DOCUMENT RESUME

ED 073 537 EA 004 892

AUTHOR	Picardi, E. Alfred
TITLE	Closing the Credibility Gap in Construction Cost Estimating.
PUB DATE	72
NOTE	13p.; Paper presented at INBEX Annual Meeting (3rd, Louisville, Kentucky, October 31-November 2, 1972)
EDRS PRICE	MF-$0.65 HC-$3.29
DESCRIPTORS	*Computer Programs; *Construction Costs; Construction Industry; *Estimated Costs; Input Output Analysis; *Models; Prediction; Probability; *Simulation; Speeches
IDENTIFIERS	Monte Carlo Simulation

ABSTRACT

The construction cost estimate, often expressed as an absolute cost, leads to misunderstanding between client, designers, and builders. If estimates are to be used as adequate cost indicators, their probabilistic nature must be recognized and they must be expressed not as absolute numbers but in terms of a number with some indication of the magnitude of the risk that that number may be expected to change by some stated amount. In order to develop this new type of estimate, a model is constructed that assigns a probability distribution to each cost component of the project and ends up with a probability distribution of the total cost of the project rather than a single answer. This bracketing of the cost provides a better in-depth look at the current knowledge of the project cost than does the traditional estimate. (Author)

CLOSING THE CREDIBILITY GAP in CONSTRUCTION COST ESTIMATING

by: E. Alfred Picardi, Vice President
The Perkins & Will Corporation

1972

CLOSING THE CREDIBILITY GAP IN CONSTRUCTION COST ESTIMATING

One of the most misleading elements in the construction industry today is the Construction Cost Estimate. Estimates are made on practically all construction projects from the time of their inception, periodically until the final general contract bidding. In some instances, where there is no general contract under the so-called Construction Management System, the Cost Control procedures involve further estimating during the actual buying out and construction phases until the day when all financial commitments for purchase of the completed work have been made.

The word "estimate" implies the judgment of value, the assessing of the monetary worth by an expert, the "educated guessing" of cost. All of these definitions indicate the probabilistic nature of the word, yet outside the industry and in many instances within it, the estimate has come to denote absolute cost at any point in time from inception to completion of the work. It most often is expressed as an absolute cost. This often leads to misunderstanding between client, designers, and builders. In many instances, it results in either unnecessary modification of the scope of a project or failure to modify scope in time to meet an unchangeable budget.

It seems that if estimates are to be used as adequate cost indicators and even cost control tools, their probabilistic nature must be recognized and they must be expressed not as absolute numbers but in terms of a number with some indication of the magnitude of the risk that that number may be expected to change by some stated amount. For example, we may state that at this point in time during the development of a design for a project, we expect its cost to be "X" dollars and there is an 85% probability that actual cost will not be less than "Y" dollars or more than "Z" dollars.

This bracketing of the cost, the determination of the probability of the cost falling within the bracket, the variation of the expected cost from the mode and the "tightness" of the bracket, provide a far better "in-depth" look at our current knowledge of the expected project cost than the traditional estimate which would have been an absolute value with, at best, a contingency factor having no really good relation to the numerous variables and conditions which will influence the final cost.

In order to develop this new type of estimate, it is necessary to construct a probabilistic cost model of the project. The model assigns a probability distribution to each cost component of the project and ends up with a probability distribution for the total cost of the project rather than a single answer.

Each cost component in the model is assumed to have a shifted, lognormal probability distribution. This assumption is based on intuition and empirical observations: costs are always greater than zero, costs are more likely to increase than decrease, and a three-parameter lognormal distribution offers great flexibility in fitting data to it. Figure 1 shows an example of a lognormal distribution and its three associated parameters.

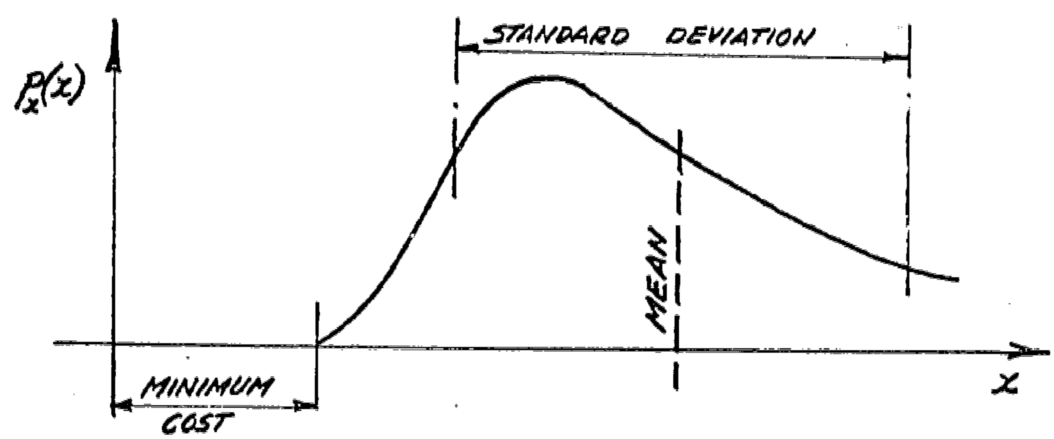

FIGURE 1

After a distribution for each cost component is assumed, the resulting total building cost distribution is determined by use of a Monte Carlo simulation which approximates the total distribution through a series of experiments. It involves simulating the cost of a large number of identical buildings. This is accomplished by generating a series of random numbers, on a computer, which are transformed into the cost for each of the items in the estimate. These are then summed to arrive at the total cost for building number one. The process is then repeated many times until the total cost distribution is derived.

The choice of the low, high and expected prices is a subjective process. The simulation and the resulting cost distribution will, at best, be as "good" as these parameters. This part of the process should be done by experienced estimators. The expected price should be their best guess based on current unit price information, forecasts for the future, etc. The high and low prices should be the prices which they feel there is only one chance in ten of being "higher than" or "less than."

Expected cost might correspond to the mode, median, or the mean, depending on the individual's interpretation. As is shown in Figure 2, in the lognormal distribution, the mode is the value that occurs the most, the median is the value that divides the cost distribution into two equal areas; i.e., $F_X (x_{median}) = 0.50$, and the mean is the "center of gravity" of the distribution. This program assumes the expected price corresponds to the mode. The median or mean could be used, but within the accuracy or confidence of this model, it would have very little effect on the outcome.

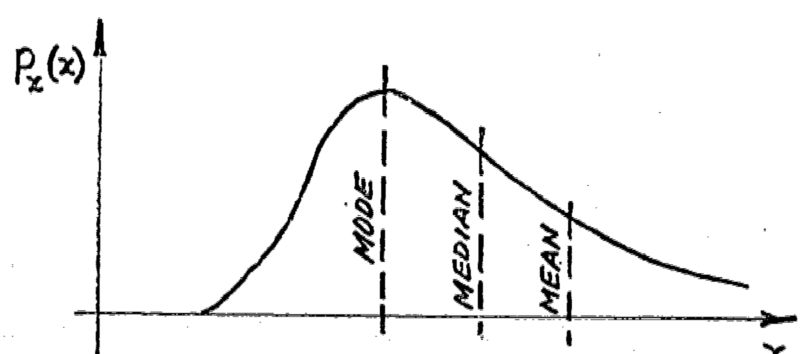

FIGURE 2

Essentially the program consists of a main program which takes the input and performs the Monte Carlo simulation, computes the sample statistics, and sets up the cost intervals for the histogram and cumulative distribution, and two subroutines which output a histogram and a cumulative distribution respectively.

It is necessary to specify several variables for the program. These include:

1. Number of cost components in the model. The program will accept up to 500 components.

2. The number of iterations in the Monte Carlo Simulation. The "accuracy" of the final cost distribution is a function of the number of iterations. It is believed that 1000 is a satisfactory number -- it is a large number that can be handled relatively easily on a computer.

3. The number of intervals in the histogram. The readability of a histogram is very sensitive to the number of intervals in it. An empirical equation that has been found to give good results is:

$$k = 1 + 3 \log_{10} n,$$

where k is the number of intervals and n is the number of observations of the sample. For the 1000 iterations that are recommended for this program, k = 11.

4. The number of intervals in the cumulative distribution. The number of intervals in the cumulative distribution will be related to the desired acxuracy in reading the distribution. This program will produce a cumulative distribution with a maximum of thirty intervals.

5. An additional set of variables. The values of an array which is used within the program to compute the parameters of the lognormal distribution.

The following is an example of the use of the program at three points during the design of a hospital project. First, in the conceptual stages when all that was known was the expected square foot area and typical, expected, high and low costs/square foot (Run No. 1). The second estimate (Run No. 2) was made at the end of design development when the scope was fully defined and expected high and low costs could be applied to the various broad trade categories. Finally, (Run No. 3) was made near the completion of Working Drawings when a detailed quantity survey and application of high, expected and low unit costs could be applied to each of the 23 items in the general estimate. The final estimate (Run No. 4) was made just prior to issuing the drawings for General Contract bidding. The computer output sheets for each run are attached to show the format of the output.

APEC is presently exploring ways in which this program may be made available to members and would appreciate indications of interest.

COST PROBABILITY MODEL

BUILDING TYPE: HOSPITAL

TOTAL SQUARE FOOTAGE: 160,000 SQUARE FEET

RUN #1: "COST PER SQUARE FOOT"

INPUT:	LOW	EXPECTED	HIGH
Cost sq/ft.:	$54	$60	$68
	$8,650,000	$9,620,000	$10,880,000

OUTPUT:

PROBABILITY	LOW	EXPECTED	HIGH
100%	$7,143,129	$9,620,000	$12,472,937
83%	$8,596,714	$9,620,000	$11,019,353

RUN #2: "PRELIMINARY ESTIMATE"

INPUT:	LOW	EXPECTED	HIGH
General	$ 4,550,000	$ 5,150,000	$ 5,820,000
Plumbing	1,260,000	1,420,000	1,610,000
HVAC	1,520,000	1,740,000	2,080,000
Electrical	1,260,000	1,420,000	1,610,000
	$ 8,590,000	$ 9,730,000	$11,020,000

OUTPUT:

PROBABILITY	LOW	EXPECTED	HIGH
100%	$ 8,306,709	$ 9,730,000	$11,375,159
88%	8,920,400	9,730,000	10,761,469

RUN #3: "PARTIAL ESTIMATE - GENERAL"

INPUT:	LOW	EXPECTED	HIGH
Excavation	$ 182,000	$ 200,000	$ 220,000
Concrete	601,000	660,000	726,000
Masonry	305,000	335,000	369,000
Millwork	191,000	210,000	231,000
Finish Hardware	82,000	90,000	99,000
Structural Steel	943,000	1,036,000	1,140,000
Metal Deck	146,000	160,000	176,000
Misc. Iron	72,000	79,000	87,000
Windows	50,000	55,000	61,000
Curtain Wall	109,000	120,000	132,000
Doors	97,000	106,000	117,000
Caulking	14,000	15,000	17,000
Roofing	37,000	40,000	44,000
Ceramic Work	32,000	35,000	39,000
Flooring	117,000	129,000	142,000
Plaster	819,000	900,000	990,000
Painting	137,000	150,000	165,000
Dampproofing	13,000	14,000	16,000
Toilet Accessories	38,000	42,000	46,000
Casework	139,000	153,000	168,000
Specialty Material	189,000	208,000	229,000
Elevators	159,000	175,000	193,000
Hospital Equipment	227,000	249,000	274,000
	$4,699,000	$5,178,000	$5,681,000

OUTPUT:

PROBABILITY	LOW	EXPECTED	HIGH
94%	$4,936,496	$5,178,000	$5,422,848

RUN #4: "FINAL ESTIMATE"

INPUT:	LOW	EXPECTED	HIGH
General	$4,930,000	$5,180,000	$ 5,420,000
Plumbing	1,370,000	1,430,000	1,500,000
HVAC	1,680,000	1,760,000	1,850,000
Electrical	1,360,000	1,420,000	1,490,000
	$9,340,000	$9,770,000	$10,260,000

OUTPUT:

PROBABILITY	LOW	EXPECTED	HIGH
100%	$9,210,639	$9,770,000	$10,412,409
83%	9,450,994	9,770,000	10,172,055

SUMMARY

PROBABILITY OF 100%:	LOW	HIGH	RANGE
Initial Estimate	$7,143,129	$12,472,937	$5,329,808
Preliminary Estimate	8,306,709	11,375,159	3,068,450
Final Estimate	9,210,639	10,412,409	1,201,770

PROBABILITY OF 83%:	LOW	HIGH	RANGE
Initial Estimate	$8,596,714	$11,019,353	$2,422,639
Preliminary Estimate	8,920,400	10,761,469	1,841,069
Final Estimate	9,450,994	10,172,055	721,061

RUN #1

*****RESULTS OF MONTE CARLO SIMULATION*****

SAMPLE MEAN = 9745100.

SAMPLE VARIANCE = 751580545024.

SAMPLE STANDARD DEVIATION = 866937.

*****COST INTERVALS FOR HISTOGRAM*****

INTERVAL NO.	INTERVAL VALUES			NO. OF OBSERVATIONS	FREQUENCY DISTRIBUTION	CUMULATIVE DISTRIBUTION
1	7143129.	TO	7627657.	2	0.00	0.00
2	7627658.	TO	8112185.	23	0.02	0.02
3	8112186.	TO	8596713.	65	0.06	0.09
4	8596714.	TO	9081241.	139	0.14	0.23
5	9081242.	TO	9565769.	200	0.20	0.43
6	9565770.	TO	10050297.	212	0.21	0.64
7	10050298.	TO	10534825.	134	0.18	0.82
8	10534826.	TO	11019353.	104	0.10	0.93
9	11019354.	TO	11503881.	48	0.05	0.98
10	11503882.	TO	11988409.	15	0.01	0.99
11	11988410.	TO	12472937.	8	0.01	1.00

85%

100%

RUN #2

*****RESULTS OF MONTE CARLO SIMULATION*****

SAMPLE MEAN = 9848057.

SAMPLE VARIANCE = 356460724224.

SAMPLE STANDARD DEVIATION = 597043.

*****COST INTERVALS FOR HISTOGRAM*****

INTERVAL NO.	INTERVAL VALUES			NO. OF OBSERVATIONS	FREQUENCY DISTRIBUTION	CUMULATIVE DISTRIBUTION
1	8306709.	TO	8613554.	14	0.01	0.01
2	8613555.	TO	8920399.	45	0.04	0.06
3	8920400.	TO	9227244.	87	0.09	0.15
4	9227245.	TO	9534089.	150	0.15	0.30
5	9534090.	TO	9840934.	207	0.21	0.50
6	9840935.	TO	10147779.	197	0.20	0.70
7	10147780.	TO	10454624.	143	0.14	0.84
8	10454625.	TO	10761469.	90	0.09	0.93
9	10761470.	TO	11068314.	47	0.05	0.98
10	11068315.	TO	11375159.	18	0.02	1.00
11	11375160.	TO	11682004.	2	0.00	1.00

88%

100%

RUN #3

*****RESULTS OF MONTE CARLO SIMULATION*****

SAMPLE MEAN = 5190389.

SAMPLE VARIANCE = 18478243840.

SAMPLE STANDARD DEVIATION = 135935.

*****COST INTERVALS FOR HISTOGRAM*****

INTERVAL NO.	INTERVAL VALUES	NO. OF OBSERVATIONS	FREQUENCY DISTRIBUTION	CUMULATIVE DISTRIBUTION
1	4797537. TO 4867016.	2	0.00	0.00
2	4867017. TO 4936495.	9	0.01	0.01
3	4936496. TO 5005974.	60	0.06	0.07
4	5005975. TO 5075453.	107	0.11	0.18
5	5075454. TO 5144932.	187	0.19	0.36
6	5144933. TO 5214411.	213	0.21	0.58
7	5214412. TO 5283890.	213	0.21	0.79
8	5283891. TO 5353369.	115	0.11	0.91
9	5353370. TO 5422848.	51	0.05	0.96
10	5422849. TO 5492327.	35	0.03	0.99
11	5492328. TO 5561815.	8	0.01	1.00

94%

RUN #4

*****RESULTS OF MONTE CARLO SIMULATION*****

SAMPLE MEAN = 9812930.

SAMPLE VARIANCE = 63835508736.

SAMPLE STANDARD DEVIATION = 252657.

*****COST INTERVALS FOR HISTOGRAM*****

INTERVAL NO.	INTERVAL VALUES			NO. OF OBSERVATIONS	FREQUENCY DISTRIBUTION	CUMULATIVE DISTRIBUTION
1	9210639.	TO	9330816.	11	0.01	0.01
2	9330817.	TO	9450993.	44	0.04	0.05
3	9450994.	TO	9571170.	88	0.09	0.14
4	9571171.	TO	9691347.	154	0.15	0.30
5	9691348.	TO	9811524.	202	0.20	0.50
6	9811525.	TO	9931701.	200	0.20	0.70
7	9931702.	TO	10051878.	150	0.15	0.85
8	10051879.	TO	10172055.	89	0.09	0.94
9	10172056.	TO	10292232.	46	0.05	0.98
10	10292233.	TO	10412409.	15	0.01	1.00
11	10412410.	TO	10532593.	1	0.00	1.00

85%

100%