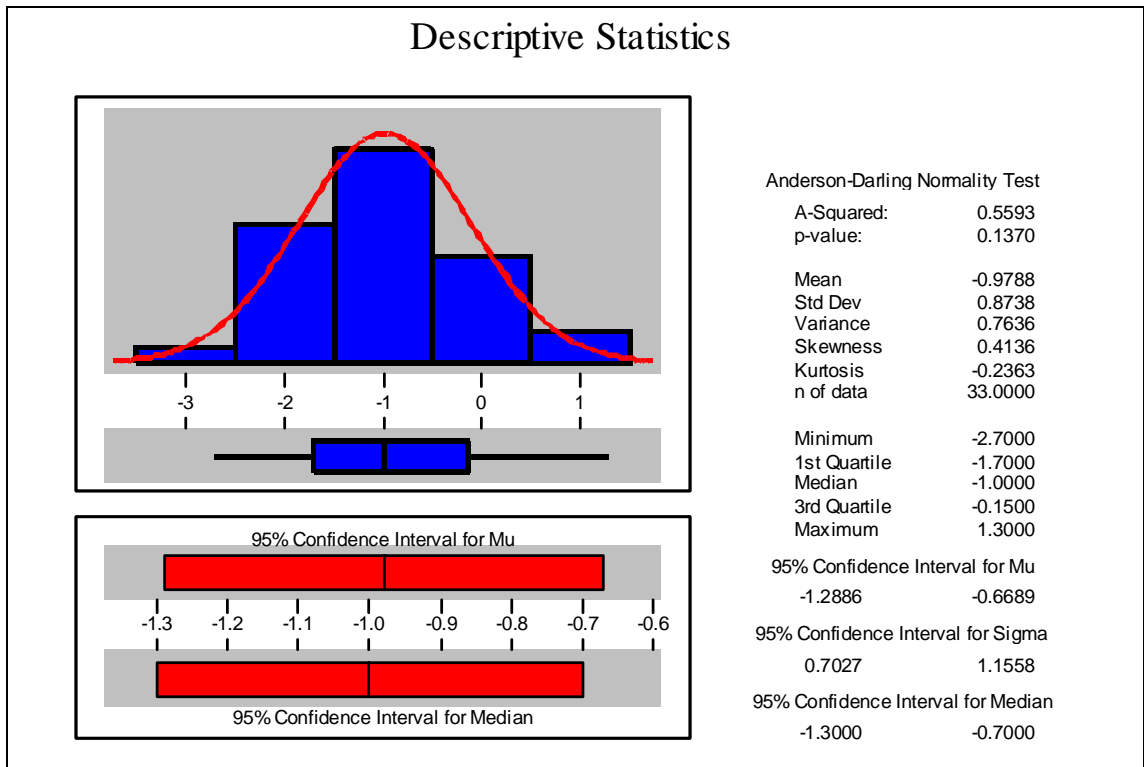


Figure 2: Distribution of responses for ‘conjunctive and disjunctive’

For the ‘conjunctive and disjunctive’ fallacy the questions are based on probabilities. Coloured bricks are drawn from a stack containing two types of coloured brick. In each case the question provides three different scenarios, ie the stack contains different proportions of the two coloured bricks and the requirement for success also changes. An example of this would be ‘...drawing a blue brick three times in a row from a stack containing 35% blue and 65% red bricks’. The respondent is asked to place the three scenarios in order of likely success, ie which is most likely to succeed and so on. Placing the scenarios in the right order gives a correct response. Clearly, this question can be worked out using standard probability arithmetic, but the question tests whether the respondent, in failing to use such a rational approach, is swayed by one aspect of the data. In the three scenarios presented the one with the least probability has the highest proportion of bricks in the stack that are the same colour as the one being drawn. Therefore, if the subject opts not to perform the mental arithmetic the question is whether they will choose on the basis of the proportions in the stack. To do so would be an error based around the conjunction of the probability and the proportion of coloured bricks.

The results for this question are similar to those for the ‘insufficient anchor adjustment’. The mean is negative at -0.9788 and the 95% confidence limits are -1.2886 and -0.6689. The standard deviation is 0.8738. Again, the sample shows systematic errors of judgement. Here the problem is that decisions are being made upon data on the basis of an assumed association between that data and the right answer. Where that association does not hold true errors are being made.

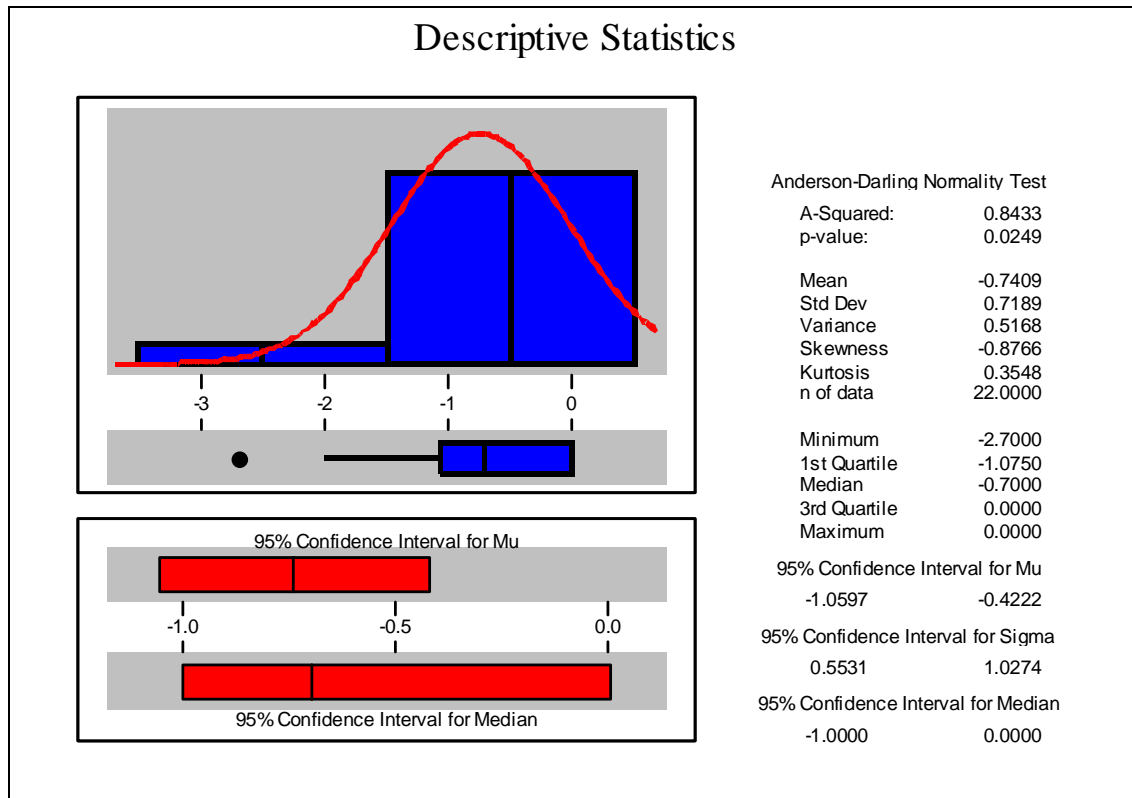


Figure 3: Distribution of responses for ‘overconfidence’

The question on ‘overconfidence’ requires the respondent to estimate five rates taken from a price book current in 1992. Once the rates have been estimated the subject is asked to give an upper and a lower limit such that they would have 95% confidence that the correct answer would lie between them. In each question there were five rates and provided that four out of the five were correct this would be taken as a correct response. Any other permutation was evidence of an error.

The distribution for ‘overconfidence’ is more negative than the previous two sub-types of error in that there were no correct responses. The mean was -0.7409 and 95% confidence limits being -1.0597 and -0.422. The standard deviation was 0.7189. There is clear evidence of systematic error. This is a potentially worrying finding in that it would indicate a lack of objectivity about potential errors in estimating which could lead to misinformed decision-making.

CONCLUSIONS

The previous discussion identified two null hypothesis scenarios - random distribution and positive skewing. The analysis of the results shows that neither of these holds true and therefore, they can be disregarded. The results show clear evidence of systematic errors of an anchoring and adjustment type. These errors have implications for the quality of strategic cost advice given to clients and, therefore, the quality of decisions made by clients when considering construction projects.

Errors of insufficient anchor adjustment will result in advice being skewed away from an objective assessment of cost and towards existing known data that is, or may be, irrelevant to the assessment. In the conjunctive and disjunctive error situation the consultant is demonstrating a reliance upon the link between two events or facts that in reality is not true. Introducing a logic flaw into the process of formulating advice

will lead to inaccurate calculations and advice that may be significantly in error. With the overconfidence error the practitioners are demonstrating a level of confidence in their estimates that is inconsistent with the facts. This error will lead to incorrect indications of the probability of the accuracy of advice.

The following are criticisms of the research;

- the sample size is small,
- the sample was a convenient sample and is not representative of the population of experts,
- the type of question used for 'overconfidence' poses a problem in that it would appear that a significant number of subjects had difficulty responding or considered the question too taxing and therefore did not attempt it. This had been recognised in the previous study and adjustments were made to reduce the difficulty attached to the question. Whilst the responses are better in the current study there is still room for improvement.

The recommendation for future research is that the new context based propensity for error test should be extended to include the other error types identified by Fortune and Lees (1997).

REFERENCES

- Bazerman, H. (1993) *Judgement in managerial decision making*. 3ed., New York; Wiley.
- Beach, L.R., Christensen-Szalanski, J. and Barnes, V. (1987) Assessing human judgement: has it been done, can it be done, should it be done? in *Judgemental Forecasting*, ed G Wright and P Ayton, John Wiley & Sons, p.49-62.
- Birnie, J. (1995) The Possible Effects of Human Bias in Construction Cost Prediction, Proceedings of 11th Annual ARCOM Conference, University of York, Sept., p.359 - 366
- Fortune, C.J. and Lees, M.A. (1996) The relative performance of new and traditional cost models in strategic advice for clients, *RICS Research Paper Series*, 2(2).
- Fortune, C.J. and Lees, M.A. (1997) Cognitive errors of judgement, learning style characteristics and clients' early cost advisors. *Procs 13th Annual ARCOM Conference*, Kings College Cambridge, Sept. 556-566
- Fortune, C.J. and Lees, M.A. (1998) Strategic cost modelling for construction projects - the role of judgement. *Engineering, Architectural and Construction Management* (in press)
- Kahneman, N., Slovic, P., and Tversky, A. (1985) *Judgement under uncertainty: heuristics and biases*. London; Cambridge University Press
- Mak, S. and Raftery, J.J. (1992) Risk attitude and systematic bias in estimating and forecasting. *Construction Management and Economics*, 10, 303-320
- Raftery, J.J. (1995) Forecasting and decision making in the presence of risk and uncertainty, Keynote Address to the *CIB W55 International Symposium on Economic Evaluation and the Built Environment*, Lisbon, 6, 1-11
- Tversky, A. and Kahneman, D. (1974) Judgement Under Uncertainty, Heuristics and Biases, *Science* 185, 1124-1131.