DEVELOPMENTS IN CONTRACT PRICE FORECASTING AND BIDDING TECHNIQUES

Chapter in

A Manual of Quantity Surveying Techniques

(ed. P S Brandon, Blackwell Scientific Publications)

by

Martin Skitmore

Department of Civil Engineering
The University of Salford
Salford M5 4WT, UK

Brendan Patchell

Bucknall Austin plc
5 Scotland Street
Birmingham B1 2RR, UK

June 1989

DEVELOPMENTS IN CONTRACT PRICE FORECASTING AND BIDDING TECHNIQUES

Martin Skitmore[1] and Brendan Patchell[2]

(A)

INTRODUCTION AND DESCRIPTION OF TECHNIQUES

The objective of this chapter is to describe and demonstrate the use of four types of design estimating systems that have been developed in recent years to the point of commercial use. These consist of:

* standard REGRESSION approach to item identification and pricing

* CPS simulation for evaluation of item interdependencies

* ELSIE, the expert system for 'front end' item and quantity generation, and

* a prototype BIDDING MODEL DEBIASER developed by the author for 'back end' estimate adjustment for bidder characteristics.

---

[1] Senior Lecturer, Dept of Civil Engineering, University of Salford, Salford M5 4WT, UK.

[2] Quantity Surveyor, Bucknall Austin plc, 5 Scotland Street, Birmingham B1 2RR, UK.

Before considering these systems however, some preliminary observations are provided concerning the type and nature of some currently available estimating techniques.

Designers' and contractors' approaches

The functional separation of design and construction has been reflected in the development of contract price forecasting techniques largely to meet the perceived needs of both these sectors of the construction industry. For designers, the need is to inform on the expenditure implications (to the client) of design decisions to help in achieving

financial targets, and
value for money

For contractors, the need is to inform on the income implications (to the contractor) should the contract be acquired. Although the nature of the competitive tendering process is such that both designers and contractors are essentially concerned with the same task - estimating the market price of the contract - the potential for contractors to access production cost information is a determining factor in the type and reliability of technique used.

Many of the methods used in practice contain a mixture of both market and production related approaches. It is common, for instance for quantity surveyors to build up rates for special items involving new materials or components from 'first principles' by finding out the manufacturer's price and making an allowance for installation and other associated costs involved in a similar manner to the contractors' estimator. It is also common for contractors' estimators to use some 'black book' method which is more intended to generate a competitive price than a genuine estimate of cost. The validity of combining the two approaches, although a popular debating point, is by no means established in academic circles. Clearly, the relationship between market prices and production costs is the determining factor, and research on this aspect has been inconclusive as yet.

It is generally accepted however that the contractors' resource based approach produces very much more reliable estimates than the designers' equivalent. Unfortunately, theoretical research in this area has until relatively recently concentrated entirely on the formulation and solution of a highly simplified mathematical construction of the contractors' bidding decision. Despite the apparent remoteness of bidding research from the real world of quantity surveying, there is an important connection with design estimating as, unlike construction estimating, bidding is directly concerned with market prices.

Strangely, bidding research has been quite opposite in many fundamental ways to estimating research. Firstly, bidding research is firmly founded in economic theory with very little empirical support whilst estimating research has no formal theoretical base at all! As a result, both research fields have run into serious logistical problems. Bidding researchers are now faced with an unmanageable number of theoretical extensions, most of which are not appropriate to real world data, and estimating researchers are now faced with an unmanageable amount of data with no theoretical basis for analysis. Secondly, bidding models are invariably probabilistic in nature whilst estimating has traditionally been treated as a deterministic matter. Perhaps the biggest emphasis in estimating research today is in the reliability of the techniques used, and statistical probability offers the greatest potential for modelling reliability. Clearly then the theories and techniques used in bidding have some relevance in estimating especially in the provision of a mathematical basis for the analysis and developments in the field.

Mathematical and typological features

Traditional estimating models can be generally represented in the form of

$$P = p_1+p_2+\ldots+p_N = q_1r_1+q_2r_2+\ \ldots\ +q_Nr_N \tag{1}$$

or, more succinctly

$$P = \sum_{i=1}^{N} p_i = \sum_{i=1}^{N} q_i r_i \qquad (2)$$

Where P is the total estimated price, p is the individual price, q the quantity of work and r the rate or value multiplier for items 1, 2, ... ,n respectively. So for a bill of quantities based estimate, q_1 q_2 etc would represent the quantities for items 1 2 etc in the bill, r_1 r_2, etc the rates attached to the respective items, and p_1 p_2 etc their individual products. For example, an item of say concrete in floors may have a quantity q of 10 m^3 and rate r of £60/m^3, giving an item price p of qr=£600, the sum of all such prices Σp giving the total estimated price P.

The differences that occur between traditional estimating methods are usually in the number and type of items and the derivation and degree of detail involved in estimating their respective q and r values.

(insert Table 1 here)

From Table 1 it can be seen that the number of items (N) may range from one single item (UNIT, FUNCTIONAL UNIT, EXPONENT, INTERPOLATION, FLOOR AREA, CUBE, and STOREY ENCLOSURE) to very many (BQ PRICING, NORMS, RESOURCE).

The UNIT method is used on all types of contracts and involving any comparable unit such as tonnes of steelwork or metres of pipeline

The FUNCTIONAL UNIT method being similar but restricted to use on buildings with units such as number of beds or number of pupils

The EXPONENT method, used on process plant contracts, involves taking the contract price of an existing contract and multiplying by the ratio of some relevant quantity, such as plant capacity, of a new to existing contract, raised to a power, r, determined by previous analysis

The INTERPOLATION method takes the price per square metre of gross floor area of two similar contracts, one more expensive and one cheaper, and interpolates between them for a new contract

The FLOOR AREA method is less specific in that a price per square metre floor area is derived *somehow*, depending on any information available

Similarly the CUBE method, involves the calculation of a quantity representing the cubic content of a building for the application of a price per cubic metre

Perhaps the most sophisticated single item method is the STOREY ENCLOSURE method, involving the aggregation of major building features such as floor, wall, basements and roof areas into a composite index known as a 'storey enclosure unit'

BQ PRICING and NORMS methods on the other hand involves the detailed quantification of a great number of items, usually prescribed by an external contract controlling institution such as the Joint Contracts Tribunal (in the UK) or Government (in Eastern Block countries)

Several methods fall between these two extremes using either a few items (GRAPHICAL, PARAMETRIC, FACTOR, and COMPARITIVE) or a reduced set of BQ PRICING/NORMS type items (APPROXIMATE QUANTITIES, ELEMENTAL).

The GRAPHICAL method, used on process plant contracts, involves plotting the quantities of each of a few items of interest against the contract sum of previous contracts, for a visual analysis of possibly non linear relationships

PARAMETRIC methods, also used on process plant contracts, adopt a multivariate approach in using a function of several process related items such as capacity, temperature, pressure in combination

FACTOR methods, again used for process plant contracts, involve pricing only a portion of the contract which is then multiplied by a factor derived from previous similar contracts. Versions of the FACTOR method include the Lang method, which uses a single factor; the Hand method, which uses different factors for different parts of the contract; and the Chiltern method, which uses factors given in ranges

The APPROXIMATE QUANTITIES method, involves the use of composite groups of BQ PRICING items which have similar quantities, eg. floors and walls, or similar physical functions, eg. doors and windows, with rates being derived from BQ/price book databases

The ELEMENTAL method, perhaps the first of the non traditional approaches, also uses items representing physical functions of buildings, but with quantities and rates expressed in terms of the building gross floor area.

The reliability of estimates generated by the traditional model is a function of several factors

(1) the reliability of each quantity value, q

(2) the reliability of each rate value, r

(3) the number of items, n, and

(4) the collinearity of the q and r values

This last factor is often overlooked in reliability considerations which tend to assume that q and r value errors are independent.

Traditional thinking holds that more reliable estimates can be obtained by more reliable q values (eg by careful measurement), more reliable r values (eg by use of bigger data bases) or more items, all else being equal - a proposition questioned by Fine's (1980) radical approach to BQ PRICING involving the generation of random values for both q and r values. More detailed theoretical analyses also support the view that traditional thinking may be an oversimplification. Barnes (1971), for example, investigated the implication of the proposition that different r values have different degrees of reliability, specifically that the reliability of $q_i r_i$ is an increasing function of its value. By assuming a constant coefficient of variation for each item, he was able to show that a selective reduction in the number of low valued items would have a trivial effect on the overall estimate reliability. The empirical evidence in favour of Barnes' assumption is quite strong and has culminated in the SIGNIFICANT ITEMS method now being used by PSA.

Another break with tradition has been to develop entirely new items based on a more conceptual classification of contract characteristics.

INTRODUCTION AND DESCRIPTION OF TECHNIQUES (continued)

ELEMENTAL estimating is perhaps the first sophisticated example of this, involving as it does the reorganisation of traditional BQ PRICING items into composite groups considered to represent mutually exclusive building functions. (A development of this using Cost Planning Units (CPU) provides an alternative.) Most types of approximate estimating methods fall into this category, the single rate methods such as FLOOR AREA, FUNCTIONAL UNIT, CUBE, and STOREY ENCLOSURE, or the multiple rate methods such as APPROXIMATE QUANTITIES. A further and most important characteristic of all these methods is that they were all developed (except for Ross', 1983, alternative APPROXIMATE QUANTITIES methods and the STOREY ENCLOSURE method) in the absence of any reliability measures by which to assess their value.

Research over the last 20 years has developed with different emphasis on all of the four factors influencing estimating reliability although systems development has been centred at the item level involving

(1) the search for the best set of predictors of tender price

(2) the homogenisation of data base contracts by weighting or proximity measures

(3) the generation of items and quantities from contract characteristics, or

(4) the quantification of overall estimate reliability from assumed item reliability.

The first of these is typified by the REGRESSION approach, involving the collection of data for any number of potential predictors (floor area, number of storeys, geographical location etc) and then by means of standard statistical techniques to isolate a best subset of these predictors which successfully trades off the costs of collection against the level of reliability of estimate. The second is typified by the HOMOGENISATION AIDS provided by the BCIS 'on line' and BICEP systems, and the fuzzy set based automatic procedure contained in the LU QIAN system. The third is typified by the *Holes', Calculix and expert systems* approach such as ELSIE. This is essentially a 'front end' to a conventional estimating system where items and quantities are derived from basic project information by either a known or assumed correspondence between the two. The fourth approach, typified by the use of probabilistic (statistical) models such as PERT-COST or simulation models such as RISK ESTIMATING or the Construction Project Simulator (CPS), goes beyond the standard REGRESSION approach by introducing more complicated relationships than those assumed by the standard regression method in accommodating some interdependency between and variability within r values (PERT-COST and RISK ESTIMATING) and q values (CPS).

Bidding models can be represented by

$$B = Cm \qquad (3)$$

where B is the value of the bid to be made by a contractor, C represents the estimated production costs that would be incurred should the contract be obtained, and m is a mark up value to be determined by the bidder. In bidding theory the major interest is in estimating a suitable value of m which will provide the best trade off between the probability of obtaining the contract and the anticipated profit should the contract be obtained. The model for C is similar to the design estimate model for P (eqn (2)) in that it consists of the sum of a series of quantified item and rate products, say

$$C = \sum_{j=1}^{k} q_j' r_j' \qquad (4)$$

where q_j' and r_j' represent the quantity and rate respectively for the jth item. If the same items are used in providing both design and construction estimates, this simplifies to, in terms of the design estimate model

$$B = \Sigma\, q_i' r_i' m = \Sigma\, q_i r_i = P \qquad (5)$$

as both B and P are essentially estimates of the same value - the market price of the contract.

The idea that bids and profits can be modelled in a probabilistic way ie. by treating them as random variables, has a long tradition in bidding theory, but it is generally assumed that it is the actual rather than estimated costs that contain the random component. Like design estimates however, it is becoming increasingly popular to treat the estimated costs also as a random component in the model. Also the similarity of approaches does not end at this point as recent empirical studies strongly suggest the existence of a close relationship between C and P.

(A)

STANDARD REGRESSION

Regression, or multiple regression analysis as it is usually called, is a very powerful statistical tool that can be used as both an analytical and predictive technique in examining the contribution of potential new items to the overall estimate reliability. Perhaps the most concentrated research on the use of regression in estimating was in a series of post graduate studies at Loughborough University during the 1970's and early 1980's. The earliest of these studies developed relatively simple models with new items to estimate tender prices of concrete structures, roads, heating and ventilating installations, electrical installations and offices. More recent studies at Loughborough have been aimed at generating rates for standard bill of quantities type items by regressing the rates for similar items against the quantity values for those items.

The regression model usually takes the form

$$Y=a+b_1X_1+b_2X_2+ \ldots + b_nX_n \tag{6}$$

where Y is some observation that we which to predict and X_1, X_2, ... , X_n are measures on some characteristics that may help in predicting Y. Thus Y could be the value of the lowest tender for a contract, X_1 could be the gross floor area, X_2 the number of storeys, etc. Values of a, b_1, b_2, ... , b_n are unknown but are easily estimated by the regression technique given the availability of some relevant data and the adequacy of some fairly reasonable assumptions about the data.

The regression model in eqn (6) is quite similar to the estimate model in eqn (1), as Y, X, and b values can be thought of as representing the P, q, and r values. The major difference in approach however is that in applying the regression technique *no direct pre-estimates are needed of the values of item rates, r,* as these are automatically estimated by the technique. This then obviates the need for any data bank of item rates thus freeing the researcher to examine any potential predictor items for which quantities are available. The implications of this are quite far reaching for without the need to have an item rate data base the research task is simply an empirical search for the set of quantifiable items which produces the most reliable estimates for Y.

Problems and limitations

Although the task appears to be straightforward, several problems have been encountered. These problems concern the model assumptions and limitations, data limitations, and reliability measures.

Model assumptions

The major assumptions of the basic regression model are that

(1) the values of predictor variables are exact

(2) there is no correlation between the predictor variables

(3) the actual observations Y are independent over time, and

(4) the error term is independently, randomly and identically normally distributed with a zero mean

In terms of eqn (1) this implies

that the quantities q are exact rather than approximate quantities that quantities for items do not change in tandem, as does floor area and number of storeys or concrete volume and formwork area for example

that the tender prices for one contract is not affected by the tender prices for the previous contract, and

that the differences between the regression predictions and the actual tenders are purely unaccountable random 'white noise', unrelated in any way to the variables used in the model.

Violation of these assumptions is not necessarily fatal to the technique. The type and degree of effects of violations depends on the type and degree of violation. Unfortunately however this is a rather specialist area in which statistical theory is not yet fully complete. The usual pragmatic approach to this is to try to minimise violations by a combination of careful selection of variables and tests on the degree of resulting violations.

(c) Data limitations

Although there is no theoretical limit to the number of predictor variables that may be entered into the regression model, at least one previous contract is needed for every variable. For reasonably robust results it is often recommended that the number of previous contracts is at least three times the number of variables in the model. Thus for the full traditional bill of quantities model of say 2000 items this means about 6000 bills of quantities would be needed for analysis, even assuming that each bill contains identical items. Fortunately however there is a kind of diminishing return involved in the introduction of each new variable into the model so that there comes a point at which the addition of a further variable does not significantly contribute to the reliability of the model. This property of regression analysis is often utilised by researchers a technique called forward regression which involves starting with only one variable in the model. A second variable is then added and checked for its contribution. If it is significant, a third variable is added, checked and so on until a variable is encountered that does not significantly contribute. This variable is then left out of the model and the analysis is completed. An extension of this method is stepwise regression which leaves out any non significantly contributing variable already in the model and enters significantly contributing ones until completion. Although stepwise regression works fine if the predictor variables are not correlated as required by model assumption 2 above, violations of this assumption results in different models depending on the order of variables entered and removed from the model. This can be overcome by yet another technique called best subset regression, which examines all possible combinations of predictor variables for their joint contribution, selecting the best set of predictor variables which significantly contribute.

A key issue of course is in specifying a criterion for distinguishing between significant and non significant contributions to the model. In regression formulations it is usual to concentrate on the behaviour of the error term, ie. the difference between the actual values of Y and those predicted by the regression model. If for example 10 contracts have been used in the analysis, then there will be 10 actual lowest tenders (contract values) and 10 model predictions. The differences between each of these 10 pairs of values is then squared and added together, the resulting total being called the residual sums of squares (RSS). As each new variable is entered into the equation, the RSS decreases a little. Two possible significance criteria therefore are the minimum total RSS or the minimum decrease in RSS as a new variable is entered. Clearly if this figure is set at zero, then all variables will be entered into the model. If on the other hand the minimum is set at some high level, very few if any variables will be entered into the model. Another possibility is to use the proportion of RSS to the total sums of squares that would be obtained if no variables were entered into the equation (TSS). Most standard regression packages use this latter method. For construction price-cost estimates this may not be appropriate. The decision is ultimately an arbitrary one. In most construction price-cost estimating research, the number of variables entered into the model before cut off using various sensible criteria levels is seldom more than 10, which suggests that data limitations is not likely to be a serious problem in practice.

(c) Reliability measures

This has been perhaps the biggest area for problems and misunderstandings in all of regression based research to date. The standard test statistic given by regression packages is the F value which is a measure of the proportion of RSS to TSS mentioned above. This tests the assumption (null hypothesis) that the predictor variables used in the model have no real predictive ability, such apparent ability revealed by the regression technique being more attributable to chance than some underlying correlation. Thus if we use the F value to test a one variable model regressing contract value against gross floor area it is incomprehensible that the null hypothesis should hold, and there is no practical advantage in testing against the incomprehensible.

Another very common failing has been to confuse measures of a model's fit with measures of the model's predictive ability. Measures of the extent to which the model fits the data are readily available in most standard regression packages, the most popular measures being the F test mentioned above which offers evidence on whether the model's fit is due to chance, and the multiple correlation coefficient which indicates the degree to which the model fits the data. However, as the model has been derived from the same data by which it is being tested for fit, it is not at all surprising that the fit should often be a good one. The real test of reliability of the regression model is to see how it performs in predicting some new data. The obvious way of doing this is to obtain the regression a and b coefficients from the analysis of one data set, collect some more X data for some more contracts, apply the old a and b values to obtain estimates of Y, and then use these estimates against the actual values of Y as a means of measuring the models predictive ability. A more subtle approach, called Jackknife Validation, is to omit one contract from the data, calculate the a and b coefficients, estimate the Y value of the omitted contract, and compare this with the actual Y value for that contract. This procedure is then repeated for all the contracts and then conducting the residual analysis as before.

B

Applications

The regression method involves six operations

(i) data preparation and entry

(ii) selection of model

(iii) selection of predictor variables

(iv) estimation of parameters

(v) application of parameter estimates to specific task

(vi) reliability analysis

In practice, operations (ii) and (iii) are executed concurrently.

c) <u>Data preparation and entry</u>

Data are prepared in the form of a matrix in which the rows correspond to contracts and columns correspond to the lowest tender (contract value) and values of potential predictor variables. Most regression packages can handle a few missing data which need to be flagged by the use of a special number such as 0 or 999. Contract values are normally updated by one the tender price indices although this is not strictly necessary (the tender price index applicable to each contract could be entered as a predictor variable for instance). Another possibility is to include some approximate or even detailed estimate of the contract value as a potential predictor variable also. If the latter is used then the regression is essentially a *debiasing* rather than estimating technique (these are examined later).

It is advisable to carefully check that the data is correctly prepared and entered before analysis as errors may not be immediately apparent. If errors do exist in the data, it is quite likely that a great deal of time and effort could be wasted in abortive analysis prior to their correction. Most regression packages offer a facility to reproduce the data in hard copy form for checking purposes.

(c) Selection of model and predictor variables

These operations are very well documented in many intermediate level statistics texts and regression package manuals. It is usual to carry out both operations concurrently, seeking the best model and subset of predictor variables together. Two fundamentally different approaches exist that have great significance for academic work. One, called the deductive approach, involves the proposition of specific pre analysis (*a priori*) hypotheses based on some theoretical position derived from an examination of extant ideas on the subject. The consequent models and predictor subsets are then tested against each other for primacy, any potential new models and subsets however obvious from the data being strictly excluded. The other approach, called inductive, is strictly empirical in that the intention is to find patterns in the data that can be used to generate some future hypotheses and, hopefully, theoretical foundations. For all practical purposes the separation of these approaches is hardly relevant except as a stratagem for dealing with the logistical problems that are invariably encountered in regression analysis. Construction price-cost estimation lacking any formal theoretical base tends to preclude the deductive approach and we are usually left to look for patterns in the data. This means trying out many possible models using not only the simple additive models described here but those involving transformations of variables (eg log, powers, reciprocals) together with combinations of several variables (eg products, powers) in either raw or transformed states. Even with only two predictor variables the number of combinations of transformations and combinations are quite substantial, with a large number of variables, some simplification is needed. In the absence of theory such simplification is bound to be an arbitrary process. As a result, research in regression is far from comprehensive and, because of the arbitrary means employed, has not been reported well enough to allow any incremental progress to be made. Also, with the confusion over reliability measures, it has not been possible to properly evaluate such progress that has been made.

Experience suggests that the best and easiest starting point is the simple regression model with raw (untransformed) values and no interaction terms (combinations of variables). Then, by using stepwise or best subset regression, this model can be trimmed down to significant predictors. The next stage is to try to find any other variable set or transformation that will significantly improve the model.

Parameter estimation

Once a satisfactory model is fitted to the data, the standard regression package will automatically calculate values of the a and b coefficients for use in estimating. This an extremely simple process by which the value of the predictor variables for the new contract are just multiplied by the b coefficients and added together with the a coefficient. The resulting total is the regression estimate for the contract.

Reliability analysis

The reliability of the regression model can be estimated in terms of both *bias* ie. the average error of the forecast, and *variability* ie. the spread of the forecast errors. The regression technique is designed to give unbiased predictions and therefore the average error of prediction is assumed to be zero. Two relevant measures of the likely variability of the forecast are the 95% confidence limits and the coefficient of variation (see Appendix A)

(B) Example

(insert Table 2 here)

Table 2 contains a set of data extracted from the RICS Building Cost Information Service's Brief Analysis files. A total of 28 contracts are included for office blocks from a variety of geographical locations in the UK for the period 1982 to 1988. The dependent variable contract sum (CONSUM) was standardised by dividing by the all-in tender price index and location factor to remove inflationary and location effects. The independent (predictor) variables chosen were the gross floor area (GFA) in square metres, the number of storeys (STOREY), air-conditioning (ACOND) valued at 1 if present and 0 if not present, contract period (PERIOD) in months, and the number of tenders (BIDDERS) received for the contract. The contract period and number of tenders received was not known for contract number 15 and these were given a special 'missing' value of 99.

The problem now is to provide an estimate for a new contract, not included in the data set, which has a gross floor area of 6000 square metres, 5 storeys, no air-conditioning, an 18 month contract period and 6 tenders

The first task is to fit a suitable model to the data set. From experience and a few trials a reasonable model was found by (1) using the price per square metre value (CONSUM/GFA ratios), and (2) taking the natural logs of all the variables except ACOND.

The stepwise procedure was used to enter the independent variables into the regression model one at a time. The resulting sequence of independent variables entering the model was firstly BIDDERS, then GFA, PERIOD, STOREY, and finally ACOND. Table 3 gives the results obtained as each variable was entered into the model.

(insert Table 3 here)

Up to step 5 the results are very much as expected with the number of bidders and gross floor area producing a negative regression coefficient indicating a drop in price per square metre as these variables increase due to intensity of competition and economies of scale respectively. Similarly the contract period and number of storeys produce a positive regression coefficient indicating a rise in price per square metre as these variables increase due to the effects of complexity of design and construction. The negative value for air-conditioning at step 5 however is not expected, as it indicates that air-conditioned buildings are cheaper. This, together with the increase in variability as measured by the 95% confidence limits and coefficient of variation suggests that the ACOND effect is spurious. The best looking model here seems to be that obtained at step 3, which has a better standard error of forecast and better coefficient of variation than the other models.

The model at step 3 is

$$\log(CONSUM/GFA)=a+b_1\log(BIDDERS)+b_2\log(GFA)+b_3\log(PERIOD) \quad (7)$$

where a b_1 b_2 b_3 are the constant and regression coefficients respectively. Thus the forecast of log(CONSUM/GFA) for the new contract is, from Figure 1(b)

$$\begin{aligned}\log(CONSUM/BIDDERS)&=1.7040-0.2897\times\log(6)-0.3071\times\log(6000)+0.8814\times\log(18)\\&=1.7040-0.2897(1.7918)-0.3071(8.6995)+0.8814(2.8904)\\&=1.0609\end{aligned}$$

The antilog of 1.0609 is 2.8890 which is the standardised pounds per square metre estimate, and the standardised total estimate is therefore 2.8890x6000=17334. This can now be converted into a current estimate in a given location by multiplying by the current tender price index and location factor. If the new contract is for January 1989 (TPI=300) in Salford (Location Factor=0.90), the estimate will be 17334x300x0.90= £4680180.

The 95% confidence limits of £2565918 and £8536266 (see Appendix A) are very large - a reflection to some extent on the limited amount of data used in this example.

(insert Figs 1(a) to 1(d) here)

A

CONSTRUCTION PROJECT SIMULATOR

The Construction Project Simulator (CPS) was developed between 1980 and 1984 at the University of Reading by John Bennett and Richard Ormerod. It is a resource type system, but with a unique facility to model some special aspects of uncertainty particularly prevalent in building work - productivity variability and external interferences to the construction process on site. Data describing a particular project - bar charts, direct and indirect costs, resources, weather, and productivity - is input and fed to a series of stochastic simulation programs which compute cost and time estimates for the project. The result is a histogram of time and associated costs which can be used as a measure of reliability of the estimates.

The system is based on the model

$$P = \sum_{i=1}^{n} t_i r_i + \sum_{j=1}^{m} u_j r_j \qquad (8)$$

where t is a stochastic (random) variable, $t_i \sim F(\mu_i, \sigma_i)$, representing the time taken to perform the ith activity, u represents the remaining non-stochastic quantities such as materials, and r the appropriate unit rates for converting the quantities into costs. The actual model used by the system is rather more sophisticated than this as activities may be delayed by stochastic interferences and the overall effects automatically rescheduled to enable dependent items such as preliminaries to be costed.

3 **Problems and limitations**

c <u>Model assumptions</u>

Although the CPS is undoubtedly superior in many respects to similar resource based systems, the basic model contains two particular limitations at the present time. Firstly, in assuming stochastic independence, any actual ***interdependences*** between future activity times are currently incompletely modelled by the system. Thus, under the model assumptions, any time slippage caused by some future chance delay could only be recovered by some equally chance future hastening event, whilst it is conceivable that a work team or manager may take direct remedial action in the live situation. Similarly, and on the other hand, the model makes no assumption that a chance exceptional productivity rate on one activity would be accompanied by a similarly exceptional rate on another activity, although this situation may well arise in reality due to some common characteristics of the people involved in both activities. Secondly, the non-stochastic components in the model - material quantities and unit rates - are also known to be variable to some degree and therefore may be better treated in a stochastic manner. The research and limited tests on the system to date however suggests that neither of these modelling limitations significantly affects the predictive abilities of the system.

c) <u>Data limitations</u>

As with all resource based systems, the estimation of likely resource demands requires a knowledge that is usually restricted to the constructor. Thus the very nature of design and construct, concerned with the product and the process respectively, mitigates against the use of the CPS by designers.

In gathering data on variability it is desirable to compare like with like. The quantity and quality of work carried out, and the method of execution should be identical or at least very similar. Bennett and Ormerod's attempts to quantify variability and differentiate between different activities however have been plagued by a dearth of relevant information. Their major source is the Building Research Establishment although even this has limitations on the type of building (traditional housing) and accuracy.

Relevant data on interferences is also severely restricted at present mainly due to the relatively recent recognition of interference as a major contributor factor. The extent of subcontractor non-attendance or labour supply problems for instance has not been accurately recorded by contractors, although some information on this is, somewhat surprisingly, available from designer sources and the management of interferences literature.

Perhaps the best and most reliable source of data is that issued by the Meteorological Office on the seasonal frequency of weather conditions. These can be obtained in the form of frequency distributions, an ideal representation for stochastic simulation.

c <u>Reliability measures</u>

This is certainly the strongest feature of the CPS system. The system's facilities include an excellent graphical display of both histogramic and cumulative frequencies of time and costs, enabling the probability of a particular cost being exceeded to be obtained at a glance.

Applications

The system has been developed specifically for ease of use by non-experts with a limited amount of computer equipment at their disposal. The main applications features concern the entry of appropriate bar charts, unit costs, resourcing and frequency distributions, and the simulation procedure.

Bar charts

The CPS bar charts are arranged as a hierarchy consisting of a primary level in which the user defines separate activities or Primary Work Packages (PWP) such as substructure, external envelope, internal subdivisions etc., and a secondary level or Secondary Work Packages (SWP) defining constituent activities of PWP's such as excavate over site, excavate bases, mass fill bases etc.

The bar charts are created directly on the computer screen by modifying charts for previous projects - mainly the start and end dates - up to a maximum of 39 bars each for PWP's and SWP's.

The logical links (maximum 250) acting as constraints to progress are next entered by the user and drawn on the screen as thin lines.

Holiday periods (maximum 9) can also be entered in much the same manner as the bars are entered.

Following entry or correction, the bars and links are automatically rescheduled before any further processing commences.

(c) Unit costs

On completion of the bar charting, the user is presented with a tabulated version of the activities against which unit costs are displayed for labour and materials separately. Temporary work activities (preliminaries) and their associated costs are entered separately by a special program facility.

(c) Resources

Resources are entered in a way common to most resource estimating packages by means of a resource library or data base. Typical trades and unit costs are contained in 100 categories. Gang compositions are built up in a separate file and these are then attached to the activities contained in the SWP's with the aid of a further program for the purpose. Various displays are available to enable the user to experiment with different combinations and quantities of resources to obtain the best mix.

Frequency distributions

A wide variety of theoretical distributions (uniform, triangular, normal, beta, etc.) and empirical distributions derived from live data are available through one of the CPS data bases. The user selects one of these distributions and enters it against each activity together with the appropriate parameters for the theoretical distributions. Interference factors are dealt with in a similar manner by entering a number between 0 and 99 to define the percentage chance that an interference will occur at any link position on an activity bar.

Twelve location specific histograms for the weather for each month of the year are entered into a separate program.

The simulation procedure

Simulation can be carried out at various stages, so that data input can have different degrees of detail and simulations still performed. A simulation of the primary level only may be carried out after the primary bar chart and weather data, and optionally the preliminary schedule and primary cost table, have been entered. This allows a quick assessment of schemes at an early stage in their development, or at tendering stage.

A simulation of one or more secondary plans may be carried out after a secondary bar chart, and optionally the secondary cost table and resources details have been entered. This allows the assessment of single PWP's when design details become available as the scheme develops, or allows the re-assessment of a PWP after some change to the construction method.

A full scale simulation of all the secondary plans, followed by a simulation of the primary level can be carried out automatically. This allows the assessment of the whole scheme, at a level consistent with a good contract programme.

The manner in which this is achieved is via the use of pseudo random numbers (RNs) generated by the computer to the frequency distributions specified. The procedure used by the program on one simulation iteration is to take each bar segment defined by logical links in turn and to choose an actual duration via RNs, the assigned distribution and the variability amount. Each link position is then examined to see if an interference occurs there and the duration of the interference is chosen by an RN if an assigned interference distribution is present. This process is repeated for each bar and the whole chart is then rescheduled by the logical restraints. The cost of each bar is then calculated from the ratio of the actual bar duration to the original bar duration multiplied by the labour cost and added to the material cost, or if resources are involved the labour cost is calculated from the resource costs. The final duration and cost are then recorded. The program allows up to 200 such iterations, and the resulting different duration and cost estimates are displayed at the end.

Further facilities allow the results of the secondary level simulation and the effects of weather and preliminaries processing to be fed into the primary level to allow the fine detail to affect the overall result.

3

Example

Figure 2(a) shows the primary bar chart with 17 PWP bars for a contract extending over a period from 1980 to 1982. Figure 2(b) shows the secondary level bar chart for PWP 2 (substructure) containing 33 SWP bars. Figure 2(c) shows the cost screen for PWP 2 with the labour, material and total costs. Figure 2(d) shows the preliminaries schedule superimposed on the primary bar chart, the heavy dashed lines representing the duration of each preliminary category. Figure 2(e) shows part of the resource library which is used to indicate the man and simulation of the primary level only may be carried out after the primary bar chart and weather data, and optionally the preliminary schedule and primary cost table, have been entered. This allows a quick assessment of schemes at an early stage in their development, or at tendering stage.

A simulation of one or more secondary plans may be carried out after a secondary bar chart, and optionally the secondary cost table and resources details have been entered. This allows the assessment of single PWP's when design details become available as the scheme develops, or allows the re-assessment of a PWP after some change to the construction method.

A full scale simulation of all the secondary plans, followed by a simulation of the primary level can be carried out automatically. This allows the assessment of the whole scheme, at a level consistent with a good contract programme.

The manner in which this is achieved is via the use of pseudo random numbers (RNs) generated by the computer to the frequency distributions specified. The procedure used by the program on one simulation iteration is to take each bar segment defined by logical links in turn and to choose an actual duration via RNs, the assigned distribution and the variability amount. Each link position is then examined to see if an interference occurs there and the duration of the interference is machine power types needed for each activity, and Figure 2(f) shows the resources allocation screen which is used to indicate the quantity of each man or machine power type needed for each secondary level activity. Thus activity 5 (mass fill bases) requires three of resource set 24 (concreting only) which happens to comprise resource code 28 (concretor) at a unit cost of 100 (from Figure 2(f)). In other words the mass fill bases activity PWP 5 requires three concretors at 100 cost units each, ie. 300 units.

Figure 2(g) shows the results obtained after 200 iterations of the simulation procedure. The average estimated cost and 95% confidence limits can be read off the cumulative frequency curve at the appropriate 50 ,2.5 and 97.5 percentage points, in this case £5710600 average with a range of £5591000 to £5852500, or £5710600 -2.1% +2.5%.

(insert Figs 2(a) to 2(g) here)

(A)

LEAD CONSULTANT EXPERT SYSTEM (ELSIE)

In 1982, a government committee under the chairmanship of John Alvey issued a report concerning Japanese research into 'Fifth Generation' computers, a general term covering several new computer technologies such as artificial intelligence and expert systems. As a result, the government created the Alvey Directorate, with a budget of £350 million, to operate a five year collaborative research programme for the UK. One of the ensuing collaborative activities within the Intelligent Knowledge Based Systems (IKBS) sector was a 'community club', formed in 1986, comprising the members of the Royal Institution of Chartered Surveyors' Quantity Surveying Division (represented by a small group of practising chartered quantity surveyors) and a research team from the University of Salford (consisting of two knowledge engineers and a quantity surveyor). The original intention of the club was to examine the potential of Expert Systems in quantity surveying and this rapidly developed into an exercise in system building around the four core tasks involved in early stage strategic planning of construction projects - initial budget estimates, procurement choices, development appraisals, and duration estimates. One of the major achievements of the 18 month research was the initial budget estimating facility of the system - a unique 'smart front end' to what is essentially a conventional approximate quantities estimating system. This is done by means of a knowledge based expert computer program which acts upon information more normally found in a project brief and automatically converts it into approximate quantity type information by making some reasonably intelligent assumptions. Like all expert systems, an important feature of ELSIE is that it can be interrogated about the assumptions it has made and these can be checked and amended where needed.

The major objective of the initial budgeting facility of ELSIE is to allow a budget to be set for the development of an office building, at an early stage in the project. In particular, it is designed to be used before Scheme Design drawings are available, but information such as sketch drawings can be utilised if available. It is designed to be used by semi-expert personnel, rather than novices, and in a way which will often complement rather than replace their activities.

(5) Problems and limitations

In general the problems and limitations of the estimating part of the system are no different to those associated with any approximate quantities type technique - item rates need to be current and appropriate to the type of project involved (in ELSIE's case the latter is overcome by restricting present applications to office contracts only). The unique front end feature of the system however does involve some special considerations concerning the validity of the knowledge base.

The major problem concerning the validity of the knowledge base is whether and to what extent the design assumptions generated by the system correctly anticipate those of the designer. Innovative design or just simply changing fashions for example cannot be adequately handled by the system at present without intervention of the user. In general however the system is claimed to make design decisions of a kind that will be made up to 5 years after the start of the research, ie. 1991. Obviously the validity of this claim remains to be seen.

Although in many cases the fashion element of design assumptions is made explicit so as to isolate the pertinent bits of knowledge base when modification is necessary, the facilities for user modifications are at present limited to unit costs and some certain default values within the data base. Changes to the knowledge base generally can be made only by the software support team.

One further limitation of the system as a means of obtaining early stage estimates is that no information is produced to enable any kind of objective assessment of the reliability of the estimates. The only documented claim that has been made in this respect is that the estimates produced in testing the system (on about 40 building projects) were in general "within about 5% of that predicted by the expert quantity surveyor" (Brandon *et al*, 1988, p51), and in a number of cases better than the human expert. This figure should be treated with some caution however as it is likely that several of these tests were used in order to fine tune the program, and therefore are equivalent to the regression model derivation data. It is quite possible though, with this type of system, that the combination of human and computer 'expertise' working in tandem may well be greater than each working in isolation.

8

Applications

The initial budget facility of the ELSIE system, like most expert systems, operates in an interrogative manner asking the user a series of questions concerning the basic characteristics of the project from which the system can deduce a set of element type unit quantities and thence the likely building cost-price. The questions asked depend on the answers given to previous questions, but usually around 25 questions need to be answered. Once an answer is given the user has a chance to change some of the assumptions that have been made, before arriving at what is called the What-Now point.

Two basic options are available at the What-Now point, (1) make changes (in up to 150 items), (2) obtain reports. The facility to change the assumptions made, or the answers originally given to the questions, gives a comprehensive 'What-if' procedure, allowing the user to investigate the cost-price implications of various options and enabling the user to gain an impression of the sensitivity to design variation. The reports contain in varying degrees of detail - cost breakdowns, explanations, and graphical representations of relationships such as quality with cost-price.

5

Example

This example illustrates the procedure used to produce a November 1986 estimate for a five storey speculative insurance company office building in the UK South East region for which the following information is available

(i) 3342 m^2 gross floor area, 100% lettable

(ii) Brick elevation

(iii) No basement

(iv) 15 months contract period

(v) 900 m^2 level town site with access problems

(vi) VAV air-conditioning in offices only

(vii) Very noisy location

(viii) 5 car parking spaces

Figures 3(a)-(e) shows some of the screens involving the question and answer Brief elicitation phase. Figure 3(f) gives the resulting cost-price estimate generated by the system (£2.38m) at the 'What now' stage. Figures 3(g)-(i) shows the first 3 pages of the explanations report (12 pages in total were produced), including an elemental breakdown of the estimate.

(insert Figs 3(a) to 3(i) here)

(A)

BIDDING MODEL DEBIASER

Although not strictly estimating techniques, estimate debiasers constitute a collection of very new 'back end' techniques, still in the research phase, aimed at improving or fine tuning estimates generated by other techniques. Three types of debiasers are under current development

(1) REGRESSION debiasers, which are identical to the usual regression estimating techniques except that the estimate is specifically included in the predictor variable list

(2) CONTROL CHART debiasers, using dynamic time series detrending techniques to detect real time biasing of recent estimates, and

(3) BIDDING MODEL debiasers, which utilise bidding theory to assess bias and reliability in estimates

This section describes one of the bidding model debiasers being developed by the author.

The purpose of bidding models is to enable a bidder to assess the best mark up value to use in an auction given some information concerning the likely bids to be entered by his competitors. This information can range from a knowledge of the identity and past bids of all the competitors in the auction through to virtually no information at all. Many formulations have been proposed to model this situation. The model adopted here is the multivariate model (Appendix B). As the probability of entering the lowest bid with a given bid is exactly the same as the probability that the lowest bid will be less than an estimate of the lowest bid, the model can be applied without modification.

Problems and limitations

The problems and limitations are similar to those of the regression estimating approach in concerning model assumptions and data limitations.

Model assumptions

Three major assumptions concerning the validity of eqn (B.1) need to be addressed

(1) the log normal assumption

(2) the independence assumption, and

(3) the consistency assumption

The first of these assumptions is dealt with in Appendix B.

Violations of the independence assumptions may have much more severe consequences. Independence is however very difficult to establish with data of these kind, and possible violations tend to be ignored for this reason. It is probably now a truism that possible lack of independence is the reason for many models of this kind failing to achieve commercial status.

The consistency assumption, ie. that bidders behave in much the same way irrespective of the type, size and other characteristics of contracts, is a reflection on the simplicity of the model (of which the independence assumption is a special case). Some research is currently proceeding on this aspect.

C <u>Data limitations</u>

Although most bidding models have very heavy demands on data, particularly on the frequency with which certain specified bidders compete against each other, the multivariate model is relatively undemanding in this respect.

The data consist of any previous designer's estimates together with contractors' bids. It is not necessary that designer's estimates are available for all the contracts in the data, nor that the designer's estimates and the specified contractors' bids are recorded for the same contract. All that is necessary is an indirect link between the designer and the specified bidders. For example, if contractors A, B, and C are bidding for a new contract, it is not important that the designer has produced estimates for previous contracts on which A, B, and C have entered bids, nor that any of the people involved have bid against each other before. All that is required is that all the people involved have bid at least once against another bidder who has bid against at least once against the current competitors, or have bid at least once against another bidder who has bid at least once against another bidder who has bid at least once against the current competitors etc.

There is a price to pay however for the relative lack of data restriction, reflected in the independence and consistency assumptions mentioned above.

Applications

The bidding model debiaser involves three stages

data preparation and entry

estimation of parameters, and

calculation of probability of lowest bid values

(c) Data preparation and entry

The data consists of all available designer's estimates, bids and associated bidder' names for a set of historical construction contract auctions. These are entered into an auction data base in the form of a contract number, designer's estimate value/designer code, and bid value/bidder code

(c) Estimation of parameters

The computer program automatically calculates the required model parameters from eqns (B.2) and (B.3) and stores the results in a computer file. This operation is only necessary when new historical auction data is entered into the auction data base, and takes a few seconds of computer time.

(c) Calculation of the probability of the lowest bid values

The computer program automatically calculates the unbiased estimated probability values for each of a sequence of m values and plots the resulting curve in terms of P'. Further probability estimates are then obtained via estimates of μ and σ^2 obtained by stochastic simulation, each iteration generating a different probability value. These additional values provide an indication of the variability of the probability estimates and are plotted as points on the graph. The resulting graph therefore enables the user to gain an impression not only of the reliability of the debiased estimate but also of the reliability of the reliability!

Example

This example contains data collected from a London building contractor (coded number 304) for an incomplete series of construction contracts auctioned during a 12 month period in the early 1980's (Table 4). For the purposes of this example, the bids entered by bidder 304 are treated as designer's estimates.

(insert Table 4 here)

The program requests the value of the estimate for the new contract together with the identity of the bidders. In this example the estimate of the new contract is £3m and the bidders are code 55, 73, 134, 150 and 154. The program then automatically proceeds as follows

(i) transforms the data to the log values $y_{ik}=\log(x_{ik}-\lambda x_{(1)k})$, in this case $\lambda=0.6$

(ii) calculates the required model parameter estimates, α_i, β_k, and s^2_i

(iii) calculates the probability of code 304 'underbidding' the other bidders with m=-0.70,-0.69, etc. by substituting α_i and s_i for μ_i and σ_i in eqn (B.6)

(iv) generates a value for μ_{304}, σ^2_{304}, μ_{55}, σ^2_{55}, μ_{73}, σ^2_{73}, μ_{134}, σ^2_{134}, μ_{150}, σ^2_{150}, μ_{154}, σ^2_{154}, by stochastic simulation

(v) calculates the probability of code 304 'underbidding' the other bidders with m values obtained by stochastic simulation

(vi) repeats (iv) and (v) 600 times

(vii) detransforms the m values to P' by eqn (B.10) and plots the curve resulting from step (iii) and points resulting from steps (iv) to (vi)

The resulting graph is shown in Figure 4. The graph is interpreted by drawing a horizontal line at the 50, 2.5 and 97.5 percentage probability points across to the curve and thence down to the estimate axis as shown to obtain the unbiased estimate (£2827300) and 95% confidence limits (£2566700 to £3138000, ie. £2827300 +10.99% -9.22%) due to the variability of the designer's estimates and contractors' bids *as predicted by the model.*

The surrounding points indicate the effect of the size of the data base on the reliability of the parameter estimates in the model - the true curve will be contained somewhere within these points. With the small amount of data used in this example, the points are quite widespread. The existence of a larger data base should have the effect of decreasing this spread.

(insert Fig 4 here)

(A)

CONCLUSION - THE FUTURE?

Contract price forecasting techniques clearly comprise a large topic area worthy of a book in its own right. As Table 1 indicates, the field is rapidly developing out of the older deterministic approaches into methods which specifically accommodate the inherent variability and uncertainties involved in forecasting the price of construction work. One result of this is that the traditional distinction between 'early stage' or 'conceptual' estimating and 'later stage' or 'detailed' estimating is being replaced by the more fundamental distinctions concerning the reliability of forecasts and their components - items, quantities and rates. This has focussed attention on the means of modelling and predicting reliability - statistically for simplicity and stochastically for complexity. The CONSTRUCTION PROJECT SIMULATOR, for instance, contains stochastic elements for item quantities, whilst RISK ESTIMATING utilises both statistical and stochastical techniques. Little has been done to treat the items themselves in this way, although some relatively new QUANTITY GENERATION systems, such as ELSIE, are clearly capable of extension. The logical conclusion of these approaches will be a technique which combines all three elements of the forecasting equation into the same non deterministic, item, quantity, rate, (NDIQR) system.

Although a somewhat daunting prospect for practitioners, the development of NDIQR systems will mark a new and exciting phase in the evolution of construction price forecasting systems generally. Firstly, current deterministic requirements will still be accommodated as deterministic forecasts are simply a special case for a non deterministic system. A rate with mean say £5 and standard deviation £0 is effectively a deterministic rate. Also the 'best guess' of a non deterministic system *is* a deterministic answer. Thus the range of forecasts provided by a non deterministic system can be regarded as secondary information to the deterministic forecast, to be divulged or not as the user wishes. Secondly, NDIQR systems will allow forecasts to be made *at any stage of the design process*. Treating the items themselves as random variables, for example, means that the standard deviation simply reduces as we become more certain that the item will be appropriate. Thus a BQ

PRICING NDIQR system will commence with a notional bill of quantities that will gradually firm up as the design progresses. Thirdly and perhaps most importantly, the reliability measures provided by NDIQR systems will enable comparisons to be made between alternative systems. For example, if a practice uses several systems to provide price forecasts for the same contract and obtains the following results

System	Forecast Range
A	£3.2m to £3.9m
B	£3.5m to £4.0m
C	£3.7m to £4.5m

we may select the inner range of these three systems, ie. £3.7m to £3.9m, tc be the best range of forecasts.

A

REFERENCES

Barnes, N.M.L. (1971) *The design and use of experimental bills of quantities for civil engineering contracts* PhD thesis, University of Manchester Institute of Science and Technology.

Bennett, J. and Ormerod, R.N. (1984) *Construction Management and Economics* 2, 225-263. "Simulation applied to construction projects"

Brandon, P.S., Basden, A., Hamilton, I.W., Stockley, J.E. (1988) *Application of expert systems to Quantity Surveying: the strategic planning of construction projects* The Royal Institution of Chartered Surveyors. ISBN 0 85406 334 X

Fine, B (1980) *Construction Management Laboratory*. Fine, Curtis and Gross

Gilmore, J. and Skitmore, M. (1989) *Chartered Quantity Surveyor*. May, "A new approach to early stage estimating"

Lu Qian (1988) *Construction contracting in China*. Hong Kong Polytechnic, p113-125 "Cost estimation based on the theory of fuzzy sets and predictive techniques"

Southgate, A. (1988) *Chartered Quantity Surveyor*. Nov. p35-36 "Cost planning - a new approach"

Ross, E. (1983) *A database and computer system for tender price prediction by approximate quantities*. MSc project report, Loughborough University of Technology.

Appendix A Reliability of Regression Forecasts

(a) 95% confidence limits

This is obtained from the standard error of the forecast SE(Y) where

$$SE(Y)^2 = S^2(1+1/N) + \sum_{i=1}^{k} (x_i-\bar{x}_i)^2 S^2 c^{ii} + \sum_{i=1}^{k-1} \sum_{j>i}^{k} 2(x_i-\bar{x}_i)(x_j-\bar{x}_j) S^2 c^{ij} \qquad (A.1)$$

where S^2 is the mean square of the residuals, N is the number of previous cases, $\bar{x}_1$, $\bar{x}_2$, $\bar{x}_3$ etc. are the mean values of the independent variables x_1 x_2 x_3 etc, and S^2c^{ij} are the variance-covariances of the regression coefficients.

The 95% confidence limits are then approximated by

$$\pm t_{(0.025)(N-n-1)} SE(Y) \qquad (A.2)$$

where $t_{(0.025)(N-n-1)}$ is obtained from the Students t distribution tabulated in most elementary statistical texts.

(b) Coefficient of variation

The coefficient of variation, cv, can be obtained from the distribution of jackknife deleted residuals as follows

$$cv = 100S_d/(\bar{x}_p-\bar{x}_d) \qquad (A.3)$$

where x_d and S_d represent the mean and standard deviation of the deleted residuals and x_p represents the mean prediction. This statistic, though not conventionally used in regression applications, has the advantage of being directly comparable with the variability measures associated with other techniques and studies.

(c) Reliability of forecast for Salford Offices example

The standard error of forecast is obtained from eqn (A.1) as

$$SE(Y)^2=S^2(1+1/N)+(x_1-\bar{x}_1)^2S^2c^{11}+(x_2-\bar{x}_2)^2S^2c^{22}+(x_3-\bar{x}_3)^2S^2c^{33}+2(x_1-\bar{x}_1)(x_2-\bar{x}_2)S^2c^{12}+2(x_1-\bar{x}_1)(x_3-\bar{x}_3)S^2c^{13}+2(x_2-\bar{x}_2)(x_3-x_3)S^2c^{23} \qquad (A.4)$$

where x_1 x_2 x_3 represent the log values of BIDDERS GFA and PERIOD respectively. The regression output (Figures 1(a) and (b)) at step 3 shows that

$S^2=0.07648$, $N=27$, $\bar{x}_1=1.703$, $\bar{x}_2=7.103$, $\bar{x}_3=2.336$, $S^2c^{11}=0.01211$, $S^2c^{22}=0.00858$, $S^2c^{33}=0.07333$, $S^2c^{12}=-0.00390$, $S^2c^{13}=0.01126$, $S^2c^{23}=-0.02226$.

Substituting into the above equation gives

$$SE(Y)^2=0.07648(1+1/27)+(\log6-1.703)^2 0.01211+(\log6000-7.103)^2 0.00859+(\log18-2.336)^2 0.07333-2(\log6-1.703)(\log6000-7.103)0.00390+2(\log6-1.703)(\log18-2.336)0.01126-2(\log6000-7.103)(\log18-2.336)0.02226$$
$$=0.0793+0.0001+0.0219+0.0225-0.0011+0.0011-0.0197$$
$$=0.0844$$

so $SE(Y)=0.2905$

The 95% confidence limits for the standardised log forecast is then, from eqn (A.2)

$$Y\pm t_{(0.025)(27-3-1)}0.2905$$
$$=1.0609\pm2.069\times0.2905$$
$$=1.0609\pm0.6010 \text{ ie. } 0.4599 \text{ to } 1.6619$$

which is 1.5839 to 5.2693 for the standardised forecast, and £2565918 to £8536266 for the Salford January 1989 contract.

The coefficient of variation is obtained by dividing the standard deviation of the forecast error ratios by their mean. As a log model has been used, the antilog of the deleted residual will give the required ratio. At step 3 the standard deviation is 0.3339 and the mean 1.0419, giving a coefficient of variation of 32.05%.

Some statistical packages make the calculations easier by providing a direct forecast for the new contract. In this example we have entered the new contract into the data base with a dummy value of £0.01 but excluding it from the model by a special select instruction. As Figure 1(c) shows, the correct forecast of 1.0609 is obtained. This package also gives the standardised deleted residual SDRESID, which is the deleted residual DRESID divided by its standard error. The standard error of the forecast can therefore be obtained very quickly as

$$SE(Y)=DRESID/SDRESID \qquad (A.5)$$
$$=-14.3656/-48.4444=0.2905$$

The residual statistics (Figure 1(d)) also enable a quick approximation of the coefficient of variation to be made by

$$cv=100\times\text{std dev DRESID}/(\text{mean PRED}-\text{mean DRESID}) \qquad (A.6)$$
$$=100\times0.3751/(1.0880-0.0187)=35.08$$

Although this figure is the coefficient of variation of the log values, it is proportional to the raw coefficient of variation and therefore indicative of the relative values. Also, as the model becomes more reliable, then the coefficient of variation of the logs values becomes closer to the coefficient of variation of the raw values

A)

Appendix B The Multivariate Bidding Model Debiaser

The multivariate bidding model is

$$\log(x_{ik}) = y_{ik} \sim N(\mu_i+\mu_k, \sigma_i^2) \tag{B.1}$$

where x_{ik} is bidder i's bid (i=1,2,...,r) entered for auction k (k=1,2,...,c), the log of which is normally distributed with mean μ_i, a bidder location parameter, plus μ_k, an auction size datum parameter, and with a unique variance parameter $\sigma^2{}_i$ for each bidder. These parameters can be estimated quite easily by an iterative procedure solving

$$\beta_k = \hat{\mu}_k = \sum_{i=1}^{r} \delta_{ik} \ (y_{ik}-\alpha_i)/n_i \tag{B.2}$$

$$\alpha_i = \hat{\mu}_i = \sum_{k=1}^{c} \delta_{ik} \ (y_{ik}-\beta_k)/n_i \tag{B.3}$$

and thence

$$s_i^2 = \hat{\sigma}_i^2 = \sum_{k=1}^{c} \delta_{ik} \ (y_{ik}-\alpha_i-\beta_k)^2/\{(n_i-1)(1 - \frac{c-1}{N-r})\} \qquad \text{for } n_i>1 \tag{B.4}$$

$$= \sum_{k=1}^{c} \delta_{ik} \ (y_{ik}-\alpha_i-\beta_k)^2/(N-c-r+1) \qquad \text{for } n_i=1 \tag{B.5}$$

where Kronecker's δ_{ik} = 1 if bidder i bids for auction k
= 0 if bidder i does not bid for auction k

$n_i = \sum_{k=1}^{c} \delta_{ik}$, the number of bids made by bidder i

$N = \sum_{k=1}^{c} n_i$, the total number of bids by all bidders

The probability of a bidder, say i=1, underbidding a set of specified competitors on a contract is then given by

$$\Pr(y_1+m<y_i, i\neq 1) = \int_{-\infty}^{\infty} \{(2\pi)^{1/2}\}^{-1} \exp(-\tfrac{1}{2}y_1^2) . \left\{ \prod_{i=2}^{n} \int_{y_i=(\sigma_1 y_1+\mu_1+m-\mu_i)\sigma_i^{-1}}^{\infty} \{(2\pi)^{1/2}\}^{-1} \exp(-\tfrac{1}{2}y_i^2) \, dy_i \right\} dy_1 \tag{B.6}$$

where m is a 'decision' constant used by bidder 1 to bring about a desired probability state (the usual approach is to use bidder 1's cost estimates in the analysis in preference to his bids on the assumption that the m value will reasonably approximate his likely profit should he acquire the contract). To observe the effects of the limited accuracy of the parameter estimates due to the sample size, values for μ and σ are obtained from α and s^2 from their sampling distributions

$$\mu_{\hat{i}} \sim N\ (\alpha_{\hat{i}},\ s_{\hat{i}}^2/n_{\hat{i}}) \tag{B.7}$$

$$\sigma_i^2(n_i-1)s_i^2 \sim \chi^2(n_i-1) \qquad \text{for } n_i > 1 \qquad \text{(B.8)}$$

$$\sigma_i^2(N-c-r+1)/s_i^2 \sim \chi^2(N-c-r+1) \qquad \text{for } n_i = 1 \qquad \text{(B.9)}$$

Now if we substitute the designer's estimator for bidder 1 in the above formulation, the value of m which results in a probability of 0.5 represents the bias in his (log) estimate and thus the amount that needs to added to his estimates to give an unbiased estimate of the lowest bid. Also the values of m which result in probabilities of 0.025 and 0.975, will give the 95% confidence limits.

The log normal assumption has been tested with three sets of UK construction contract bidding data indicating that a three parameter log normal model may be more appropriate than the general two parameter model proposed here. The modifications necessary to convert the formulation are quite straightforward however, involving a prior transformation of the data before applying the iterative procedure. The resulting m values have however to be detransformed before plotting the final probability graph. Thus, for data transformed by $y_{ik} = \log(x_{ik} - \lambda x_{(1)k})$, the unbiased estimate P' is given by

$$P' = \frac{\lambda P e^{\omega}(1-e^{m})}{1-\lambda+\lambda e^{\omega}} + Pe^{m} \qquad (\omega = m \mid Pr=0.5) \qquad \text{(B.10)}$$

```
21 MAR 89     example of regression analysis estimating for bcis offices
09:55:05      University of Salford          Prime 9955       rev 20.1.0

                              * * * *   M U L T I P L E   R E G R E S S I O N   * * * *

Listwise Deletion of Missing Data
   Selecting only Cases for which    CASENO LE    28.00

             Mean  Std Dev  Label

PRICE       1.088     .395
GFA (x2)    7.103    1.279
STOREY      1.082     .566
ACOND        .333     .480
PERIOD(x3)  2.336     .436
BIDDERS(x1) 1.763     .536

N of Cases =     27
```

Figure 1 (a) Regression output at step 3

```
21 MAR 89    example of regression analysis estimating for bcis offices
09:55:07     University of Salford          Prime 9955       rev 20.1.0
                                  * * * *   M U L T I P L E   R E G R E S S I O N   * * * *

Equation Number 3     Dependent Variable..   PRICE

Descriptive Statistics are printed on Page    3

Beginning Block Number  1.  Method:  Enter      BIDDERS   GFA       PERIOD

Variable(s) Entered on Step Number  1..    BIDDERS
                                    2..    PERIOD
                                    3..    GFA

Analysis of Variance
                      DF        Sum of Squares       Mean Square
Regression             3               2.28993            .76331
Residual              23               1.75910      (s²)  .07648

F =        9.98020        Signif F =  .0002

Var-Covar Matrix of Regression Coefficients (B)
Below Diagonal:   Covariance     Above:  Correlation

              BIDDERS          PERIOD          GFA

BIDDERS  (s²c¹¹)  .01211          .37795      -.38235
PERIOD   (s²c¹³)  .01126  (s²c³³) .07333      -.98752
GFA      (s²c¹²) -.00390  (s²c²³) .02226  (s²c²²) .00858

------------------ Variables in the Equation ------------------

Variable              B          SE B         Beta          T    Sig T

BIDDERS (b₁)     -.289683     .110059     -.393112     -2.632    .0149
PERIOD  (b₃)      .881413     .279786      .974902      3.255    .0035
GFA     (b₂)     -.307112     .092623     -.995027     -3.316    .0030
(Constant)(a)    1.704044     .357162                   4.771    .0001

End Block Number   1   All requested variables entered.
```

Figure 1 (b) Regression output at step 3

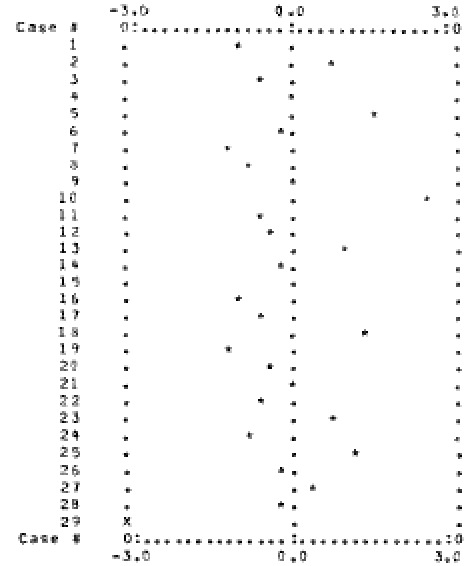
21 MAR 89   example of regression analysis estimating for bets offices
09:55:07   University of Salford   Prime 9955   rev 20.1.0

* * * *   M U L T I P L E   R E G R E S S I O N   * * * *

Equation Number 3   Dependent Variable..   PRICE

Casewise Plot of Standardized Residual

*: Selected   M: Missing   X: Unselected



Case #	PRICE	*SDRESID	*DRESID	*ADJPRED	EXP ADJRRED/PRICE	EXP − EXP − DRESID
1	.88	−.9860	−.2823	1.1508		1.33
2	1.15	.7307	.2103	.9370		0.81
3	.99	−.5942	−.1737	1.1681		1.19
4	1.00	.1031	.0399	.9733		0.97
5	2.68	3.0287	1.1739	1.5075		0.31
6	.90	−.1312	−.0392	.9437		1.04
7	.64	−1.3530	−.3909	1.0239		1.46
8	.86	−.8505	−.2501	1.1093		1.28
9	1.21	−.0137	−4.0407E−03	1.2109		1.00
10	1.78	3.0960	.7677	1.0117		0.46
11	.75	−.5638	−.1671	.9190		1.18
12	1.07	−.3665	−.1054	1.1785		1.11
13	1.12	.9905	.2922	.8259		0.75
14	.81	−.1842	−.0555	.8650		1.06
15	.79	.	.	.		–
16	.99	−1.0747	−.3048	1.2919		1.36
17	.85	−.5639	−.1722	1.0198		1.19
18	1.29	1.6465	.4569	.6248		0.63
19	1.30	−1.7940	−.6667	1.9628		1.95
20	.89	−.4103	−.1216	1.0123		1.13
21	.90	.0121	3.5238E−03	.8930		1.00
22	.79	−.5588	−.1615	.9553		1.18
23	.93	.9884	.3222	.6052		0.72
24	.97	−.9481	−.2892	1.2637		1.34
25	1.20	1.3244	.3945	.8013		0.67
26	1.17	−.3167	−.1036	1.2716		1.11
27	1.36	.5284	.1692	1.1927		0.85
28	.90	−.1523	−.0444	.9446		1.05
29	−13.30	−49.4444	−14.3656	1.0609		–
Case #	PRICE	*SDRESID	*DRESID	*ADJPRED		

$\bar{x}$ 1.0419
σ_{n-1} 0.3339
cv. 32.05%

Figure 1 (c) Regression output at step 3

21 MAR 89 example of regression analysis estimating for b.c.s offices
09:55:37 University of Salford Prime 9955 rev 20.1.0

* * * * M U L T I P L E R E G R E S S I O N * * * *

Equation Number 3 Dependent Variable.. PRICE

Residuals Statistics:

Selected Cases: CASENO LE 28.00	Min	Max	Mean	Std Dev	N
*PRED	.6952	2.2392	1.0886	.2959	27
*ZPRED	-1.3234	3.8792	.0000	1.0000	27
*SEPRED	.0650	.2184	.0994	.0338	27
*ADJPRED	.6052	1.9628	1.0693	.2618	27
*RESID	-.3548	.6954	.0000	.2611	27
*ZRESID	-1.2829	2.5145	.0009	.9435	27
*SRESID	-1.7132	2.6425	.0252	1.0928	27
*DRESID	-.6667	1.1730	.0187	.3751	27
*SDRESID	-1.7940	3.2960	.0585	1.1846	27
*MAHAL	.4749	15.2555	2.8889	3.5218	27
*COOK D	.0000	2.8055	.1564	.5474	27
*LEVER	.0183	.5869	.1111	.1355	27

Unselected Cases: CASENO GT 28.00	Min	Max	Mean	Std Dev	N
*PRED	1.0609	1.0609	1.0609	.0000	1
*ZPRED	-.0913	-.0913	-.0913	.0000	1
*SEPRED	.0891	.0891	.0891	.0000	1
*ADJPRED	1.0609	1.0609	1.0609	.0000	1
*RESID	-14.3656	-14.3656	-14.3656	.0000	1
*ZRESID	-51.9448	-51.9448	-51.9449	.0000	1
*SRESID	-49.4444	-49.4444	-49.4444	.0000	1
*DRESID	-14.3656	-14.3656	-14.3656	.0000	1
*SDRESID	-49.4444	-49.4444	-49.4444	.0000	1
*MAHAL	1.5725	1.5725	1.5725	.0000	1
*COOK D	.6164	.6164	.6164	.0000	1
*LEVER	.0582	.0582	.0582		

Total Cases = 28

Durbin-Watson Test = 1.00664 (Sig. pos. corr.)

Figure 1 (d) Regression output at step 3

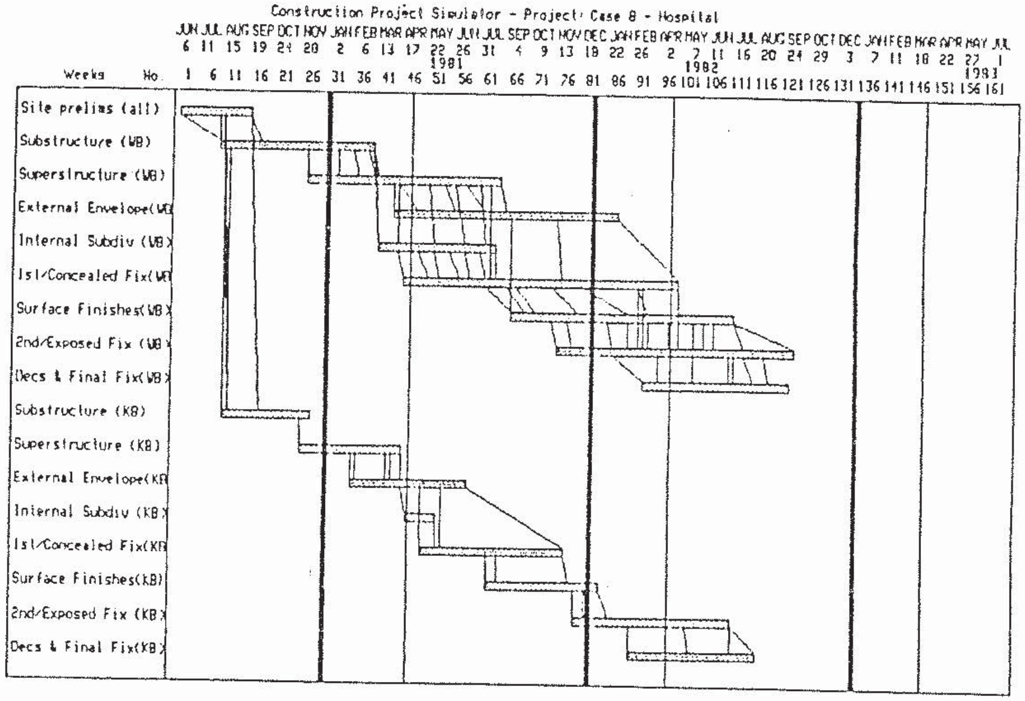


Figure 2(a) Primary Bar Chart.

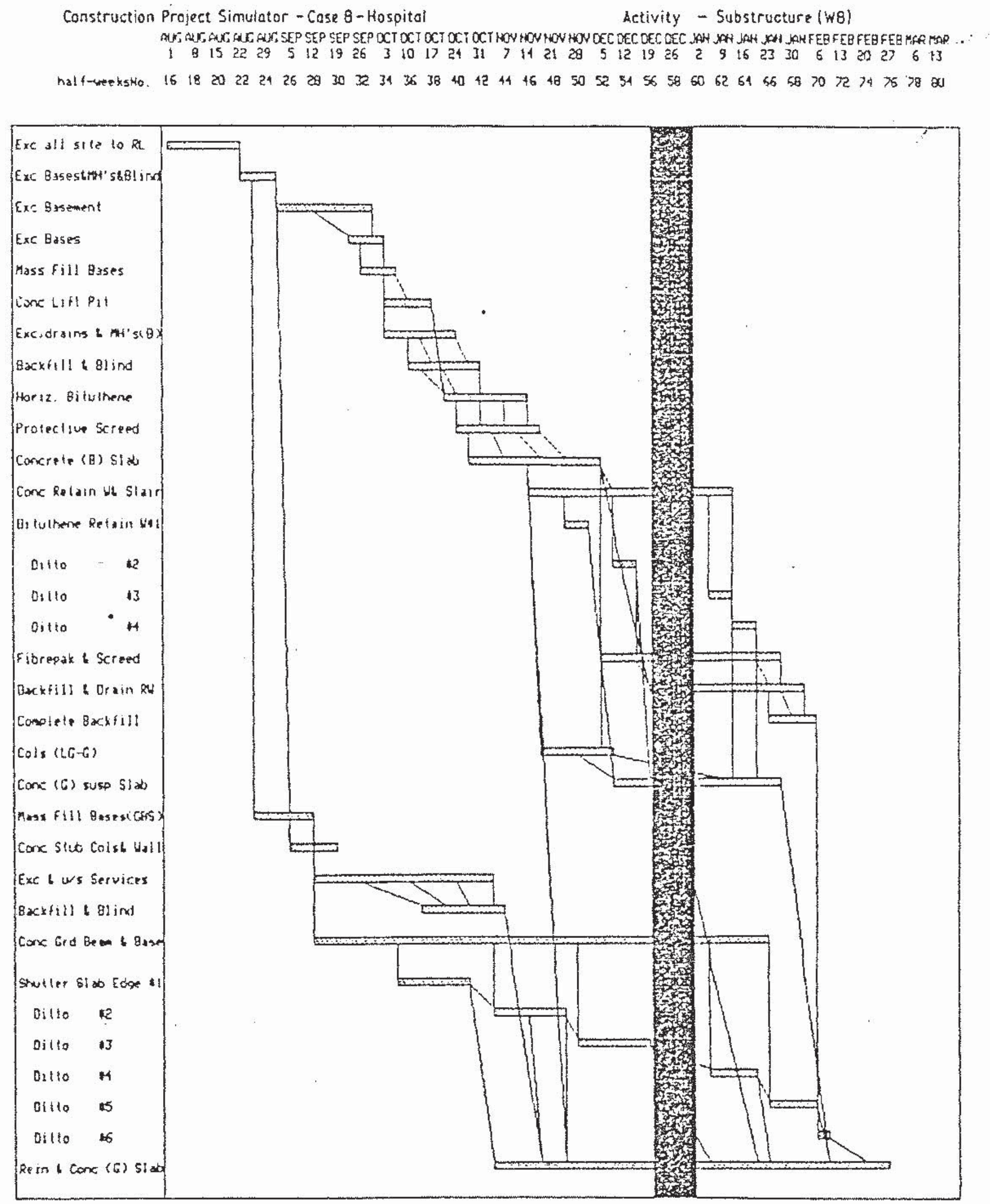


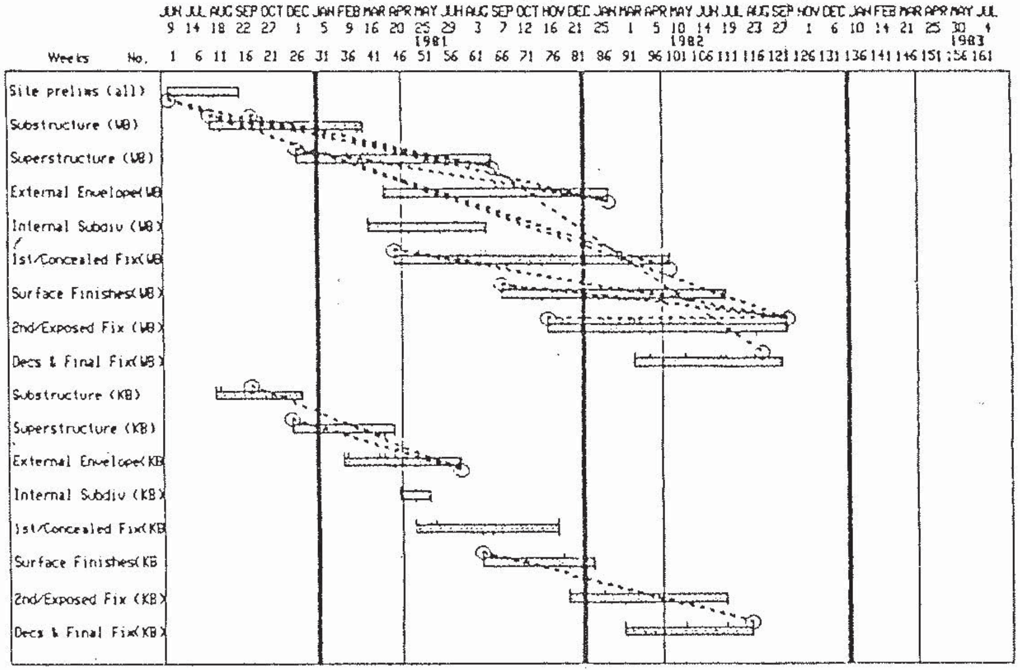
Figure 2(b) - Secondary level bar chart (extended plan of WP 2 in Figure 2(a)).

PWP No. 2 **AMEND COST DATA** Substructure (WM)

Activity	Labour Cost	Material Cost	
Exc all site to RL	<11610 >	<9395 >	
Exc Bases&MH's&Blind	<8645 >	<6600 >	
Exc Basement	<14925 >	<12875 >	
Exc Bases	<6635 >	<5365 >	
Mass Fill Bases	<520 >	<1300 >	
Conc Lift Pit	<1615 >	<3565 >	
Exc.drains & MH's(B)	<7900 >	<11705 >	**LABOUR TOTAL**
Backfill & Blind	<1065 >	<450 >	< 110520>
Horiz. Bituthene	<775 >	<1590 >	
Protective Screed	<440 >	<905 >	**MATERIAL TOTAL**
Concrete (B) Slab	<3500 >	<7735 >	< 160300>
Conc Retain W8 Stair	<4040 >	<8920 >	
Bituthene Retain W#1	<250 >	<455 >	**GRAND TOTAL**
Ditto #2	<245 >	<455 >	< 270820>
Ditto #3	<245 >	<455 >	
Ditto #4	<245 >	<455 >	
Fibrepak & Screed	<1445 >	<2955 >	
Backfill & Drain RW	<12020 >	<10725 >	
Complete Backfill	<3510 >	<2045 >	
Cols (LG-G)	<2150 >	<4755 >	

Figure 2(c) Cost screen

Construction Project Simulator – Project Preliminaries: Case B – Hospital



Construction project simulator – Preliminaries description and costs

Week start		Week end		Description	Labour	Materials
1. Site prelims (all)	(0)	to 2nd/Exposed Fix (WB)	(115)	Staff, hutting, power, clean, other, profit	2480	333 285
2. Site prelims (all)	(0)	to External Envelope (WB)	(81)	Structural foreman 2 No.	380	0
3. Substructure (WB)	(16)	to Superstructure (WB)	(60)	Tower Crane No. 1	655	5150
4. Substructure (WB)	(8)	to 1st/Concealed Fix (WB)	(92)	Assistant engineer + chain boy	315	0
5. Substructure (WB)	(8)	to Superstructure (WB)	(60)	Engineer + assistant	400	0
6. Superstructure (WB)	(60)	to Decs & Final Fix (WB)	(110)	Hoists 2 No.	255	800
7. Superstructure (WB)	(25)	to External Envelope (WB)	(81)	Scaffold	890	0
8. 1st/Concealed Fix (WB)	(42)	to 2nd/Exposed Fix (WB)	(115)	Services coordinator	230	0
9. Surface Finishes (WB)	(62)	to 2nd/Exposed Fix (WB)	(115)	Finishing foremen 3 No.	485	0
10. 2nd/Exposed Fix (WB)	(71)	to 2nd/Exposed Fix (WB)	(115)	Completion agent	205	0
11. Substructure (KB)	(16)	to External Envelope (KB)	(54)	Tower Crane No. 2	660	4650
12. Superstructure (KB)	(24)	to External Envelope (KB)	(54)	Scaffold	155	0
13. Surface Finishes (KB)	(58)	to Decs & Final Fix (KB)	(108)	Finishing foreman	165	0

```
                         CREATE RESOURCE SETS
   TITLE                       COST
23.  <Grdwk bricklyr >    <150      >
24   <Grdwrk plantdvr>    <120      >
25.  <Drainlayer      >   <185      >
26.  <Formwork Carp   >   <125      >
27.  <Steelfixer      >   <120      >
20.  <Concretor       >   <100      >
29.  <Conc. finisher  >   <185      >
30.  <Scaffolder      >   <165      >
31.  <Bricklayer      >   <185      >
32.  <Bricklyr Labour>    <190      >
   RESOURCE SET                        CODE REFERENCES
11. <RC Stairs : TL/P                 > <27>    <26>    <28>    <  >    <  >
12. <RC Stairs : H/ D                 > <27>    <26>    <28>    <  >    <  >
13. <RC Foundations                   > <26>    <27>    <28>    <  >    <  >
14. <Formwork                         > <26>    <  >    <  >    <  >    <  >
15. <Steel fixing 8 concreting        > <27>    <20>    <  >    <  >    <  >
16. <Structural steel erection        > <60>    <  >    <  >    <  >    <  >
17. <Asphalt roofing                  > <35>    <  >    <  >    <  >    <  >
18. <Tiled roofing                    > <36>    <  >    <  >    <  >    <  >
19. <Scaffolding                      > <30>    <  >    <  >    <  >    <  >
20. <Brickwork                        > <31>    <32>    <90>    <  >    <  >
```

Figure 2(e) Resources library screen.

PWP No. 2 **ALLOCATE RESOURCES AND COSTS** Substructure (WB)

ACTIVITY	RESOURCE SET	(TRADES)	(NUMBERS)	(TRADES)	(NUMBERS)	(TRADES)	(NUMBERS)	(TRADES)	(NUMBERS)	(TRADES)	(NUMBERS)	(NUMBERS)	
4. [Exc Bases]	(26)	24	1	91	1	22	1						298
5. [Mass Fill Bases]	(24)	28	3										388
6. [Conc Lift Pit]	(13)	26	2	27	2	28	3						798
7. [Exc, drains & MH's(B)]	(26)	24	1	91	1	22	2						396
8. [Backfill & Blind]	(27)	12	1	13	2								418
9. [Horiz. Bituthene]	(27)	12	1	13	3								548
10. [Protective Screed]	(27)	12	1	13	3								548
11. [Concrete (B) Slab]	(2)	27	4	26	2	28	3						1838
12. [Conc Retain WB Stair]	(8)	27	2	26	2	28	3						798
13. [Bituthene Retain W*ll]	(27)	12	1	13	2								418

RESOURCE SET	CODE REFERENCES				
18. [Tiled roofing	[36]	[ ]	[ ]	[ ]	[ ]
19. [Scaffolding	[38]	[ ]	[ ]	[ ]	[ ]
20. [Brickwork	[31]	[32]	[90]	[ ]	[ ]
21. [Window fixing	[37]	[38]	[ ]	[ ]	[ ]
22. [Plumbing	[43]	[ ]	[ ]	[ ]	[ ]
23. [Mastic pointing	[39]	[ ]	[ ]	[ ]	[ ]
24. [Concreting only	[28]	[ ]	[ ]	[ ]	[ ]
25. [Excavation (large scale]	[24]	[92]	[ ]	[ ]	[ ]
26. [Excavation (small scale]	[24]	[91]	[22]	[ ]	[ ]
27. [General labouring	[12]	[13]	[ ]	[ ]	[ ]

RESOURCE TOTALS

24. Grdwrk plantdvr	(2)
92. Excavator	(1)
91. JCB	(2)
22. Grdwk Labourer	(2)
28. Concretor	(3)
26. Formwork Carp	(4)
27. Steelfixer	(4)
12. Ganger	(1)
13. Labourer	(3)

Figure 2(f) Resources allocation screen.

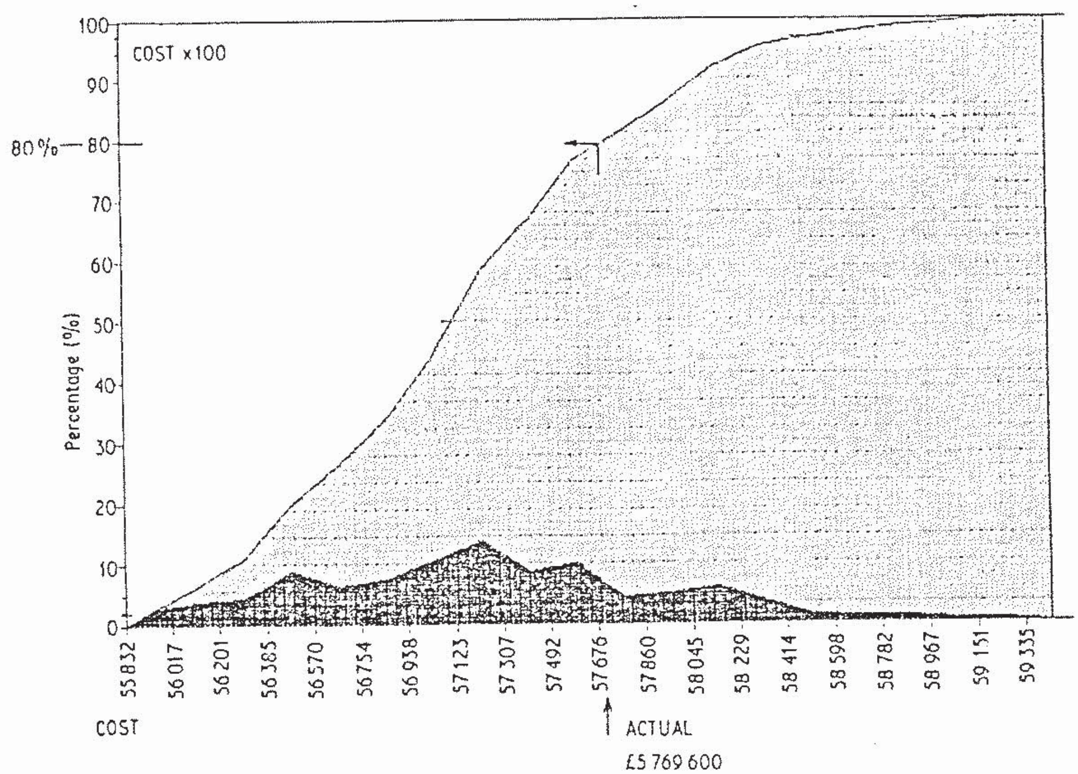


Figure 2(g) Simulation result, including major and 10% inflation, for project final cost.

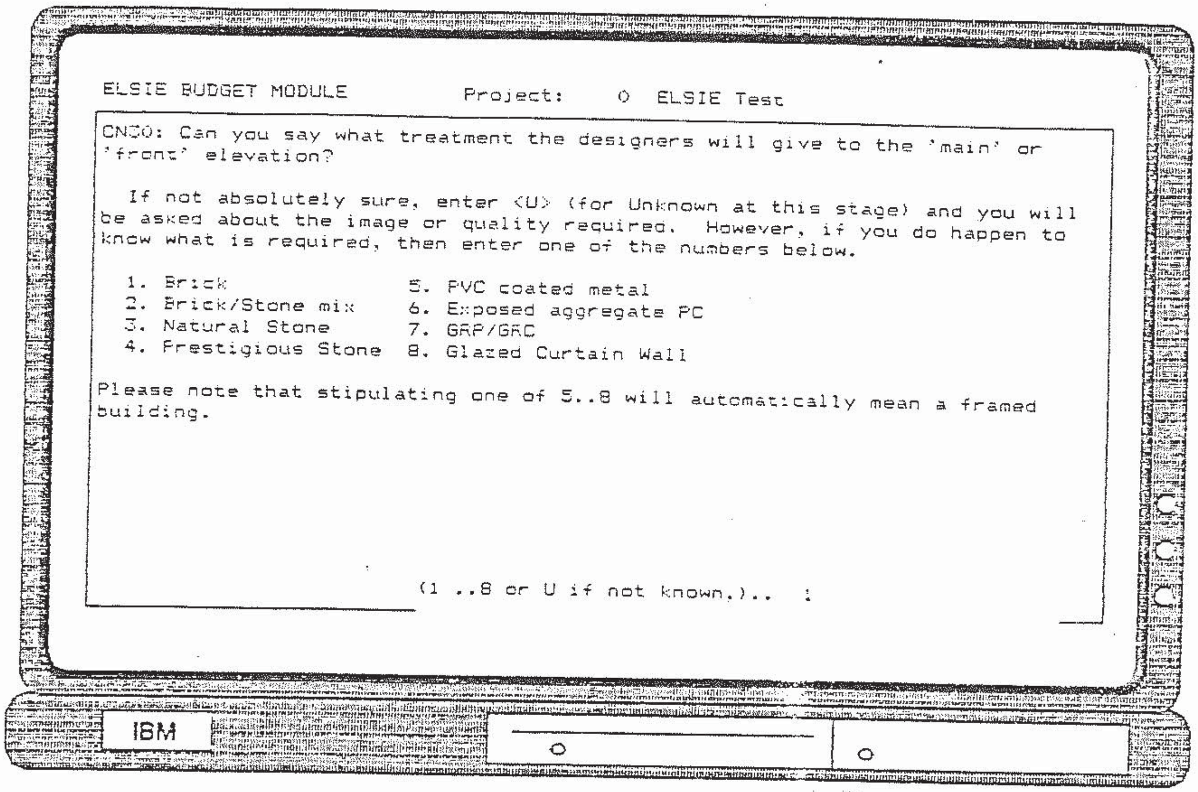


Figure 3(a) Wall appearance input.

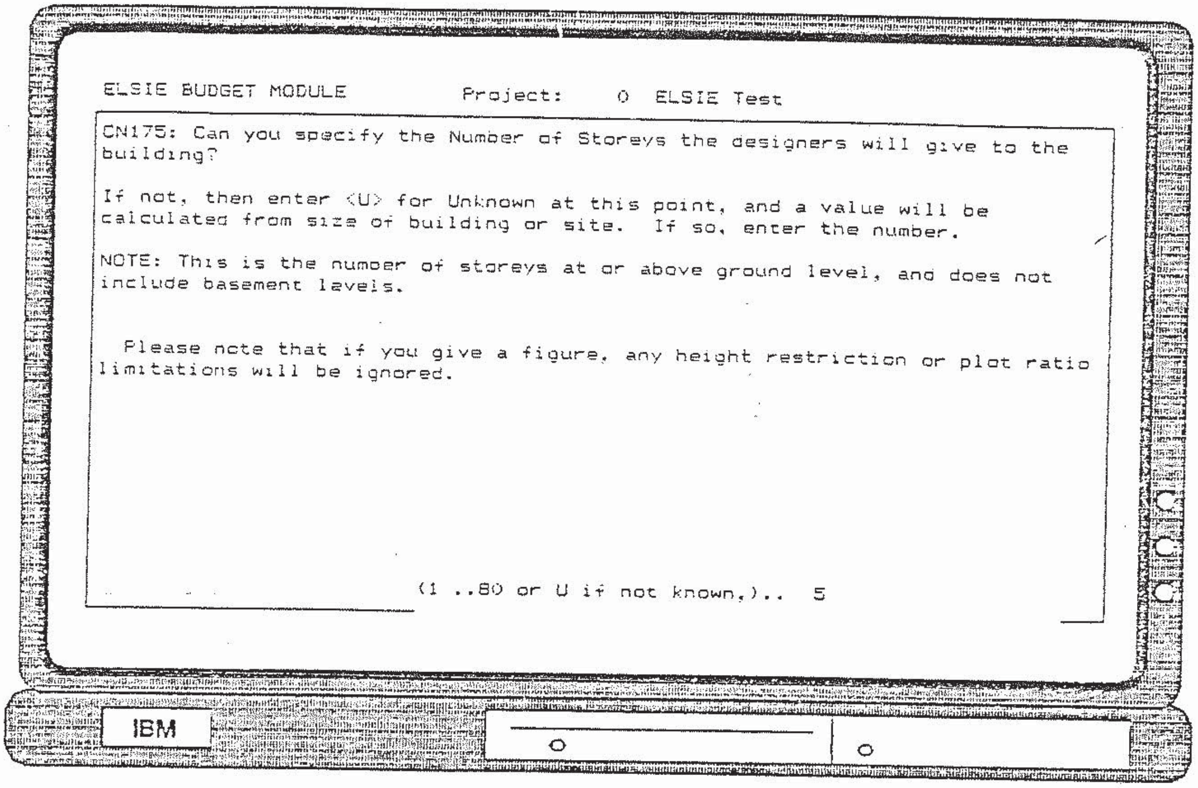


Figure 5(b) Number of storeys input.

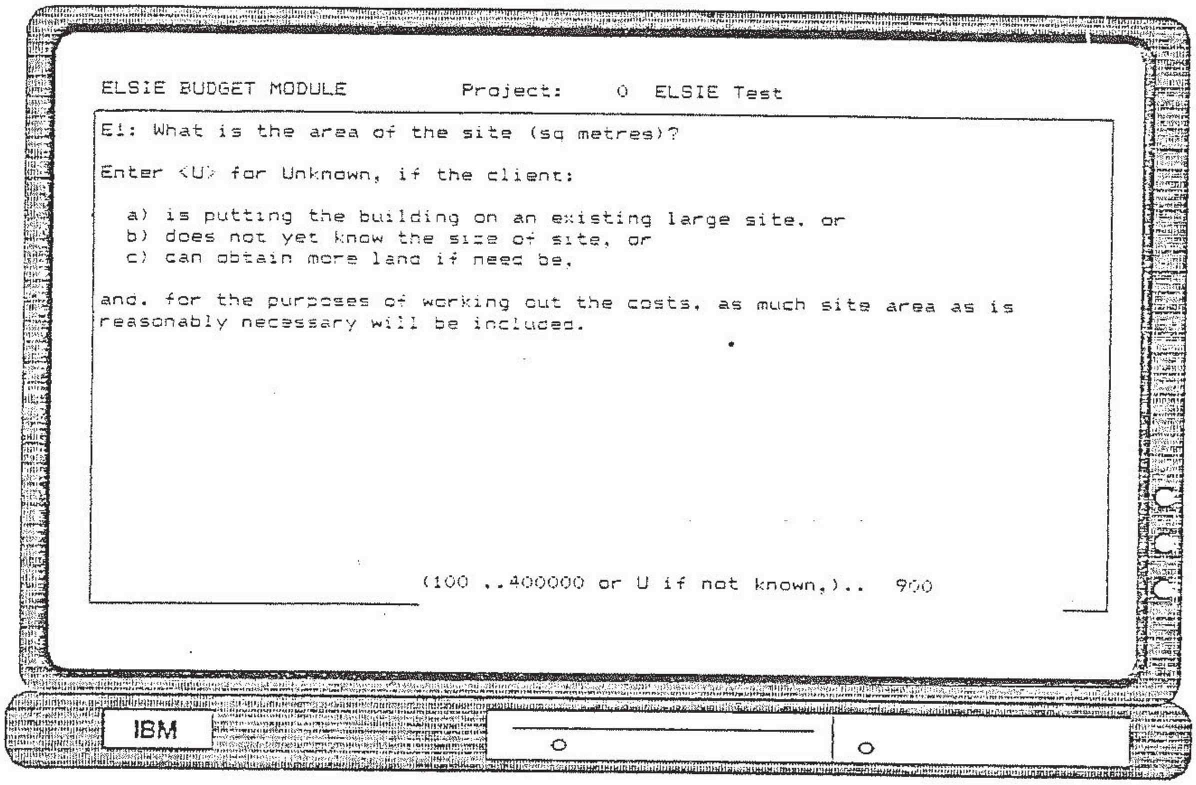
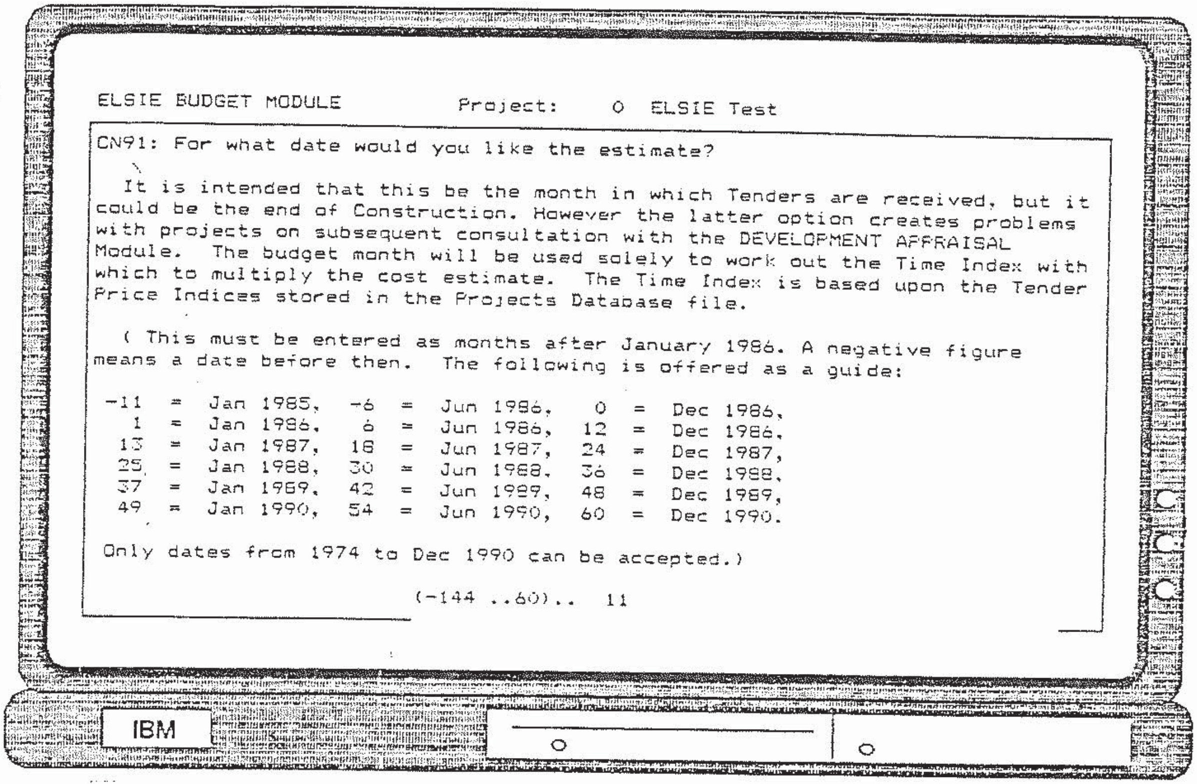


Figure 3(c) Site area input



ELSIE BUDGET MODULE Project: 0 ELSIE Test

CN91: For what date would you like the estimate?

It is intended that this be the month in which Tenders are received, but it could be the end of Construction. However the latter option creates problems with projects on subsequent consultation with the DEVELOPMENT APPRAISAL Module. The budget month will be used solely to work out the Time Index with which to multiply the cost estimate. The Time Index is based upon the Tender Price Indices stored in the Projects Database file.

(This must be entered as months after January 1986. A negative figure means a date before then. The following is offered as a guide:

-11	=	Jan 1985,	-6	=	Jun 1986,	0	=	Dec 1986,
1	=	Jan 1986,	6	=	Jun 1986,	12	=	Dec 1986,
13	=	Jan 1987,	18	=	Jun 1987,	24	=	Dec 1987,
25	=	Jan 1988,	30	=	Jun 1988,	36	=	Dec 1988,
37	=	Jan 1989,	42	=	Jun 1989,	48	=	Dec 1989,
49	=	Jan 1990,	54	=	Jun 1990,	60	=	Dec 1990.

Only dates from 1974 to Dec 1990 can be accepted.)

(-144 ..60).. 11

Figure 3(d) Tender date input.

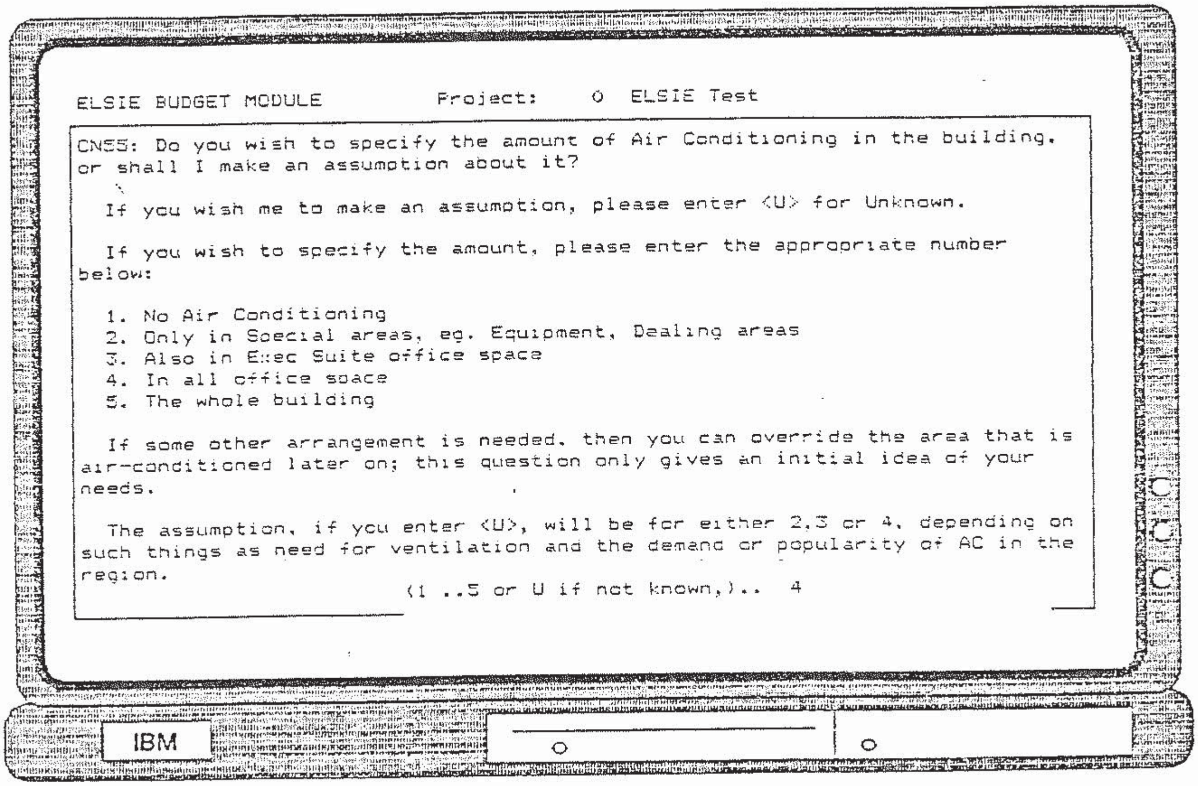


Figure 2(e) Air-conditioning requirements input.

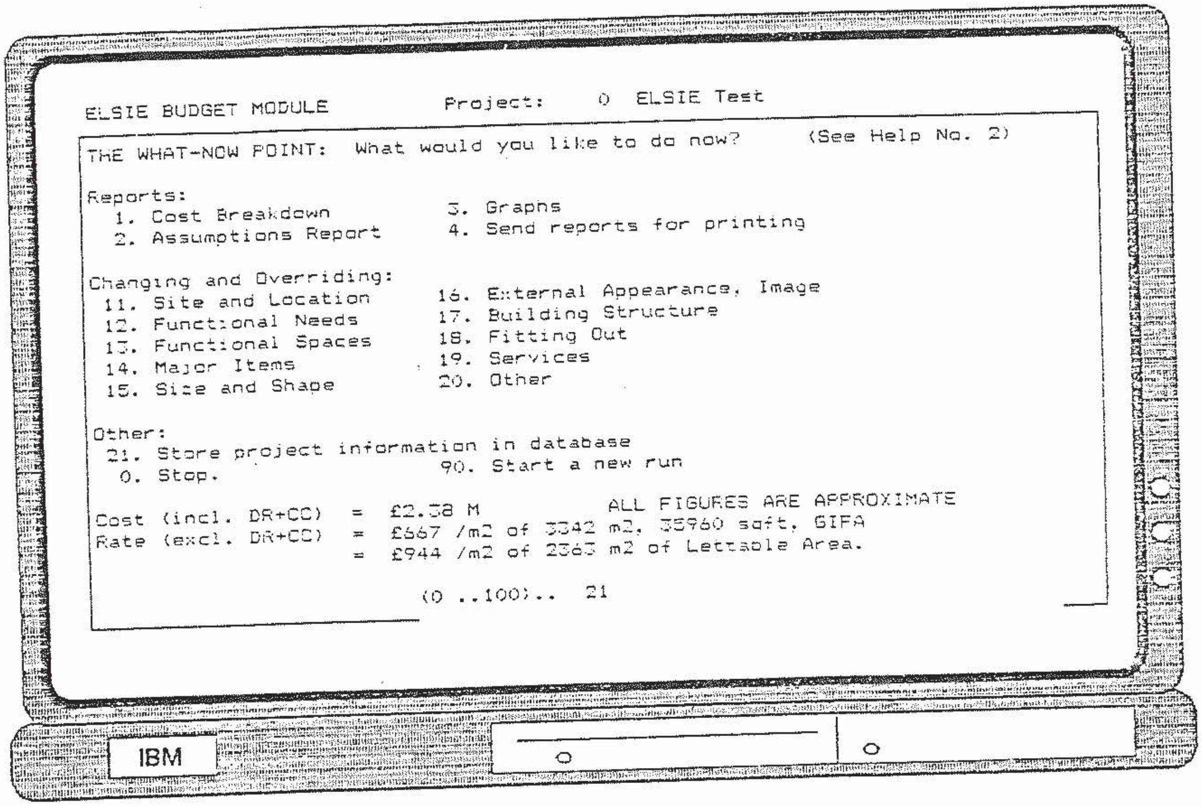


Figure 3(f) The 'What-Now' point.

REPORT ON PROJECT

INPUT INFORMATION:

('X .. Y' below means the information was not asked for.)

No.	Item	Value
1.	type of client	4 Speculative
2.	sector	4 Insurance
3.	office status	1 ..5 Sole Office to Main HQ
4.	office function	1 ..3
10.	size derivation	3
11.	office staff needed	10 ..10000
13.	office space wanted	100 .. 5000000 m2
170.	GIFA wanted	3342 m2
175.	no of storeys	5
14.	max storeys	0 ..80
21.	basement needed?	False
176.	basement area	0 m2
20.	no of cars	5
30.	wall appearance	1: Brick
191.	wall detailing	
32.	image	1 ..7
50.	level of fitting out	2
51.	open plan %	100%
52.	need for large spaces	0
53.	column free ?	False
54.	airtightness	0 ..10
55.	AC where ?	4: Office Space
150.	AC type	5: VAV
70.	region	6 South-East region (excl. London)
71.	type of locality	1 ..5 City centre to Rural/village
72.	neighbour distance	4
73.	noise level	10
179.	site area	900 m2
81.	site known ?	True
82.	site difficulty	7
83.	state of site	2
84.	ground bearing capcty	1 ..4
85.	special foundation ?	True
86.	amount of rock	0
87.	water problems	0
88.	slope	1
89.	can build on rock	Unknown
90.	constrn months	15 months
91.	start month	Nov 1986

Figure 8(a). Explanation report

SIZE DERIVATION

The size of the building (GIFA) was directly specified.
The Height of the building was specified as 5 storeys.
The size of the site was specified directly.

MAIN RESULTS

Approximately, the construction costs will be £2.38 M, including Design Reserve and Construction Contingencies of 6.5 %.

Without these contingencies the cost is approximately £2230468, and gives average rates over the Gross Internal Floor Area of:

£	667 per sq metre over	3342 m2
£	62 per sq foot	35960 sqft

This price is based on Conventional Procurement Path with Competitive Tendering. It includes Preliminaries of around 17 %. These have been increased above a typical of 12% to take into account limitations on the construction process due to Site Constraints.
Location and Tender Price Indices have been included as follows:
Location: 1.00 for South-East region (excl. London)
(1.0 = South East region)
Budget Date: 1.03 for Nov 1986 (1.0 = Jan 1986)
You can alter the Budget Date to be start or end of construction period, or at any other time, to suit your requirements.

ELEMENTAL BREAKDOWN	%	0........50	£/m2	£
1 Substructure	4		25	92346
Basement	0		0	0
2A Frame	8		56	187053
B Upper floors	5		32	107279
C Roof	3		21	69018
D Stairs	3		19	63460
E External Walling	7		44	145512
F Windows + Ext Doors	7		47	158539
G Internal walls/doors	2		11	37341
3 Finishes	8		51	167533
4 Fitting and Furnishings	1		6	20583
5F Heating and Ventilation	17		114	380378
H Electrical	9		59	198607
J Lifts	3		20	67189
M Special Installations	1		4	14411
Other Services and BWIC	4		29	95866
6 External services/works	2		13	44174
7 Prelims	17		113	379180
Total (less contingencies)	100		667	2230468

Figure 8(h) Explanatory report (continued)

SITE USAGE

	Metric	Acres
Landscaped Area	122	0.0
Hard Area (eg. Car Park	110	0.0
	-------	-------
External site area	232	
Building footprint	668	
Site fit	0	0
	-------	-------
Site area	900	0.2
Total Car Parking Area needed	110	0.0
No. of Car Spaces	5	

(The assumed Paved Area does not allow for any paved concourses nor approach roads; if you wish to include these then override the Paved Area.)

SIZE AND SHAPE

The building is reckoned to be simple rectangle in Plan Shape, 5 No. storeys high. The major areas are as follows:

	Metric	Imperial
Ground Floor Area	668	7192
Upper Floor Area	2674	28768
Basement Area	0	0
Atrium Area	0	0
	--------	--------
Gross Internal Floor Area	3342	35960
Average Floor-floor height	3.8	12.6
Width of building	13.6	45
Overall Length	49	163
Wall/Floor ratio	0.72	

The average plan area for upper floors is 668 m2.

EXTERNAL APPEARANCE

The AA Quality and external complexity factors are:

Aesthetic/Amenity Quality
Complexity of Form
of Walls
of Roof

The 'Main' elevation may be something like Brick on frame, with a Window-to-Wall ratio of 0.33.

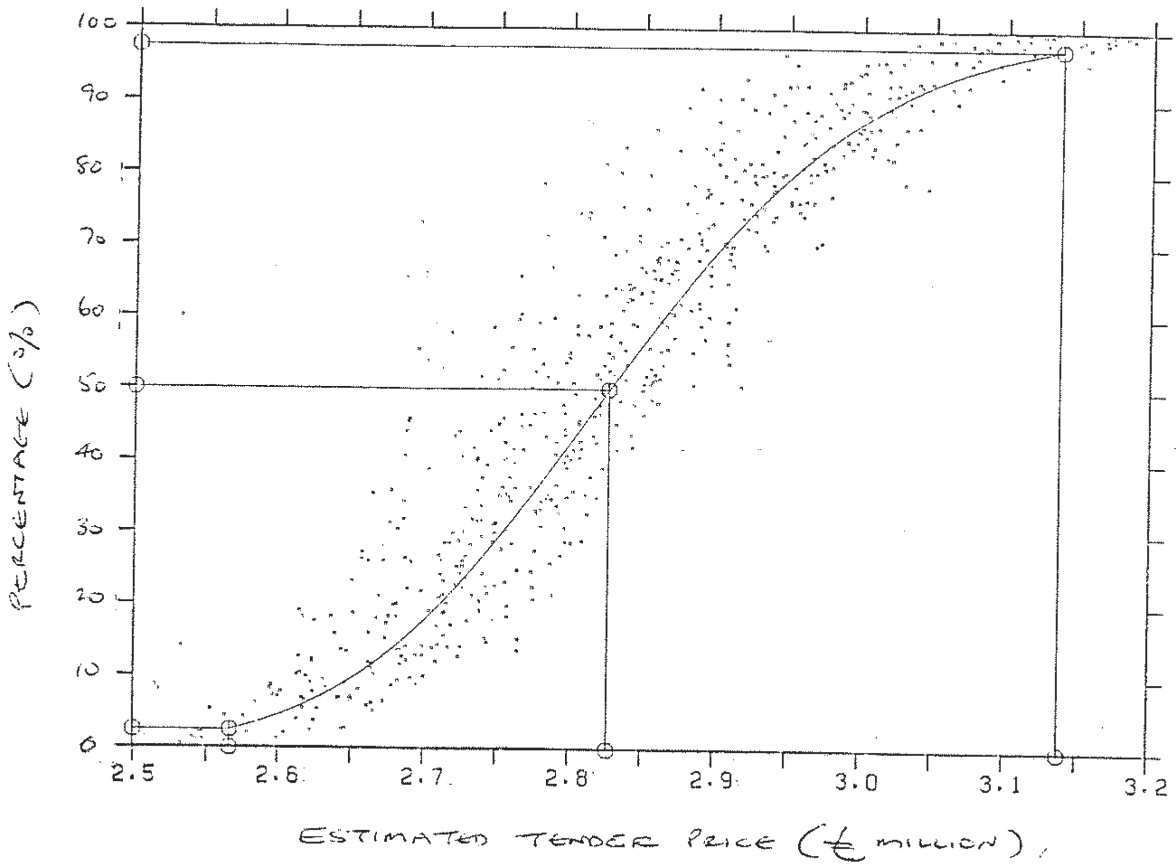


FIGURE 4. Bidding model [illegible] [illegible] results

Estimate Technique	Model	Relevant Contract Type	General Accuracy (cv)	Det/ Prob	Items (i) Number	Items (i) Type	Quantities (q) Derivation	Quantities (q) Det/Prob	Rates (r) Derivation Data Base	Rates (r) Weighting	Rates (r) Current	Rates (r) Quantity trended	Rates (r) Det/ Prob
1 UNIT	$P = qr$	All	25-30%	det	single	any comparable unit, eg tonne steelwork, metre pipeline.	Brief	det	averaged price-cost unit	?	direct	none	det
2 GRAPHICAL	$P = f_r(q)$	Process Plant	15-30%	det	few	ditto	Brief	det	trended price-cost	?	interpolated	objective	det
3 FUNCTIONAL UNIT	$P = qr$	Buildings	25-30%	det	single	ditto eg number of beds number of pupils	Brief	det	averaged/ rule price-cost unit	?	direct	none	det
4 PARAMETRIC	$P = f_r(q_1\ q_2\ q_3\ \ldots)$	Process Plant	15-30%	det	few	process parameters eg capacity, pressure, temperature, materials, cost index	Brief	det	averaged/ rule price-cost parameter	?	direct	none	det
5 EXPONENT	$P_2 = P_1 \frac{q_2}{q_1}^r$	Process Plant	15-30%	det	single	size of plant or equipment eg capacity	Brief	det	averaged/ rule price-cost exponent	crude subjective	direct/ interpolated	objective	det
6 FACTOR	$P = \sum_j^m fact_j \sum_i q_i r_i$ a) $m = 1$ (Lang method) b) $m > 1$, $fact_1 \neq fact_2$ etc (Hand method) c) $fact_j = U(\alpha_j, B_j)$ (Chiltern method)	Process Plant	10-15%	det	few	any	Brief/ measure	det	averaged/ rule price-cost	?	factored	a)none b)none c)subjective	det det det
7 COMPARATIVE	$P_2 = P_1 + \sum_i (p_{2i} - p_{1i})$	All	25%-30%	det	few	depends on differences	Brief	det	price-cost items	crude subjective	adjusted		det
8 INTERPOLATION	$P = qr$	Buildings	25%-30%	det	single	gross floor area	Brief	det	price/m^2	crude subjective	interpolated	-	det
9 CONFERENCE	$P = f(P_1\ P_2\ \ldots)$	Process Plant	?	det	any	any	Brief/ measure	det	-	-	negotiated	-	det
10 FLOOR AREA	$P = qr$	Buildings	20-30%	det	single	gross floor area	Brief/ measure	det	averaged price/m^2	crude subjective	direct	subjective	det
11 CUBE	$P = qr$	Buildings	20-45% (based on 86 cases)	det	single	volume	measure	det	averaged price/m^3	crude subjective	direct	none	det
12 STOREY ENCLOSURE	$P = qr$	Buildings	15-30% (based on 86 cases)	det	single	floor/wall area/ basement/roof	measure	det	averaged price/SE unit	crude subjective	direct	none	det
13 BQ PRICING a)(Conventional)	$P = \sum q_i r_i$	Construction	10-20% (5-8% for builders)	det	v many	SMM	measure	det	a)averaged BQ's	crude subjective	direct	subjective	det
b)(B Fine)	$P = \sum q_i r_i$	Buildings	15-20%	prob					b)$r_i \sim U(r_{min}, r_{max})$		direct	subjective	prob
14 SIG. ITEMS	$P = \sum q_i r_i$	PSA Buildings	10-20%	det	medium	SMM	measure	det	averaged BQ's/rule	crude subjective	direct	objective	det
15 APPROXIMATE QUANTITIES a)(Conventional)	$P = \sum q_i r_i$	Construction	15-25%	det	medium/ few	SMM combined	a)measure	det	a)averaged BQ/price book	crude subjective	composited	subjective	det
b)(Gleeds)	$P = \sum q_i r_i$	Buildings	15-25%	det			b)Brief/ measure	det	b)averaged BQ/price book	crude subjective	composited	subjective	det
c)(Gilmore)	$P = \sum q_i r_i$	Buildings	15-25%	det			c)Brief/ measure	det	c)averaged BQ/price book	crude subjective	composited	subjective	det
d)(Ross 1)	$P = \sum q_i r_i$	Buildings	25% (based on 17 cases)	det/ prob			d)measure	det	d)50 BQ's averaged	none	direct	none	det
e)(Ross 2)	$P = \sum q_i r_i$	Buildings	50% (based on 17 cases)	det/ prob			e)measure	det	e)50 BQ's $r_i = a + bq_i + e$, $e \sim N(o, \sigma^2)$	none	mathematically	objective	prob
f)(Ross 3)	$P = \sum p_i$ ($p_i = a + bq_i + e$, $e = N(0, \sigma^2)$	Buildings	30% (based on 17 cases)	det/ prob			f)measure	det	f)50 BQ's	none	mathematically	objective	prob
16 ELEMENTAL	$P = \sum q_i r_i$	Buildings	20-25%	det	medium	BCIS/CI sfb entities	Brief/ measure	det	averaged BQ's/ BCIS/m^2	crude subjective	composited/ direct	subjective	det
17 CPU	$P = \sum q_i r_i$	Buildings	20-25%	det	medium	Similar	Brief/ Measure	det	averaged BQ	crude subjective	composited	subjective	det
18 ELSIE	$P^2 = \sum q_i r_i$	Offices		det	medium	DBE	Brief	det	averaged BQ/ rule	none	direct	none	det
19 NORMS (schedule)	$P^2 = \sum q_i r_i$	Buildings	10-20%	det	v many	SMM type eg PSA schedule	measure	det	cost based rules	none	direct	none	det
20 REGRESSION	$P = a + \sum q_i b_i + e$ $e = N(0, \sigma^2)$	All	15-25%	det/ prob	few	usually contract characteristics eg floor area, number of storeys	Brief	det	any	crude subjective	mathematically	objective	prob
21 LU QIAN	$P = \sum q_i r_i$	Buildings	?	det	few	usually contract characteristics eg floor area, number of storeys	Brief	det	any	mathematically	mathematically	none	det
22 RESOURCE (Activity, operational, scheduling)	$P = \sum q_i r_i$	All	5-8% (builders)	det	v many	resource eg man hours, materials, plant	production plan	det	average costs	crude subjective	direct/ analytical	subjective/ objective	det
23 PERT-COST	$P = \sum p_i$ where $p_i \sim N.(q_i r_i, \sigma 2_i)$	All	N/A	prob	varies	usually time resources eg man hours	production plan	prob (time)	-	-	-	-	-
24 CPS	$P = \sum t_i r_i + \sum m_j r_j$ $t_i \sim F(\mu_i, \sigma^2{}_i)$	Buildings	6.5% (based on 4 cases)	prob	usually few	resource eg manhours, materials, plant	production plan	prob (time)	average cost	crude subjective	direct	none	det
25 RISK ESTIMATING	$P = \sum q_i r_i$	Construction	N/A	prob	usually few	any	any	det	theoretical frequency distributions of cost	crude subjective	random selection $r_i \sim F(\mu_i, \sigma^2)$	none	prob
26 HOMOGENISED ESTIMATING (BCIS on line) (BICEP etc)	$P = \sum q_i r_i$	Building	N/A	det	any	any	any	det	average BQ	aided subjective	direct	objective	det

Notes

F ( ) some (unspecified) probability function
N ( ) normal probability function
U ( ) uniform probability function

Table 1: Resume of estimating techniques

Contract number CASENO	Standardised contract sum* CONSUM	Gross floor area (m2) GFA	Number of storeys STOREY	Air-conditioning (0 = no) (1= yes) ACOND	Contract period (months) PERIOD	Number of tenders BIDDERS
1	1085.95	452	2	0	8	6
2	5042.91	1601	7	0	11	8
3	2516.59	931	3	1	11	7
4	18290.60	6701	7	1	17	6
5	3195.81	219	3	0	12	1
6	8894.68	3600	6	0	15	9
7	932.06	490	2	0	7	6
8	979.93	415	1	0	8	8
9	1684.94	504	3	0	9	6
10	1896.39	320	2	0	7	7
11	3789.05	372	2	0	6	7
12	2445.12	837	2	0	9	4
13	1501.91	491	3	0	6	6
14	1114.31	496	1	0	6	6
15	943.48	430	2	0	99	99
16	3670.98	1368	4	0	12	4
17	1094.75	469	2	0	6	4
18	4584.87	1260	2	0	8	5
19	10942.28	2994	8	1	15	1
20	760.29	312	2	0	6	6
21	3002.67	1225	2	0	9	7
22	2720.44	1230	2	0	10	8
23	58365.39	23089	7	1	20	7
24	11323.40	4273	4	1	20	7
25	37 357.91	11300	5	1	18	6
26	46309.12	14430	3	1	30	6
27	1704.17	437	3	1	10	9
28	6792.04	2761	5	1	12	6

Contract sum ÷ tender price index ÷ location factor e.g. contract number 2 contract sum = 1514385 ÷ in London (location factor 1.30) on Sept 1987 (TPI 231). Standardised contract sum = 514385 + 231 ÷ 1.30 = 5042.91

Table 2: Regression data

	Regression coefficient	Constant	Forecast	95% confidence limits for new forecast	cv
Step 1					
BIDDERS	–0.4355	1.8298	1.0495	±0.6844	38.31
Step 2					
BIDDERS	–0.4251				
GFA	–0.0395	2.0929	0.9873	±0.7090	39.10
Step 3					
BIDDERS	– 0.2987				
GFA	– 0.3071				
PERIOD	0.8814	1.7040	1.0609	±0.6011	32.05
Step 4					
BIDDERS	–0.2665				
GFA	–0.3293				
PERIOD	0.8480				
STOREY	0.1032	1.7887	1.0634	±0.6075	33.99
Step 5					
BIDDERS	–0.2646				
GFA	–0.3253				
PERIOD	0.8848				
STOREY	0.1108				
ACOND	–0.0667	1.6847	1.1166	±0.6836	35.06

Table 3: Stepwise regression results

PROJ	BID	DDR	BID	DDR	BID	DDR	BID	DDR	BID	DDR	BID	DDR	BID	DDR	BID	DDR	BID	DDR
1	1454515.	150	1514865.	55	1475398.	304	1460775.	134	1447867.	154	1457977.	73						
2	535608.	304	502042.	291	529744.	154	516376.	157										
3	1333142.	75	1331156.	217	1366863.	304	1266892.	281	1276787.	115	1277652.	93	1865545.	360				
4	696743.	304	696972.	292	701062.	237	637815.	79	697826.	361	637815.	280						
5	404110.	55	422297.	304	413224.	97	389196.	117	417489.	362	389848.	157						
6	2116877.	134	2169966.	99	2187991.	293	2161120.	304	2198655.	221	2296108.	137	2165611.	8	2153344.	117	2133608.	294
7	3065742.	304	3119689.	150	3141641.	170	3153800.	134	3249927.	191	3269768.	55	3335993.	107				
8	7925257.	221	7351929.	304	7374650.	247	6900000.	20										
9	871520.	118	899935.	137	902378.	304	914393.	291	950737.	83	996483.	221						
10	1063337.	304	1154023.	251	1102272.	173	1079637.	201										
11	1759614.	154	1792123.	281	1838532.	157	1910066.	170	1947733.	304	1784215.	308						
12	1126816.	304	1146398.	201	1169795.	154	1227296.	24	1312527.	280	1399472.	221						
13	698005.	304	625501.	268	630288.	308	666545.	55										
14	588810.	364	584833.	365	639229.	79	646341.	145	682802.	304	691474.	154						
15	1429218.	303	1493849.	291	1511033.	304	1521628.	12	1526377.	366	1717715.	55						
16	842319.	6	870894.	304	883617.	185												
17	284947.	367	292692.	356	294694.	368	303700.	152	307282.	85	313203.	134	315727.	369	333597.	118	334353.	370
	348969.	304																
18	461444.	150	483862.	304	482241.	308	447021.	55	493417.	154	455480.	311						
19	2858191.	280	2947007.	371	2950723.	134	2999999.	304	3093587.	60	3099528.	6	3278229.	266	3325198.	170	3333793.	55
20	7831865.	276	7837276.	304	7859122.	256	7904172.	55	8047230.	152	8145323.	293	8279564.	117	8657685.	134		
21	3971051.	55	3854074.	304	4724785.	372	3955007.	154	3944772.	373	3731543.	79	4001188.	237				
22	573485.	292	596737.	291	597730.	134	613528.	201	615015.	304	621223.	170						
23	1610942.	304	1623447.	163	1646286.	173	1663742.	268	1700000.	152								
24	1196036.	64	1199328.	374	1206837.	187	1226509.	304	1262082.	291	1271000.	152	1295954.	170	1302161.	254		
25	2636397.	137	2654728.	150	2673906.	187	2685127.	55	2762123.	304	2845567.	152						
26	469663.	24	476784.	268	485870.	286	486485.	55	504026.	122	529468.	263	540814.	304				
27	1526553.	201	1533719.	152	1698797.	148	1076612.	304										
28	2106139.	201	2175928.	304	2210065.	308	2223710.	280	2255246.	221	2296623.	137	2331830.	266				
29	499888.	102	559596.	55	592026.	217	602042.	170	608957.	304	619065.	134						
30	2639525.	304	2842407.	308	2874130.	280	2861665.	55	2736300.	152	2770720.	256						
31	732572.	304	599429.	365	623906.	145	691759.	79	744332.	154	607065.	364						
32	546641.	134	539565.	268	608242.	55	538382.	24	599934.	170	559351.	304						
33	792966.	221	811788.	99	819971.	308	847621.	55	847892.	137	853793.	304						
34	2085151.	152	2130217.	107	2150583.	280	2203956.	115	2219653.	137	2241687.	154	2325900.	304				
35	821617.	268	844579.	115	848459.	303	871927.	304	872215.	106	935765.	375						
36	792474.	304	747374.	24	778559.	217	743788.	252	808345.	268	835465.	170						
37	7279854.	304	7650271.	60	7029448.	308	6631664.	150	7089879.	193	7230120.	170	6986341.	247	7143710.	191	6794553.	266
38	592096.	304	573997.	150	518613.	217	508985.	121	544480.	376								
39	538600.	348	567031.	377	621365.	378	699839.	268	825451.	72	991468.	190	1001254.	304				
40	2087946.	154	2104017.	276	2183122.	186	2205359.	304	2212382.	280	2267987.	112	2332476.	221	2400000.	294		
41	1503739.	247	1536654.	24	1576905.	304	1583595.	154	1616432.	294	1704995.	157						
42	3624453.	191	3694803.	221	3732133.	304	3751115.	170	3773967.	193	3866339.	55	3922937.	134	4122448.	281		
43	629164.	157	695284.	173	723315.	311	729305.	266	743578.	304	768189.	379						
44	2252833.	304	2264310.	24	2274380.	112	2323385.	191	2384494.	55								
45	1202916.	163	1268733.	55	1291365.	221	1294986.	304										
46	2968891.	217	2772626.	280	2822857.	186	2972189.	134	2821600.	276	2857275.	304	2793000.	221				
47	1398400.	286	1401500.	152	1427140.	237	1436804.	304	1453070.	301	1511643.	55	1591986.	371	1665760.	83		
48	698161.	294	709676.	24	758565.	291	789355.	304	797926.	134	842684.	55	751677.	252				
49	248733.	31	251007.	291	251415.	252	261286.	380	264933.	304								
50	358840.	293	362370.	217	386983.	304	421797.	381	456272.	154								
51	527692.	317	570874.	311	588854.	75	609221.	173	636451.	308	694297.	304						