USING BIDDING STATISTICS TO PREDICT COMPLETED CONSTRUCTION COST

MICHAEL G. WRIGHT
Rutgers University

TREFOR P. WILLIAMS
Rutgers University

ABSTRACT

The completed cost of a competitively bid construction project often exceeds the original low bid. This paper presents two models to predict completed construction cost based upon characteristics of the submitted bids. Data on completed projects were obtained from New Jersey Department of Transportation for 298 highway construction projects. Median bid and normalized median absolute deviation (NMAD) were selected from various bid characteristics as the best predictors of completed construction cost. Regression and neural network models were developed from the data. Both models have similar utility to predict completed costs. Due to ease of use, the regression model is preferred over the neural network model.

INTRODUCTION

Construction involves bringing equipment, materials and labor to a unique site and building a unique product. The product remains fixed at the site after the construction team leaves. Construction differs from manufacturing where product lots or batches are produced at fixed locations and finished products are distributed to the user [12]. Most public works projects are procured through competitive bidding. Constructors submit bids based upon a defined scope of work, and contracts are awarded to the lowest responsive, responsible bidder. The constructor's compensation can be based upon a fixed lump sum fee for the defined scope of work but frequently a unit-price contract is used. Through change orders, additions and deductions to the scope of work are made, and the constructor's compensation is increased or decreased respectively.

The completed cost to the owner of a competitively bid project often exceeds the original low bid. Factors that contribute to cost overruns include bidding errors, poor design, design constructibility, project complexity, poor construction management, location, weather, labor relations and material

availability. The impact of these factors is difficult to predict. Large increases in cost present a risk to the owner because they can exceed the project budget. If the completed construction cost could be predicted, the financial risk could be objectively evaluated. The following actions could be taken if the projected completed cost exceeds the construction budget:

1.   accept the low bid and the increase in the project budget
2.   reject all bids and solicit new bids
3.   implement tighter construction management controls.

RELATED RESEARCH

Recent research has focused on the development of complex and mathematically rigorous models to evaluate the effects of many different factors on completed construction costs [1] [9] [16]. Other researchers have used common statistical techniques. Pedwell, Hartman and Jergeas [13] used multiple linear regression to evaluate the effect of contractual complexity, design completeness and contract type on construction schedules and costs in the oil industry. Bacon, Besant-Jones and Heidarian [2] used multivariate regression to identify correlation between project characteristics, and schedule and cost overruns on World-Bank-financed, power generation projects in developing countries. Their dependent variables characterized plant technology and size, procurement environment and host country. These two studies used the results to make recommendations about the procurement processes studied. Brandon [3] reported on using stepwise linear regression to estimate the contract cost of building construction in the U.K. Building size, contract duration and number of bidders were used as cost predictors. Williams, Miles & Moore [17] used linear regression to develop models to predict the completed cost from the low bid amount for highway construction projects in the U.K. and the U.S.A. They concluded that there is a distinct multiplicative relationship between low bid and final cost that indicates that final cost increases as a power of the low bid.

Smith and Mason [14] compared the effectiveness of neural network and regression models for parametric cost estimating. They concluded that regression models have significant advantages with respect to accuracy, variability, model creation and model examination. This is true when an appropriate model can be discerned beforehand; however, neural networks have advantages when dealing with data that does not fit low order polynomials. Simple models that are easily understood and applied by practitioners are desired. This study attempts to show that a simple model can be developed. The

goal was to predict completed construction cost of a project based upon characteristics of the bids received. In this study, both neural network and regression models were developed. Their effectiveness to predict completed construction cost based upon characteristics of all the submitted bids were compared.

BID CHARACTERISTICS

Bid characteristics studied include the number of bidders, lowest bid, mean bid, standard deviation, median bid, normalized median absolute deviation (NMAD) and spread. These characteristics are discussed in the sections below.

NUMBER OF BIDS

The number of bids submitted on a project shows the competitiveness of the construction market. High numbers of bids suggest a very competitive market. The number of bids received on the projects included in this study ranged from two to seventeen. (Projects with only one bid submitted were excluded because standard deviation, NMAD and spread cannot be calculated.) In a highly competitive market, bidders may be more willing to take risks by submitting unusually low bids. They hope to regain profit sacrificed to "buy" the contract by submitting claims [18]. This practice, if successful, increases the completed cost over the low bid and, potentially, over the budget.

LOW BID

A high correlation was anticipated between the low bid and the completed construction cost. If there were no change orders, there would be no difference between the bid and completed cost on a lump sum contract. On a unit price contract, the only difference between the low bid and the completed cost would be due to differences in measurement of quantities.

MEAN BID AND MEDIAN BID

The mean (or average) of the bids represents the market's consensus of the true cost of the project. Variations about the mean represent differences in judgment, assumptions or minor bidding errors [5]. Crowley [6] suggests that the median is a better estimate of consensus cost because its robust nature eliminates the influence of spurious bids.

STANDARD DEVIATION AND NORMALIZED MEDIAN ABSOLUTE DEVIATION

Standard deviation of bids measures the variation about the mean. It is an indicator of the bidders' uncertainty about the value of the project. Projects where high standard deviations occur indicate that there is considerable uncertainty among the bidders about the project cost. It can be postulated that a project with a high standard deviation may be prone to larger cost escalation during construction than a project with a low standard deviation and little disagreement among the bidders. This uncertainty could be due to vagueness of the contract documents. For example, the scope of work or field conditions may be poorly documented. In addition, uncertainty could be due to other variables such as availability of labor, equipment or materials; availability of right-of-way or site access; or third party involvement (railroads, utilities, etc.) Usually, uncertainty causes individual bidders to submit higher bids.

Alternatively, the normalized median absolute deviation (NMAD) can be used. This statistic measures variation about the median. Like standard deviation, this statistic can be used as an indication of the variations between bidders. It is calculated as follows:

$$NMAD = \text{median}\left(\left|Bid_i - M_d\right| / 0.6745\right) \qquad (1)$$

where Bid_i is the ith bid, M_d is the median of the bids and the constant 0.6745 is a normalizing constant corresponding to the Z-score partitioning 25% of the normal distribution into the right tail. Crowley [6] has found that NMAD is a superior estimator of bid variation.

SPREAD

Gates [8] defines spread as the difference between the low bid and the second low bid or "the money left on the table." This measure assumes that the second low bid is a reasonable bid and that the second low bid is not a spurious bid. Alternatively, spread can be calculated as the difference between the low bid and the mean bid or consensus value.

Spread is an indicator of another risk to the owner. Occasionally, an unusually low bid is received. A bidder who does not recognize uncertainty can submit these unknowingly. In addition, a low bid can be submitted knowingly by a bidder who chooses to exploit uncertainty to generate claims for additional compensation [18]. While the owner may want to contract with these bidders to obtain a bargain, the owner may be accepting additional risk. The first type of bidder may have trouble completing the work on time, if at all. When the error is discovered, the constructor may seek additional compensation through claims

and change orders to compensate for the inadequate bid. The second bidder type counts on change orders and claims to make a profit. In both cases, the owner will incur additional costs due to the claims themselves, processing and negotiating the claims and completion delays.

COST ESCALATION AND COMPLETED COST

The purpose of the study was to develop models to predict increases in construction cost above the low bid. Two dependent variables were selected for prediction: completed cost and escalation. Completed cost is the amount paid to the constructor after all change orders are negotiated and executed. Largely, the completed cost depends on the project size. Escalation is defined, in the study, as the ratio of the completed cost to the low bid. The ratio was selected so that comparisons over a range of different project sizes could be made. A ratio greater than one represents an increase in cost to the owner. Conversely, a ratio less than one represents a cost saving for the owner.

DATA COLLECTION

Data for this study was obtained from the New Jersey Department of Transportation (NJDOT). The Bureau of Roadway Plans and Specifications provided bid summary tables for highway construction projects advertised between February 1989 and January 1996, inclusive. The tables list the amount of each bid submitted on the projects. A database table was created manually by entering each project and its bids into a database program.

Information on completed projects was obtained from NJDOT's Bureau of Construction Engineering. Database files were provided from their computer records of completed projects. The database files were imported into the database program and cross-tabulated with the bid data. After records with missing or suspect data were eliminated, a table containing 298 records was created. For each construction project, the database contained fields with an identifying number (DPNUM), project location, construction type, project size, bid data, completion date, original contract amount, completed contract amount and all bids received. This table was exported to a spreadsheet program for further data manipulation and preparation of data sets.

The projects included in the study have a total value of $1.51 billion (in 1999 dollars) or approximately 50 percent of the value of all NJDOT projects bid over the period. They include new highway and bridge construction, bridge and highway reconstruction, widening, resurfacing, bridge repair, intersection improvements, safety and traffic control, miscellaneous and unique projects. It

was assumed that these projects represent a random sample of all project types bid by the NJDOT.

The bids were received by NJDOT over a seven-year period. To eliminate the effects of inflation, all bids and costs were converted to 1999 dollars using the ENR Construction Cost Index [9] and the following equation:

$$C_2 = C_1 \times (I_2 / I_1) \tag{2}$$

where C_2 = cost in year two dollars, C_1 = cost in year one dollars, I_2 = ENR Construction Cost Index for year 2 and I_1 = the ENR Construction Cost Index for year 1. TABLE 1 presents univariate statistics for the data set in 1999 dollars.

TABLE:1. Univariate Statistics of 298 NJDOT Projects (1999 dollars)

Description	Lowest	First Quartile	Median	Third Quartile	Highest
Number of Bidders	2	5	6	9	17
Low Bid	$23,786	$737,266	$1,545,520	$3,449,208	$66,567,024
2nd Low Bid	$26,829	$812,874	$1,622,628	$3,676,656	$73,264,210
Mean Bid	$33,992	$906,284	$1,782,257	$4,024,154	$72,080,807
Median Bid	$30,448	$886,041	$1,708,598	$3,943,882	$70,348,202
Standard Dev.	$12,309	$102,849	$ 210,590	$ 468,696	$ 7,349,627
NMAD	$ 3,018	$ 79,485	$ 177,619	$ 408,199	$ 7,373,985
Spread (Low Bid)	$ 107	$ 30,821	$ 82,735	$ 216,824	$ 6,697,186
Spread (Mean Bid)	$ 6,492	$ 91,222	$ 212,723	$ 485,484	$10,040,701
Completed Cost	$ 8,626	$747,789	$1,639,458	$3,712,551	$73,036,689

The technique of data-splitting (cross-validation) [11] was used, as described below, to validate the models. The complete data was separated into a model subset and a validation subset. The validation data, consisting of 50 projects, was created by selecting every sixth project. This selection method eliminated effects of potential, hidden, temporal trends in the data.

REGRESSION MODEL

The first step in the development of the regression model was to test the correlation between the bid characteristics (independent variables) and escalation and completed construction cost (dependent variables). TABLE 2 lists the Pearson Correlation Coefficients for each bid characteristic. From

inspection of the table, the correlation between bid characteristics and escalation is very weak so prediction of escalation was abandoned. The correlation between the remaining bid characteristics and completed cost is very strong except for the number of bidders.

TABLE: 2. Correlations Between Bid Characteristics and Escalation and Completed Cost

Bid Characteristic	Escalation	Completed Cost
Number of Bidders	0.003	0.241
Low Bid	0.079	0.982
2nd Low Bid	0.083	0.979
Mean Bid	0.083	0.979
Median Bid	0.085	0.980
Standard Deviation	0.104	0.877
NMAD	0.128	0.911
Spread (Low Bid)	0.090	0.656
Spread (Mean Bid)	0.109	0.833

Second, a stepwise linear regression analysis was performed using the data set. The analysis revealed that all the bid characteristics contribute information to the regression model except spread and standard deviation. High multicolliniarity existed between median bid, mean bid, and low bid. This suggested that these statistics contributed redundant information to the model so mean and low bid statistics were dropped. Median bid and normalized median absolute deviation were selected for further analysis.

Third, using the method of least squares, a linear regression model was constructed using 248 projects in the model subset. The following equation was derived:

$$CCost = 0.9510*(Median) + 0.9094*(NMAD) - \$278{,}973. \qquad (3)$$

where $CCost$ = completed construction cost, Median = median bid and NMAD = normalized median absolute deviation of the bids. The multiple coefficient of determination was 0.992 and the root mean square of error was $1,001,006.

Fourth, a residuals analysis was performed. All tests were satisfactory except a plot of residuals versus predicted values of completed cost suggested heteroscedasticity. Natural logarithm variance-stabilizing transformation was applied to the dependent variable: completed cost; however, the resulting

models produced inferior results. A multiplicative model was constructed by transforming both independent and dependent variables using the natural log function. The following model was developed:

$$\ln(CCost) = 1.0979 * \ln(\text{Median}) - 0.0857 * \ln(NMAD) - 0.5127 \qquad (4)$$

which can be transformed into:

$$CCost = \frac{\text{Median}^{\wedge}1.0979}{1.67^{*}(NMAD)^{\wedge}0.0857} \qquad (5)$$

This model did not exhibit heteroscedasticity as the linear model did. The multiple coefficient of determination was 0.967.

Fifth, the validation subset was used to make predictions and the coefficient of determination was calculated.

NEURAL NETWORKS

Artificial neural networks (ANNs) attempt to imitate the learning ability of biological brains. In ANNs, processing elements called nodes or neurons are arranged and interconnected in a network. Many topologies with differing characteristics have been developed. Well-known examples are the Hopfield network and the Kohonen network. ANNs can be "hardwired" in electronic circuits or simulated using software on conventional serial computers. For this study, simulation software was used.

A multi-layer perceptron (MLP) model was selected for the study because MLP's are optimized for prediction applications [15]. In MLP's, neurons are arranged in layers. Each node is connected to all nodes in the next layer. Each connection has a weight (w) that determines how strong the connection is within the network. Higher weighted connections contribute more to the solution. Data are fed to the top layer or input layer. Each node processes the data and passes its output to the next lower level. Data is passed to lower layers until it reaches the lowest level or output level. Intermediate layers are called hidden layers. One input node is required for each datum in the problem set, and one output node is required for each datum in the solution set. Nodes process the input data using the following equations [3]:

$$IN = \sum_{i=1}^{n} w_n^{\,*} x_n \qquad (6)$$

where: IN = the net input to the node, w_n = the weight of the nth connection and x_n = the input from the nth connection, and;

$$OUT = 1/(1+e^{-k*IN}) \qquad (7)$$

where OUT = the output from the node and k = a constant, and e = the base of the natural logarithm. The latter equation is called the activation function. The form shown here is the sigmoid function. It has an S-shape curve.

Training is accomplished by presenting training sets of input patterns (problem data) and output patterns (solutions) to the network. As the training set is processed by the network, the network learns how to estimate the correct solution. The steps in the training process are as follows [3]:

1. Present an input pattern and let the ANN produce an output using its current weights.
2. If the output is the same as the desired output, go back to step 1 using the next input pattern.
3. If the output does not match the desired output, adjust the weights associated with each active connection. Inactive connection weights are not changed because they do not contribute to the solution. In MLP's using the backpropagation method, weights are adjusted to reduce half the sum of squares of errors.
4. Repeat the process for each input pattern until the error falls below a given threshold.

Once the ANN is trained, new problem data can be fed to it and solved. One problem with training ANNs is overtraining. Overtraining occurs when the ANN "memorizes" the input and output data. When overtrained, ANNs produce exact solutions for the training set but do not produce acceptable solutions for new problem data. To avoid overtraining, a validation subset is created from the training set. During the training cycle, the error of the validation set is measured. During the early stages of learning, the error of the validation set will decrease at the same rate as the training set if the ANN is training well. When the ANN begins to learn the noise in the data, the validation set error will begin to increase while the training set error will

continue to decrease. The optimum network is the one that has the lowest validation set error [15].

MULTILAYER PERCEPTRON MODEL

The first step in selecting the MLP topology is determining the number of layers and the number of nodes in each layer. The number of nodes in the input layer must match the number of fields in the input records. The input fields consist of median bid data and NMAD data so two input nodes are required. The output field is completed construction cost so one output node is required. The number of nodes in the hidden layers and the number of hidden layers required to produce good results cannot be determined beforehand. Trial-and-error must be used to find the best combination of layers and nodes. Certain rules of thumb have been recommended based upon experience of neural network developers. For most models, one or two hidden layers is adequate [7]. Also, the maximum number of connections in the network should be one-tenth the number of training sets [15]. Given 248 records in the training set, the maximum number of connections could be 25. Therefore, the number of hidden nodes in one hidden layer can be calculated as follows:

$$W_{max} = \left(I_{node} + O_{node}\right) \times H_{node} \tag{8}$$

where W_{max} = maximum number of weights, I_{node} = input nodes and O_{node} = output nodes and H_{node} = hidden nodes. If $W_{max} = 25$, $I_{node} = 2$ and $O_{node} = 1$, then $H_{node} = 8$ max. The selected MLP topology consisted of two input nodes connected to two hidden nodes in one layer that in turn were connected to one output node resulting in six connections.

Second, training options of the MLP were configured as follows:

- The activation function defines how the net input to a node is translated into an output value. The hyperbolic tangent function was selected.
- Distribution of initial weights defines how weights are distributed at the start of training. A uniform distribution was used.
- The learning rule determine how the connection weights are changed as the neural network learns. The conjugate gradient rule was selected. It measures the error surface gradient and uses a compromise between the direction of the steepest gradient and the previous direction of change [15].

- Training stop threshold sets goals for when training is complete. RMS = 0.001 and percentage correct = 75% were used.

Training the MLP model was performed using the estimation subset. The problem set (i.e. input data) consisted of median bid and NMAD. The solution set (i.e. output data) consisted of the corresponding completed construction costs. The input and output to the neural network were the same as the independent and dependant variables used in the regression model so that a comparison between the two models could be made.

Once trained, the model was used to predict completed cost. The validation subset was run through the neural network to make predictions. The results were compared to the actual completed costs, and the coefficient of determination was calculated.

MODELING RESULTS

Data-splitting (cross-validation) was used to validate the model results. This technique calls for splitting the data into model and validation subsets and comparing measures of model validity such as coefficients of determination or mean square of errors for each subset [11]. If the validation measures are close, the model can be considered valid. Coefficients of determination were calculated for the model subset and the validation subset for both models. The results are summarized in TABLE 3. The results indicate that approximately 99 percent of the variation in completed construction cost is predicted by the median bid and NMAD of the bids. The high coefficients of determination for the estimation and validation subsets give high confidence in both models.

TABLE 3. Model Coefficients of Determination

Model	Model Subset	Validation Subset
Neural Network	0.99	0.94
Multiplicative	0.97	0.98

FIGURE 1 shows a plot of actual completed construction costs versus the models' predicted costs. The diagonal line represents a perfect match between actual cost and predicted cost. The tight clustering of points along this line shows graphically that the models give good predictions of completed construction cost.
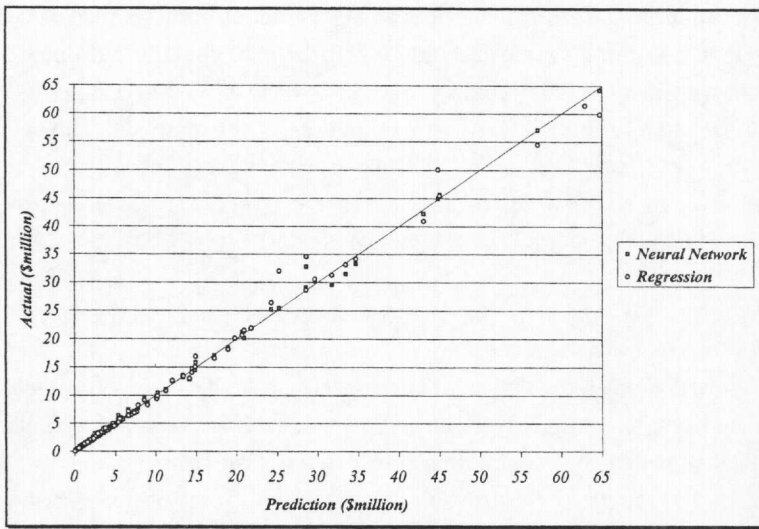
FIGURE 1: Actual Cost Vs. Predicted Cost

The following example demonstrates an application of the multiplicative regression model. Consider the bids listed in TABLE 4 for a representative project selected from the data set. The median bid is $1,650,982. Using Eq. (1), NMAD is calculated to be $171,823. Substituting these values into Eq. (5), completed cost is predicted to be $1,429,211. The low bid was $1,465,966. This suggests that the low bid was reasonable.

TABLE 4. Representative Bid Data

Description	Value
Low Bid	$1,465,966
Median Bid	$1,650,982
NMAD	$171,823
Bid 1	$1,465,966
Bid 2	$1,535,088
Bid 3	$1,575,649
Bid 4	$1,644,122
Bid 5	$1,650,982
Bid 6	$1,728,437
Bid 7	$1,963,532
Bid 8	$2,164,540
Bid 9	$2,735,988

The models are based upon highway construction projects in New Jersey, USA, that had more than one bidder. The data set contains information on projects with median bids ranging from \$30,448 to \$101,096,058 and NMAD from \$3018 to \$7,373,985. Due to the limitations of the regression and neural network models, application of the models should be limited to projects of similar size. In addition, application to other markets may produce unreliable results so the model should be used with caution.

CONCLUSIONS

This study demonstrates that a simple model can be developed to predict the completed cost of competitively bid construction projects. Bid characteristics, specifically median bid and normalized median absolute deviation, are valid predictors of completed construction cost. The multiplicative regression model and neural network model are valid models, and both models produce comparable results. The regression model produced a simple equation. It can be used to make predictions with any handheld calculator whereas the neural network model requires neural network simulation software. For this reason, the regression model is preferred.

ACKNOWLEDGEMENTS

REFERENCES

[1] ALARCON, LUIS F. and DAVID B. ASHLEY, "Modeling Project Performance for Decision Making," *Journal of Construction Engineering and Management*, ASCE, Vol. 122, No. 3, pp. 265-273.

[2] BACON, ROBERT W., JOHN E. BESANT-JONES, JAMISHID HEIDARIAN, "Estimating Construction Costs and Schedules Experience with Power Generation Projects in Developing Countries," *The World Bank*, Washington, D.C.

[3] BHARATH, R. and J. DROSEN, *Neural Network Computing*, Windcrest/McGraw-Hill, Blue Ridge Summit, PA, 1994.

[4] BRANDON, PETER S. (Ed.) *Quantity Surveying Techniques New Directions.* Blackwell Scientific Publications, London, U.K. 1992.

[5] CROWLEY, LARRY G. and DONN E. HANCHER, "Risk Assessment of Competitive Procurement," *Journal of Construction Engineering and Management,* ASCE Vol. 121, No. 2, 1995, pp. 230-237.

[6] CROWLEY, LARRY G. "Robust Statistical Estimators for Use Within Competitive Bid Data," *Journal of Construction Engineering and Management,* ASCE, Vol. 123, No. 1, 1997, pp. 53-63.

[7] FLOOD, I., and KARTAM, N. "Neural Networks in Civil Engineering. I: Principals and Understanding," *Journal of Computing in Civil Engineering,* ASCE, Vol. 8, No. 2, 1994, pp. 131-148.

[8] GATES, MARVIN, "Bidding Strategies and Probabilities," *Journal of the Construction Division,* Proceedings of the American Society of Civil Engineers, ASCE, Vol. 93 No. 1, 1967, pp. 75-107.

[9] GROGAN, T., "Cost Indexes: Listed in ENR's January 4, 1999 Issue" http://www.enr.com/cost/cost2.asp (1 Feb., 1999)

[10] HAJAK, MAKARAND, DANIEL W. HALPIN, and JORGE VANEGES, "COMPASS – New Paradigm for Project Cost Control Strategy and Planning," *Journal of Construction Engineering and Management,* ASCE, Vol. 122, No. 3, 1996, pp. 254-263.

[11] MENDENHALL, WILLIAM and TERRY SINCICH, *A Second Course in Statistics: Regression Analysis,* Prentice Hall, Upper Saddle River, NJ, 1996.

[12] OSTWALD, PHILLIP F., *Construction Cost Analysis and Estimating,* Prentice Hall, Upper Saddle River, NJ, 2001.

[13] PEDWELL, KEITH, FRANCIS T. HARTMAN, and GEORGE F. JERGEAS, "Project Capital Cost Risk and Contracting Strategies," *Cost Engineering,* AACE, 1998.

[14] SMITH, ALICE E., and MASON, ANTHONY K., "Cost Estimation Predictive Modeling: Regression Versus Neural Network," *The Engineering Economist,* Vol. 42, No. 2, 1997, pp. 137-161.

[15] SPSS and Recognition Systems, Inc. Neural Connection 1.0 User's Guide. SPSS, Chicago, IL, 1995.

[16] TOURAN, ALI. "Probablistic Cost Estimating with Subjective Correlations," *Journal of Construction Engineering and Management,* ASCE Vol. 119, No. 1, 1993, pp. 57-71.

[17] WILLIAMS, T., J. MILES, and C. MOORE, "Modeling Cost Increases in Competitively Bid Construction Projects," *Proceedings of the Construction and Building Research Conference 1998,* Vol. 1, The Royal Institution of Chartered Surveyors, Oxford, U.K., September 1998, pp. 102-110.

[18] ZACK, JAMES G., "Claimsmanship: Current Perspective," *Journal of Construction Engineering and Management,* ASCE Vol. 119, No. 3, pp. 480-497.

BIOGRAPHICAL SKETCHES

MICHAEL G. WRIGHT is a Ph.D. candidate in the Civil and Environmental Engineering Department at Rutgers University, New Brunswick, NJ, where he is specializing in Construction Engineering and Management. In addition, he is a Project Manager with the consulting firm of Arora and Associates, Lawrenceville, NJ. Mr. Wright has over 20 years of civil engineering experience. He is member of ASCE, ASHE and NSPE, and a Registered Professional Engineer in NJ and PA

TREFOR P. WILLIAMS is an associate professor in the Civil and Environmental Engineering Department at Rutgers University, New Brunswick, NJ. He holds a Ph.D. in Civil Engineering from Georgia Institute of Technology. The FAA, NJHA, NJDOT, NYDOT and USDOT have funded his research into construction management and transportation engineering. Dr. Williams is the author of more than 35 published papers, a member of ASCE and a Registered Professional Engineer in NY and NJ.