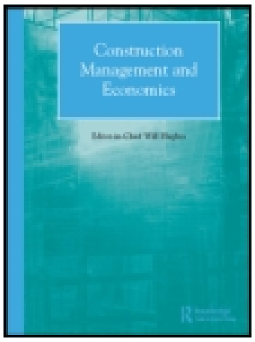

Construction Management and Economics

ISSN: 0144-6193 (Print) 1466-433X (Online) Journal homepage: www.tandfonline.com/journals/rcme20

Risk attitude and systematic bias in estimating and forecasting

Stephen Mak & John Raftery

To cite this article: Stephen Mak & John Raftery (1992) Risk attitude and systematic bias in estimating and forecasting, Construction Management and Economics, 10:4, 303-320, DOI: 10.1080/01446199200000028

To link to this article: https://doi.org/10.1080/01446199200000028

Published online: 24 May 2006.

Construction Management and Economics, 1992, **10**, 303–320

Risk attitude and systematic bias in estimating and forecasting

STEPHEN MAK and JOHN RAFTERY
School of Land and Construction Management, Thames Polytechnic, Oakfield Lane, Dartford, Kent DA1 2SZ, UK

This paper describes findings from the first phase of a research project investigating the modelling of risk and uncertainty in construction estimating and forecasting. The objectives of the project are to examine current methods and, experimentally, to explore techniques which offer the potential for the production of improved information from estimates and forecasts. Some of the causes of error and bias in both traditional deterministic and in probabilistic estimating and forecasting are described. The majority of the research in cognitive psychology which has led to the common assumption of errors and biases has been carried out with lay people thinking intuitively about problems. The present research is an attempt to test these biases in a domain-specific, non-intuitive context with individuals trained in that domain. The authors hypothesize that, if professional training has any value then they should find less evidence of bias than is the case in the general literature. Empirical results are reported and discussed. This work finds empirical support for only some of the biases commonly assumed to exist.

Keywords: Risk, uncertainty, estimating, forecasting, heuristic, bias, fallacy.

Introduction

This paper describes findings from a research project investigating the modelling of risk and uncertainty in construction estimating and forecasting. The objectives of the project are to examine current methods and, experimentally, to explore techniques which offer the potential for the production of improved information from estimates and forecasts. Some causes of error and bias in both traditional deterministic and in probabilistic estimating and forecasting are described. The majority of the research in cognitive psychology which has led to the common assumption of errors and biases has been carried out with lay people thinking intuitively about problems. This research is an attempt to test these biases in a domain-specific context with individuals trained in that domain. The authors hypothesize that, if professional training has any value then they should find less evidence of bias than is the case in the general literature. Empirical results are reported from an experimental study of biases in construction estimating. This work finds empirical support for only some of the biases commonly assumed to exist. The implications of these results are discussed.

Two further points need to be raised in this introduction. First, it should be noted that there is not a consensus in the literature on bias. The commonly held assumptions of

0144–6193

systematic error and bias which flow directly from a large body of the literature have been criticized by, among others Cohen (1981, 1986) and Hogarth (1981). That said, the study described here seeks to find evidence of these biases in a fairly realistic estimating exercise. Second, although the study described here is not about bidding, it is worth noticing that much of the literature on bidding takes as axiomatic the existence of unbiased bids (Flanagan and Norman, 1985). Empirical studies which gradually uncover the existence of construction estimating biases may result in the need to reconsider some aspects of bidding theory.

Causes of bias in reporting from cost models

Skitmore *et al.* (1990: 159) define bias in terms of the relationship between the value of the contract price forecast and the value of the lowest tender. They have devised three measures of bias; namely, raw differece, percentage difference and log difference. These are dependent variables and are affected by such factors as the target contract, the forecasting technique used, the information used, feedback, and the forecaster in the process of forecasting.

How and why do errors and biases creep into estimates and forecasts in building economics? There are at least two groups of reasons. One stems from common rules of thumb (or heuristics) and biases in the cognitive processes of human beings making judgements and forecasts in any situation. These will be explored below. Here we will reflect on a second source of error and bias which stems directly from the tendency, in building economics, to make unrealistic simplifying assumptions. In other words, the tendency to assume away real-world uncertainty and to assume that estimates and forecasts are exact and certain single points. Let us being by considering the latter.

Although subjects such as risk anlaysis and utility theory have been taught in business schools for more than a quarter of a century, they have only become embedded in building economics syllabuses in the past decade. Traditionally, in construction and property development, estimates and forecasts of, for example, time, costs, prices, rents and returns have been given as single points. There are, sometimes, very good reasons for this. Single point figures are often needed as 'budgets' for purposes of planning and control. Some projects need to be managed such that they are completed by a specified date. For example, exhibition, buildings or say, a harbour to facilitate materials imports for some large and vital infrastructure project. Other projects need to be managed – even to the extent of downsizing, redesign or reduced specification – such that the out-turn cost is equal to or less than the allowed budget as there may be no possibility of further resources. In such cases single point estimates and forecasts are indeed necessary for the purposes of day to day planning and control. Nevertheless, it remains the case that these single points are *forecasts* and targets not *certainties*. For most purposes single point forecasts conceal more then they reveal. Thus, there are fertile conditions for error when an estimate/forecast is producted by an expert or consultant and then reported to a director, partner or some other senior decision-maker. All that said, two important sources of bias in single point forecasts will be explored here. First the bias that results from conflict between, on the one hand, the need for better forecasts and, on the other, the method of managerial control using rewards and punishments for over- and under-estimation. Second, the bias and mis-interpretation which result from the differences in personal risk attitudes. These sources of bias have been long understood in business schools. Let us explore then in turn. The approach here is an adaptation to construction of work reported by Woods (1966) in *The Harvard Business Review*.

Conflicting use of forecasts

In many organizations there are reporting systems which involve managers reporting back, with explanations, when projects exceed budgets, or when returns fail to meet the initial target, by some arbitrary figure, in many cases 10%. Where a project comes in under budget, or when returns exceed forecasts, by more than the same amount the explanations, although necessary, are not examined in quite the same detail. There are not the same negative implications for the personnel who produced the model or the forecast. When an organization applies negative sanctions to managers who exceed budgets and, say, takes no action of praise or reward when projects come in on or under budget, then there will be a natural tendency to over-estimate forecasts of project costs and to under-estimate forecasts of project returns or rental. In this case, the figure which gets reported is not necessarily the most likely, it may be a very conservative figure, i.e. for project costs rather high, and for project returns a rather low figure which is quite likely to be exceeded. In either case, this type of reporting will lead, over time, to a misallocation of resources. Many projects which have reasonable likelihoods of achieving healthy returns will not get support as they do not contain a sufficient buffer to ensure* the safety of the forecaster in the short term. Many projects which have a better than even chance of coming in on or under budget will not get support because they cannot be estimated safely enough to protect the forecaster from recrimination in the short term.‡

Figures 1 and 2 illustrate, respectively, project cost and internal rate of return forecasts, showing in each case the areas within which conservative and risk-seeking forecasts lie. A conservative and self-preserving estimate of project cost will lie in the upper range of possible outcomes. We may regard as risk-seeking, any estimate from the lower end of the range with a probability of being exceeded of at least 0.5. Conversely, when we consider forecasts of project return, the conservative forecaster will draw a forecast from the lower end of the range. This is a very safe forecast. It will be relatively unlikely that the project will disappoint its sponsors by producing no or a very small return. Our choice of 0.5 as the cut-off point is purely arbitrary and for the purposes of discussion. In reality a client may choose a different cut-off dependent on it's own corporate risk attitude.

Differences in personal risk attitude

Biases and mis-interpretation occur even in situations where the output from a cost model is reported in ways which purport to take account of risk exposure. A consultant may indicate that there is a 'good' chance that the completed project in a specific location will achieve a rent of, say, £70 per square foot. Speaking *ex ante*, probably before the site has even been purchased, the consultant knows that the project will not be complete and available for at least two years. At which time the state of the property market could be the same, better or worse than today. Similarly, an advisor may state that there is a 'reasonable' chance that a project can be completed for less than £40 million. What do these statements actually mean?

* Or perhaps, effectively, *insure*.

‡ Assuming that there are a large number of projects and that they are of similar size and profitability and that they generate similar levels of absolute profit, i.e. no one project exerts excessive influence on the firm or practice. Then it is sufficient for six in ten of the projects to come in on budget. However, if the forecaster is subject to the type of controls mentioned here then four 'wrong' forecasts in ten may be a few too many for the forecaster's own career progression.

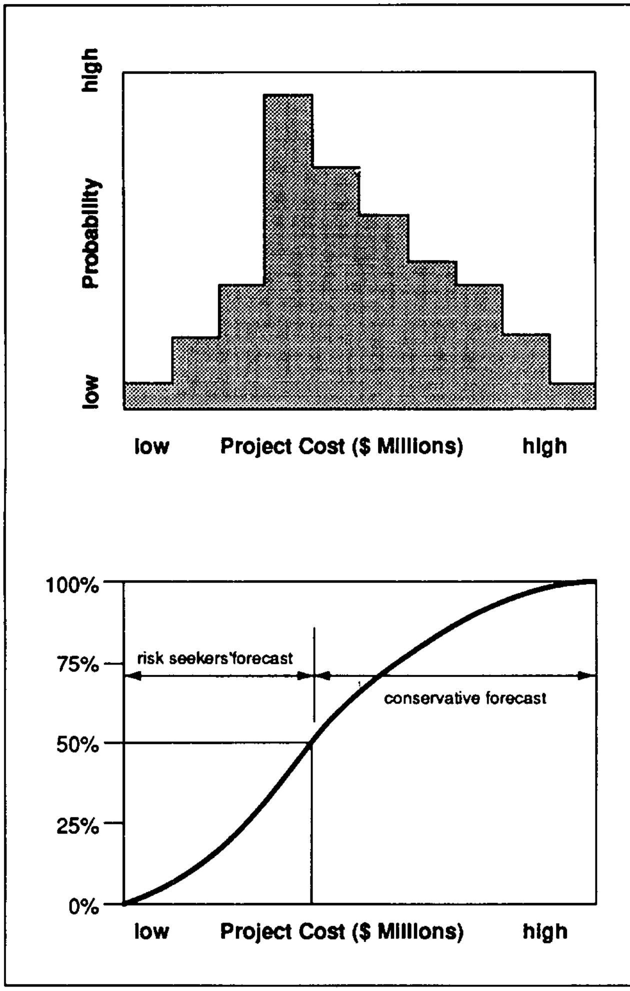

Fig. 1. Risk attitude in forecast of costs.

The language, in itself, seems reasonably clear. However, is a 'good' chance a nine in ten, an eight in ten or a six in ten? Is a 'reasonable' chance an eight in ten, a seven in ten or a five in ten chance? These differences could be very significant to a decision maker choosing between projects or between different approaches to the same project. What is the decision maker to do? It would seem rational for the decision maker to try to deduce what influences were acting upon the forecasters and to make adjustments for this before making decisions based on this information. Thus, 'this consultant is looking for an unblemished record leading to further work. Therefore this forecast is likely to be conservative so I shall add 10% to this estimate of project return.' This adjustment is, of course, arbitrary. The same sort of arguments could be applied to multiple single point forecasts such as those employing the three points, most likely, optimistic and pessimistic. The precise point at which the optimistic and pessimistic values are located is a function of the risk attitude of the forecaster. As such it is unlikely to be identical to the risk attitude of the decision maker. Therefore it may lead to

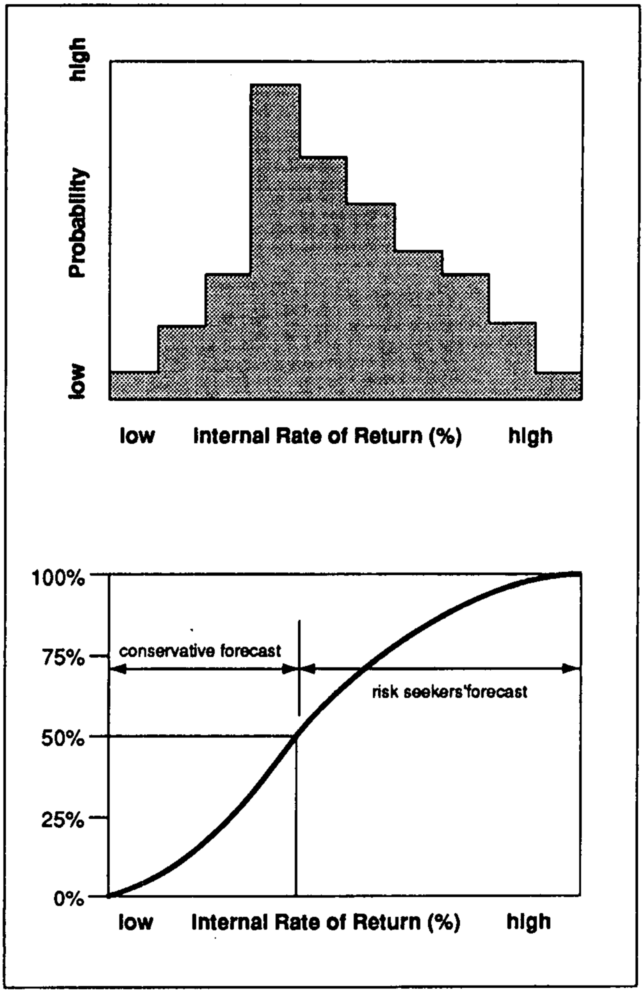

Fig. 2. Risk attitude in forecast of return.

biased decisions. Even if the decision maker decides to adjust for the bias of the forecaster the nature of the adjustment is likely to be quite arbitrary and thus may also result in bias. Let us consider an example of what can happen in practice.

What the estimator said . . .

Assume that an estimator is just completing work on a bid for a large overseas project. The estimator has to report a net cost estimate to the managing director who will make the mark-up decision. The estimating team get together on the day before the bid is due to be submitted and decide that their best estimate of the net cost is $72 million. This figure includes tangible and intangible costs, head office overheads and an allowance for the cost of recovering finance charges. It includes no profit normal or otherwise. This has been arrived at by breaking the project down as indicated in Table 1.

Table 1. Deterministic estimate of net project cost

Materials	\$25 M
Labour	10 M
Subcontractors	20 M
Overheads and finance	12 M
Supervision on site	5 M
Total	\$72 M

What the manager thought the estimator said . . .

The manager receives this figure together with the background briefing on the project from the estimators and the planning department. The decision on mark-up is a familiar problem to the manager who is accustomed to taking calculated risks in order to secure work at favourable rates for the firm. The manager knows that the estimate is a forecast of the outturn cost should the firm win the project. Thus, for the purposes of calculation, it would be rational to assume that the estimate is the most likely figure drawn from a distribution which manifests some skewness at the upper end of the range. In estimates of costs we expect a right skew. In other words, if things go very well then the costs may be at the lower end of the range. On the other hand, if things go badly they can go very badly so that the distribution is unlikely to be symmetrical. This is illustrated in Fig. 3. The mental cost model of the project held in the mind of the manager is that described by Fig. 3. The characteristics of this model are that the most likely outturn cost is \$72 million. The optimistic outcome is a project net cost of \$65 million, and the pessimistic outcome shows a cost of \$86 million.

What the estimator thought but did not say . . .

In a number of risk analysis workshops carried out on an earlier project (Raftery, 1985) it was possible to analyse in detail, deterministic project cost estimates. This exercise involved meeting the estimating teams towards the end of the bid period when they had a quite firm deterministic idea of the net cost of the project. In a 2 h session it was possible to carry out a detailed simulation of the project to produce a frequency histogram of possible outcomes and, from this, a cumulative probability graph. The experience of these analyses was that, when the simulation results were obtained it was usually found that the deterministic figure given earlier had been extremely conservative, with a probability of being exceeded often in the region of 0.1. This is not surprising given the long history of empirical studies demonstrating that business people have a tendency to be markedly risk averse. Thus, for the purpose of discussion, using that datum, we may construct what our generalized estimator thought but did not say. This is illustrated in Fig. 4 where we see that the \$72 million is a conservative estimate drawn from a distribution with lower and upper bounds of \$59 and \$79 million respectively.

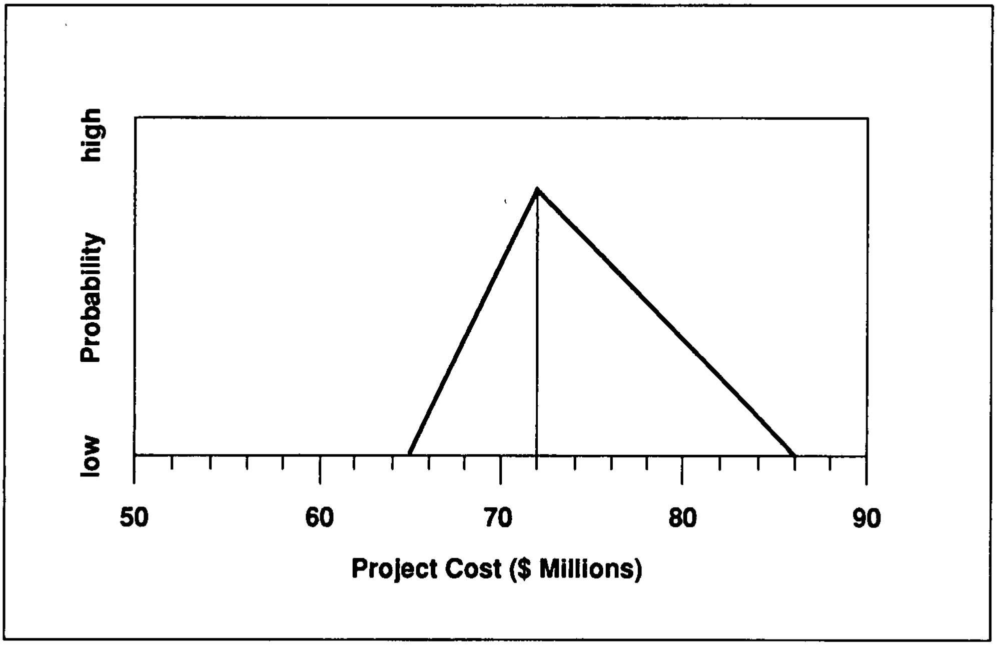

Fig. 3. Manager's perception of the single point estimate.

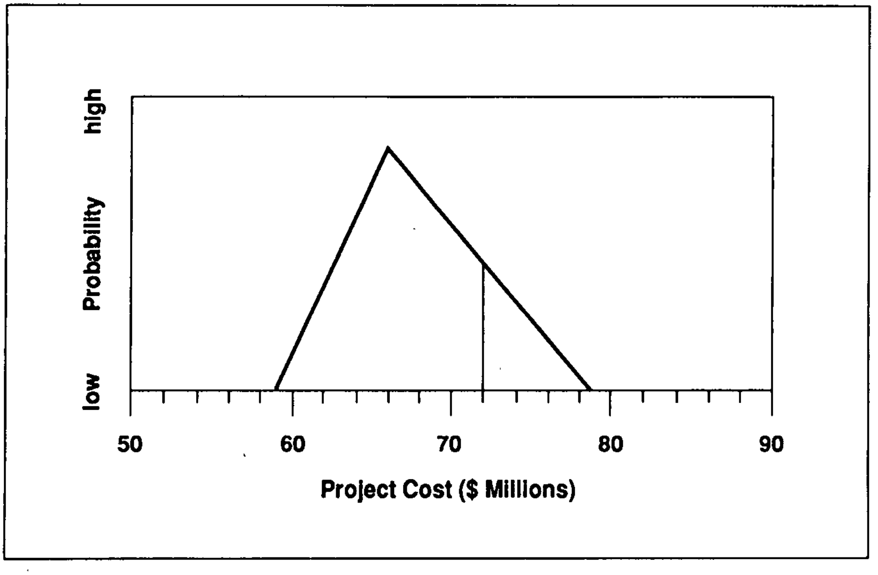

Fig. 4. Estimator's perception of project cost.

The amount of bais introduced is illustrated by combining Figs 3 and 4 into Fig. 5. This is a generalized abstracted case but it illustrates well the potential for disagreement when people implicitly adopt conservative forecasting as a method of coping with certain types of managerial control of professional work. The solution to this problem is abstractly obvious although possibly difficult to implement. Forecasters should adopt methods which explicitly deal with both risk and exposure and risk attitude.

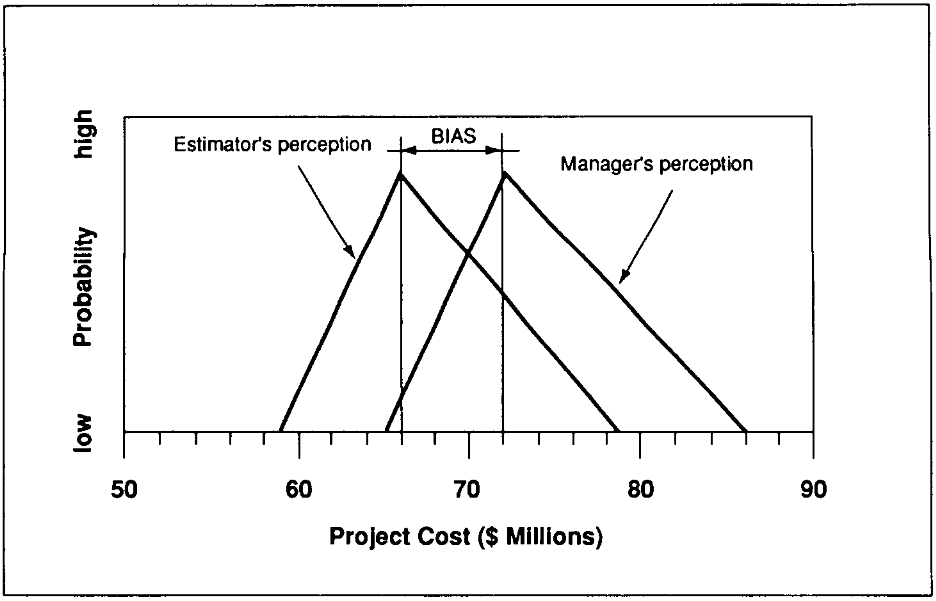

Fig. 5. Forecasting bias.

Systematic biases

There is a large psychological literature on the subject of judgemental bias and inconsistency. This has been widely reported elsewhere both in the field of decision science and, increasingly, in the field of construction economics and management (see for example Skitmore *et al.*, 1990; Birnie and Yates, 1991: 171–73). Those reports will not be repeated here. For a recent, detailed review and discussion readers are referred to Evans (1990). Here we only include sufficient descriptive material to place our empirical study, reported below, in its proper context.

The work of Daniel Tversky and Amos Kahneman (Tversky and Kahneman, 1974) has had a huge impact in many fields of work including military planning, psychology and decision science to name but three. Their work has lead to various controversies and there is some dissent from their findings (Cohen, 1979). There is however a large body of agreement on three common rules of thumb used by people making judgements. These three heuristics are known as; representativeness, availability, and anchoring and adjustment. Each of these three leads to, so called, fallacies. Let us, briefly, explore each in turn. The representativeness heuristic is the process of evaluating the probability of an uncertain event, or a sample, by the degree to which it (i) is similar in essential properties to its parent population; and (ii) reflects the salient features of the process by which it is generated. It is in essence an assessment of the degree of correspondence between a sample and a population, an instance and a category, an act and an actor or, more generally, between an outcome and a model (Tversky and Kahneman, 1983). Since information is commonly stored and processed in relation to mental models, such as prototypes and schemata, it is natural and economical for the probability of an event to be evaluated by the degree to which that event is representative of an appropriate mental model (Smith and Medin, 1981). The representativeness heuristic leads to fallacies about base rates and sample size.

Availability is a heuristic which makes use of retrievability of instances. Some event with which one has has personal contact looms more likely than it would statistically. For example, if one of us has had personal contact with an airplane accident, a relatively rare event, then we are highly likely to overestimate the likelihood of future similar mishaps.

Adjustment refers to the cases when people make estimates by starting from an initial value and adjusting it to yield the final answer. Adjustment from an anchor is usually employed in numerical prediction when a relevant value is available. The initial value may be suggested by the formulation of the problem, or it may be the result of a partial computation. Adjustments are typically insufficient in both cases. Anchoring refers to the phenomenon where final results are usually biased toward the initial values. Anchoring occurs not only when the starting point is given, but also when the assessor has to base his estimate on the result of incomplete computation.

Unresolved problems

It is known that people use heuristics (rule of thumb) in decision making. This seems sensible. It has been shown – although not without dissent – that these heuristics may lead to 'systematic and predictable' errors (Tversky and Kahneman, 1974: 1131). This project is concerned with empirically testing for the extent of these biases in a specific domain with domain specific experts who are, in the main, thinking non-intuitively.* This has been done before in other fields such as investment analysis (Slovic, 1972). As Birnie and Yates (1991: 171) put it recently in this journal, 'it is unlikely that construction estimators and forecasters are immune from these tendencies'. On the other hand, we may hypothesize that if estimating skill, training and expert judgement are worth anything then this may imply a pattern of bias different to the general population. This question is the subject of the experiment reported in the next section.

An experimental study of systematic bias in construction estimating

A study was carried out in April 1991 with a group of 62 final year undergraduate quantity surveying students at Thames Polytechnic. Many of them had reported several years'

* These questions *were* explored by Skitmore *et al.* (1990) in an empirical study of construction estimating accuracy. In this case, though, the questions on this type of bias tended to be outside the field of construction. Therefore, Skitmore *et al.*'s study, while constituting a useful testing of Tversky and Kahneman's findings, did not extend into the non-intuitive, domain specific field of the work being reported here.

industrial experience in the construction industry. These students were divided randomly into two groups and each given a set of estimating tasks in the form of questionnaire. There were 32 subjects in the control group and 30 in the experimental group. They were given 1 week to complete the tasks and were not told which group they were in. SPSS/PC was used to analyse the data.

The questionnaire

The are two sections in the questionnaire. Only Section A of it is reported here. It comprises three questions testing the three heuristics (representativeness, availability and anchoring and adjustment) which in turn lead to the various fallacies. The questions are designed in the construction context and involve the task of estimating. Both sets of questionnaires contained an appendix with extracts of cost information from various sources sufficient for all questions to be answered. The difference between the questions given to the two groups is that the control group was given a plain question while the experimental group was given the same question plus an additional piece of information, the purpose of which is to test whether the subjects can distinguish between relevant and irrelevant information in the estimating process and whether they display 'systematic and predictable' errors (Tversky and Kahneman, 1974: 1131). The three questions are listed in Table 2 below. The sentences in italics are additional information given to the experimental group of subjects. Table 3 shows the raw data collected from the subjects. Columns 2, 4 and 6 are the estimates given by the subjects while columns 3, 5 and 7 show whether the subject actually adopted the heuristic concerned. Since there are 30 subjects in the experimental group, the 32 subjects in the control group have a 'n.a.' remark in columns 3, 5 and 7.

Analysis of results

Question 1. The first question is about anchoring and adjustment. The subjects were divided into two groups. Group A, the control group, was given the question without the last paragraph; while Group B, the experimental group, was given an extra sentence describing the estimated cost of a finished project. This exta information was designed such that if the subjects were anchored to this given cost then they would tend to give an answer biased towards that cost (the anchoring effect). We might expect that a professional with some knowledge of this particular field (in this case a quantity surveyor) would be less likely than the general population to display an anchoring bias.

The means of the estimated costs of the two groups are calculated and compared. The t-statistic is used to compare the difference between the two means. The null hypothesis is that there is no difference between the estimates of the two groups. And the alternative hypothesis is that there is a difference in estimated cost between the two groups.

The results in Fig. 6 shows that with a degree of freedom of 60, the t-value of the pooled variance estimate is 1.96. This value is smaller than that corresponding to a t-value at the 5% level. This means that the null hypothesis is not rejected at 5% significance level. In turn this means there is no statistically significant difference in the estimated costs between the two groups. The anchor, which says 'the client has heard reports that a recently finished supermarket cost £1.5 million', did not have much effect on the subjects' cost estimate. Had it not been the case, then the experimental group's cost estimate would have been substantially lower, in this case, in order to fit the anchor of £1.5 million.

Table 2. The questionnaire in the experiment

1. A client approaches you and wants to build a new supermarket in Dartford Town Centre. Prepare an estimate for the construction costs for this proposed supermarket, with a gross floor area 10 000 m^2, 100 parking spaces and associated works.

   The client has heard reports that a recently finished supermarket for one of his competitors cost £1.5 million.

2. You are an estimator employed by ACME Contractors Ltd. You are asked to prepare a lump sum estimate for the pile caps as shown in Fig. 1. This estimate is used to compare the tender submissions of several sub-contractors. This pile of cap is to be let as a separate subcontract to one of the sub-contractors.

   You have checked a previous BQ of your company and found the following information:

(1) Reinforced concrete grade 25 N/mm² in formwork	*£30.00/m³*
(2) High tensile reinforcement bars cut, bent and bundled 16–32 mm diameter	*£350.00/tonne*
(3) Formwork to foundation and beds height > 1m	*£20.00/m²*

   bearing in mind that, due to your company's pricing policy, these rates are not necessarily true costs of the respective items.

3. A client approaches you and wants you to give an approximate estimate (total cost based on cost/m^2) for the construction of a factory building for his new light-weight production line. The area of the factory will be about 12 000 m^2, located in rural West Yorkshire.

   The construction will be steel columns, steel truss roof with aluminium roof coverings; pad foundations and aluminium cladding. The soil type is mainly clay. Some landscaping work will also be required but that is to be considered as a separate contract later. There will need to be builders work allowances for 24-h security guard post and CCTV installation. The installations themselves need not be considered at this stage.

Note: A separate set of information sheets, comprising extracts from BCIS, Wessex Data Bank and Spon's Price Book, were provided to the subjects.

Nonetheless, the mean cost estimate of the experimental group did show a lower figure compared to that of the control group, £6 049 867 *vs* £7 710 844 ($\approx$27.5%). To an extent this reflects that the subjects in the experimental group exterted more caution in the estimate and tended to select a relatively lower figure, whenever there was a range of figures to choose from, so that the estimate produced did not deviate 'too much' from the clients' information.

A qualitative analysis of the questionnaires was undertaken. The calculations in each of the experimental group's questionnaire were analysed to see if the anchor was really used by the subjects in arriving at the final cost estimate. A yes–no count was done and this is shown below in Fig. 7.

Figure 7 shows that all the 30 subjects in the experimental group did not use the anchor in their calculations. Indeed, in scrutinizing the calculations, some subjects (5 out of 30 of 16.7%) did comment on the anchor being of no use for them in making the estimate. Subjects mentioned the absence of information on location, size and time factors.

Question 2. The second question is about availability. Similarly, group A gets the plain question while group B gets and extra piece of information which was so designed that it

Table 3. Raw data

Sub No.	*Q1* (£'000)	Use *H*	*Q2* (£'000)	Use *H*	*Q3* (£'000)	Use *H*	Sub No.	*Q1* (£'000)	Use *H*	*Q2* (£'000)	Use *H*	*Q3* (£'000)	Use *H*
1	5016	no	16261	no	3315	no	32	13914	n.a.	3340	n.a.	3611	n.a.
2	12540	no	7000	yes	3157	no	33	14000	n.a.	3523	n.a.	3600	n.a.
3	4600	no	7000	yes	3300	yes	34	8000	n.a.	1712	n.a.	3023	n.a.
4	5016	no	10630	no	2810	no	35	10350	n.a.	30982	n.a.	3263	n.a.
5	5109	no	6100	yes	2520	no	36	5016	n.a.	24160	n.a.	3435	n.a.
6	4560	no	23372	yes	3588	no	37	5640	n.a.	23818	n.a.	3818	n.a.
7	6620	no	26186	yes	3812	yes	38	4950	n.a.	26000	n.a.	3800	n.a.
8	11500	no	15000	no	3280	yes	39	14100	n.a.	17832	n.a.	3553	n.a.
9	3950	no	14343	yes	3507	yes	40	13420	n.a.	25500	n.a.	4800	n.a.
10	5313	no	23129	no	3270	no	41	5876	n.a.	16060	n.a.	3156	n.a.
11	4888	no	6954	yes	3315	no	42	4560	n.a.	16800	n.a.	3020	n.a.
12	4560	no	14448	no	2244	no	43	6946	n.a.	18871	n.a.	4021	n.a.
13	4460	no	7447	yes	3696	no	44	4814	n.a.	18475	n.a.	3834	n.a.
14	4650	no	28200	yes	4150	yes	45	5088	n.a.	19439	n.a.	4322	n.a.
15	5820	no	5608	yes	3252	yes	46	6578	n.a.	20878	n.a.	3924	n.a.
16	4476	no	6322	yes	3653	no	47	7094	n.a.	18871	n.a.	4110	n.a.
17	11637	no	16289	yes	1310	no	48	6704	n.a.	18550	n.a.	4876	n.a.
18	4814	no	20885	no	4028	no	49	19514	n.a.	16100	n.a.	7121	n.a.
19	4560	no	16993	yes	4200	no	50	4845	n.a.	21800	n.a.	3913	n.a.
20	6741	no	19300	no	3788	no	51	5865	n.a.	24000	n.a.	2767	n.a.
21	7500	no	20800	no	4600	yes	52	5918	n.a.	19440	n.a.	3400	n.a.
22	3778	no	11212	yes	3157	no	53	4546	n.a.	17671	n.a.	3912	n.a.
23	5250	no	17700	no	3156	no	54	5368	n.a.	11703	n.a.	3020	n.a.
24	5600	no	23812	yes	3758	no	55	5890	n.a.	21800	n.a.	3306	n.a.
25	12215	no	12557	yes	3653	no	56	17271	n.a.	14247	n.a.	3500	n.a.
26	5555	no	23932	yes	3105	yes	57	6320	n.a.	19900	n.a.	4118	n.a.
27	4658	no	9587	yes	3041	no	58	5545	n.a.	12900	n.a.	3323	n.a.
28	5700	no	18500	no	3000	no	59	4455	n.a.	20636	n.a.	4794	n.a.
29	4600	n.a.	1900	n.a.	3000	n.a.	60	4884	no	32552	yes	2766	no
30	6500	n.a.	9625	n.a.	4500	n.a.	61	5526	no	29191	yes	4169	yes
31	9360	n.a.	19025	n.a.	3611	n.a.	62	4660	n.a.	16056	n.a.	3080	n.a.

```
t-test for:  A1        Sec A Q1 (Anchoring) Estimated Cost (£'000)

                       Number                 Standard    Standard
                      of Cases      Mean     Deviation      Error

        Group 1          32      7740.8438    4072.853    719.986
        Group 2          30      6049.8667    2495.024    455.527

                      | Pooled Variance Estimate  | Separate Variance Estimate
                      |                           |
        F     2-Tail  |    t    Degrees of 2-Tail |    t     Degrees of   2-Tail
      Value   Prob.   |  Value   Freedom   Prob.  |  Value    Freedom     Prob.
                      |                           |
       2.66    .010   |   1.96     60       .055  |   1.98     51.90       .052
```

Fig. 6. *t*-test of the anchoring effect in cost estimate between the two groups.

- - - - - Chi-square Test

A1E USE ANCHOR

	Category	Cases Observed	Expected	Residual
NO	2	30	30.00	.00
		--		
	Total	30		

Only one cell generated. Test abandoned.

Fig. 7. χ^2 test of the use of anchor in calculations for the experimental group.

should be discarded. The appendix provided with the questionnaire gave an easy access to cost information which should have been adopted by the subjects had they relied only on available information without resorting either to their own expert knowledge or to up-to-date data banks. The null hypothesis here is that the cost estimates done by both groups are not different. The alternative hypothesis is that the cost estimates are different between the two groups. The output is shown in Fig. 8.

t-test for: A2 Sec A Q2 (Availability) Estimated Cost (£'000)

	Number of Cases	Mean	Standard Deviation	Standard Error
Group 1	32	17237.9375	7058.695	1247.813
Group 2	30	16377.0000	7703.636	1406.485

		Pooled Variance Estimate			Separate Variance Estimate		
F Value	2-Tail Prob.	t Value	Degrees of Freedom	2-Tail Prob.	t Value	Degrees of Freedom	2-Tail Prob.
1.19	.632	.46	60	.648	.46	58.64	.649

Fig. 8. t-test of the availability effect in cost estimate between the two groups.

The calculated t-value is 0.46 with 60 degrees of freedom. This value is much smaller than the t-value at the 5% level and therefore renders the null hypothesis not to be rejected. This implies that there is no statistically significant difference in the cost estimates between the two groups and infers that the available information did not have great effect on the cost estimate process of the subjects.

A yes–no count similar to question 1 above was also done. Those subjects in the experimental group who used the available information were counted and the results, together with a χ^2 test is displayed in Fig. 9.

The χ^2 test result does not show a significant difference within the experimental group between those who used the 'available' information and those who did not. This has implications for the results of the t-test because it could be implied later that the majority of the subjects within the experimental group did not use the available information. In fact, 20

- - - - - Chi-square Test

A2E USE AVAILABLE INFORMATION

	Category	Cases Observed	Expected	Residual
YES	1	20	15.00	5.00
NO	2	10	15.00	-5.00
	Total	30		

Chi-Square	D.F.	Significance
3.333	1	.068

Fig. 9. χ^2 test of the use of available information in calculations for the experimental group.

out of 30 (66.7%) of the subjects did use the available information provided. One explanation for this could be that the subjects did not know how to interpret the information in the data bank. Accompanying the questionnaire was statistical data extracted from the Wessex data book and the rates for reinforced pile caps included excavation, formwork, reinforcement and concrete, all as described in the preamble item which was also supplied to the subjects. However, many subjects assumed the rate for a pile cap from the data bank covered concrete only and adjusted for reinforcement, formwork and excavation by using the 'available' information provided. A number of subjects did build up the rate from the available information only without resorting to the data bank provided. A reason for this may be that 'available' information provided more itemized information that was available in the data bank. The aggregate effect of the subjects on the experimential group shows a lower mean cost estimate that that done by the control group, because the rates in the 'available' information were deliberately much lower than they should have been. In this case, however, the difference has not been shown to be significant. Had the 'available' information been skewed far enough from the mean, perhaps the two groups would have produced more significant difference in the *t*-statistic. On the other hand, if the 'available' information were too skewed from the mean, the subjects would have noticed the difference and avoided using it. This counter-balancing effect is worth investigating further. One can, though, notice the much higher standard deviation in the cost estimate in the experimental group (£7 703 *vs* £7 058).

Question 3. The third question is about representativeness. Group B was given some extra descriptions which are general descriptions of a typical project of its type. The intention was to test whether the subjects would be affected by this piece of information and adjust the base rate information accordingly. It was expected that the subjects (in the experimental group) would not be affected by the description of the project, because it is so general and typical; otherwise they would have committed the base rate fallacy, i.e. neglecting the base rate because of irrelevant information. The null hypothesis is that there is no difference between the estimated costs between the two groups while the alternative hypothesis is that there is a difference. Similarly, the *t*-statistic is used first and the results are shown in Fig. 10.

The *t*-value for the pooled variance estimate is 2.35 with 60 degrees of freedom. This is much larger that a *t*-value at the 5% level. This implies that the null hypothesis is not supported and there is evidence to show that the two means are different. It is also interesting t-test for: A3 Sec A Q3 (Representativeness) Estimated Cost (£'000)

	Number of Cases	Mean	Standard Deviation	Standard Error
Group 1	32	3797.8438	823.090	145.503
Group 2	30	3353.3333	646.846	118.097

		Pooled Variance Estimate			Separate Variance Estimate		
F Value	2-Tail Prob.	t Value	Degrees of Freedom	2-Tail Prob.	t Value	Degrees of Freedom	2-Tail Prob.
1.62	.196	2.35	60	.022	2.37	58.27	.021

Fig. 10. *t*-test of the representativeness effect in cost estimate between the two groups.

to note that the mean and standard deviation of the cost estimate by the experiment group is much smaller that that by the control group, the group which did not receive the 'representative' information. This means the subjects of the experimental group consistently estimated a lower cost that the subjects of the control group. This can be explained, though, by the fact that the subjects of the experimental group confined themselves to two sets of data from the data bank provided, in addition to the 'representative' information, which are the BCIS Average Building Price for Steel Frame Factories and Factories over 2000 m^2 Gross Floor Area. Also these two sets of data had the lowest mean values among the six sets of BCIS data provided. Obviously, the control group, having no indication of the type of construction and so on, tended to have a wilder choice.

A yes–no count was again done in an attempt to distinguish whether the subjects actually used or did not use the representative information. The statistics are shown in Fig. 11.

- - - - - Chi-square Test

A3E USE REPRESENTATIVE INFORMATION

	Category	Cases Observed	Expected	Residual
YES	1	9	15.00	-6.00
NO	2	21	15.00	6.00
	Total	30		

Chi-Square	D.F.	Significance
4.800	1	.028

Fig. 11. χ^2 test of the use of representativeness information in calculations for the experimental group.

Twenty-one out of 30 subjects (70%) in the experimental group did not use the representative information provided in their estimates. This is significant as shown by the χ^2 value of 4.8 which is smaller than a χ^2 value at the 5% level with 1 degree of freedom. The hypothesis that there is no difference between the adoption of representativeness information is not supported. This means that the subjects did not randomly adopt the representativeness information. In effect, the majority of subjects did discriminate between high and low quality data. As we are not assessing the accuracy of estimates here, it is more plausible to trust the χ^2 test results as we want to test whether the heuristics approaches, as suggested by the Tversky and Kahneman (1974), have been used. In this case, they were not used by the majority of subjects within the experimental group.

Conclusions and limitations of the study

In a domain specific, context sensitive situation, subjects from the domain tend not to be making intuitive estimates. This research is an attempt to find empirical support for the systematic biases commonly assumed to afflict estimators. The findings of this study demonstrate that in this situation, subjects tended to be less error-prone in making judgements than might have been expected from the literature on bias. This should be viewed, of course, in the light of the limitation of the work which are discussed below.

The comparison of the cost estimates of the two groups of subjects does not reflect clearly whether the subjects adopted the heuristics or not. A supplementary approach, which looks into how the subjects actually did the calculations and evaluates whether they used the heuristics, provides further results. In general, no statistically significant support was found for two of the three biases. Further analysis of the data showed that many subjects (20 out of 30 in the experimental group) did actually make use of the easily retrievable but incorrect, data supplied with question 2. This, it could be argued that the subjects did demonstrate the availability bias. However, the tendency to use the incorrect available data did not produce statistically significant differences between the results of either group. It could be argued that the absence of a significant difference is because the rogue information was not 'incorrect' enough, i.e. it was too close mathematically, to the price level in the data bank. This raises a general methodological problem in empirical work in this field. Since the subjects have some expertise in the domain they are likely to notice if information, although available, is incorrect. In this case they would not use it. In our study, perhaps the information was used because it was sufficiently close to reality not to be noticed as obviously incorrect. This leads, intuitively, to the conclusion that domain-specific experts should not demonstrate a generalized tendency to adopt the availability heuristic. Table 4 summarizes the results of this study. Several factors may have contributed to these findings. First, the analysis of the

Table 4. Summary of results of study

Question	Content	Heuristic	Supported
1	Supermarket	Anchoring and adjustment	NO
2	Pile caps	Availability	YES
3	Factory	Representativesness	NO

questions was not confined purely to comparison of the given answers. Comparison of given answers and comparison of the approaches used to arrive at the answers were carried out. This, to an extent, has eliminated the danger of other considerations, such as adding different percentages of professional fees or odd items which, though correct, may affect the outcomes, in the calculations, and may lend firmer support to the results. Secondly, the questions were set in an appropriate context so that the subjects were familiar with what was asked or expected. Thirdly, the subjects were not asked to complete the questionnaire in a laboratory environment in which they had to complete the questionnaire in a hurry. They were allowed to take the questionnaire home though it was suggested that they should not spend more than 1 h to finish it. Although this may introduce the danger of collaborative work, in general the outcome appears satisfactory insofar as we have no evidence to suggest collusion.

There are several limitations to this study. For example, small number of subjects and subjects not being highly experienced in the construction industry. Nevertheless, the results lend support to the authors' hypothesis that if professional training has any value then they should find less evidence of bias than is the case in the general literature. In the work reported here, subjects with some experience, but educated to degree level, did demonstrate that conclusions of previous findings by cognitive psychologists may have been too pessimistic.

The study has not considered the extent to which the 'extra' information, given to the experimental group, affected subjects' judgements. The skewness of the 'extra' information given to the subjects may affect subjects' choice of approach. If the extra piece of information is skewed far enough, perhaps the adoption of a heuristic would be apparent. However, one should not overlook the counter balancing effect, i.e. if the given information is too skewed, the subjects would be able to spot it and discard it in the calculations. It would be worthwhile investigating this effect in further studies.

This study may be seen as less important for the results than for raising questions about the methodology of this area of research. Much of the empirical work on heuristics and biases, on decision making under risk, and the study of risk attitudes through the construction of personal utility curves had been done through the use of laboratory experiments. Amongst the advantages of this laboratory approach are, detailed control by the investigators and the ability to isolate specific variables of interest. Frequently, their advantage is also their weakness. The results are often criticized due to the unrealistic nature of the experiments themselves. They may involve hypothetical outcomes or tiny rewards. The conditions of making judgements may be very dissimilar to the real world equivalent. The work reported here did not take place in a laboratory setting. The subjects had no knowledge of the literature on biases, they did have some training in the professional area of the estimating tasks involved. They were asked to bring their considered responses back within a week. This is probably a slightly longer timescale than would be the case in industry but it is certainly more like the real world that the laboratory equivalent.

However, in attempting to resolve one problem which besets most laboratory experiments in this field we may have introduced other problems We have carried out a study which is more akin to the real world than its laboratory equivalent. We cannot prove that consultation or collusion did not take place. Although, on the other hand, we have no evidence to suggest that they did. Let us, momentarily, assume the worst. What would be the effect on the results of this study if the subjects had colluded in some way.? *A priori* it would seem sensible to expect some sort of *Delphi* type consensual effect. This would have an ameliorating effect on the results and would lessen the effect of any bias which might otherwise have shown up. Hence if this study shows only weak evidence of bias it would be safe to assume that there would be even stronger evidence in a laboratory study. Conversely, the methodology of this study would tend to suggest that it is highly unlikely that we have found evidence of a bias which was not there.

Perhaps the best that can be said is that there was little significant support for the existence of 'severe and systematic' bias in this study. Thus, the authors' hypothesis remains, for the time being, unrefuted. Our interim conclusion must be that further studies are required focusing both on substance and methodology especially with experienced practitioners in the industry and subjects from different institutions, and perhaps different countries.

Acknowledgement

The authors wish to acknowlege helpful comments, criticisms and suggestions on an earlier draft of this paper from; the members of the Reading Construction Economics Seminar, Dr Martin Skitmore of the University of Salford and an anonymous referee.

References

Birine, J. and Yates, A. (1991). Cost prediction using decision/risk methodologies, *Construction Management and Economics*, **9**, 171–86.

Cohen, L.J. (1979). On the psychology of prediction: Whose is the fallacy? *Cognition*, **7**, 385–407.

Cohen, L.J. (1981). Can human irrationality be experimentally demonstrated? *The Behavioural and Brain Sciences*, **4**, 317–70.

Cohen, L.J. (1986). *The Dialogue of reason*, Clarendon Press, Oxford.

Evans, J.St.B.T. (1990). *Bias in Human Reasoning: Causes and consequences*, Erlbaum, Brighton.

Flanagan, R. and Norman, G. (1985) Sealed bid auctions: an application to the building industry, *Construction Management and Economics*, **3**, 145–61.

Hogarth, R.M. (1981). Beyond discrete biases: Functional and dysfunctional aspects of judgemental heuristics, *Psychological Bulletin*, **90** (2), 197–217.

Raftery, J.J. (1985). *Risk analysis in Construction Tendering: A Guide for Building Contractors*, Technical Research Centre of Finland, ESPOO.

Skitmore, M., Stradling, S., Tuohy, A., and Mkwezalamba, H. (1990). *The Accuracy of Construction Price Forecasts: A study of Quantity surveyors' Performance in Early Stage Estimating*, University of Salford.

Slovic, P. (1972). Psychological study of human judgement implications for investment decision making, *Journal of Finance*, **27**, 779–99.

Smith, E.E. and Medin, D.L. (1981). *Categories and Concepts*, Harvard University Press, Cambridge, Mass.

Tversky, A. and Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases, *Science*, **125**, 1124–31.

Tversky, A. and Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment, *Psychological Review*, **90**, 293–315.

Woods, D. H. (1966). Improving estimates that involve uncertainty, *The Harvard Business Review*, July–August, 91–98.