*Jordan Journal of Civil Engineering,* Volume **2**, No. **3**, **2008**

# Prediction Model for Construction Cost and Duration in Jordan

***Ayman A. Abu Hammad*** $^{1)}$***, Souma M. Alhaj Ali*** $^{2)}$***, Ghaleb J. Sweis*** $^{3)}$***and Adnan Bashir*** $^{4)}$

$^{1)}$ Assistant Professor of Construction Engineering and Management, Department of
Civil Engineering, College of Engineering, Applied Science University, Amman, Jordan.
Email: dr-abuhammad@asu.edu.jo. (Corresponding Author)
$^{2)}$ Assistant Professor, Industrial Engineering Department, The Hashemite University, Zerqa, Jordan.
Email: souma_alhajali@hu.edu.jo
$^{3)}$ Assistant Professor, Civil Engineering Department, College of Engineering, University of Jordan, Amman 11942,
Jordan. Email: gsweis@ju.edu.jo
$^{4)}$ Assistant Professor, Industrial Engineering Department, The Hashemite University, Zerqa, Jordan.
Email: abashir@hu.edu.jo

## ABSTRACT

Risk is mitigated in the course of reliable prediction. A probabilistic model is proposed to predict the risk effects on time and cost of construction projects. Project managers and consultants can use the model in estimating project cost and duration based on historic data.

Statistical regression models and sample tests are developed using real data of 140 projects. The research objective is to develop a model to predict project cost and duration based on historic data of similar projects. The model result can be used by project managers in the planning phase to validate the schedule critical path time and project budget. Research methodology is steered per the following progression: i) Conduct nonparametric test for project cost and time performance. ii) Develop generic multiple-regression models to predict project cost and duration using historic performance data. iii) The percent prediction error is statistically analyzed; and found to be substantial; thus, iv) Custom multiple regression models are developed for each project type to obtain statistically reliable results. In conclusion, the 95% point estimation of error margin= ±0.035%. Therefore, at a probability of 95%, the proposed model predicts the project cost and duration with a precision of ±0.035% of the mean cost and time.

**KEYWORDS:** Prediction model, Construction projects, Time, Cost, Jordan.

## INTRODUCTION

Construction has rapidly developed in the past few years in Jordan with huge investments in public and infrastructure projects which amounted to 6.5 Billion JD in 2006 from a low amount of 2.5 Billion JD in 2003. The nominal GDP in 2006 was 14.3 Billion JD with an annual real growth of 6.4%, out of which the construction sector accounted for 11.6% (Jordan Investment and Finance Bank, 2007).

Construction investments are sensitive to time and cost overrun. Delay and cost escalating are considered two severe consequences of change. Variation Orders (VOs) are inevitable in all projects. Yet, change poses a substantial risk that cannot be predicted in the contract for taking on preventive measures. On the other hand, construction contracts give the owner the right to modify, add and delete work items at anytime via a VO. Thus, the

---

Accepted for Publication on 1/7/2008.

© 2008 JUST. All Rights Reserved.

scalable effects of delay and cost escalation are rarely dealt with efficiently by project managers. In many cases, change clause is used by contractors to account for their losses due to competitive underbidding practices.

## RESEARCH PROBLEM AND OBJECTIVES

Construction projects are hardly ever constructed as designed. Consequently, as built plans and consistent updating of project schedules are the current procedures for modeling project change impacting both cost and time. The overall project budget and duration of the CPM schedule should be verified against historic performance of projects. A statistical model is developed in this research to predict with significant confidence the terminal project cost and duration. Therefore, construction contracts and computerized project management tools can incorporate the statistical model results through incorporating extra float to project duration and analyzing financial contingency. Network schedules should be fine-tuned with the regression model results in order to accommodate for the risk of change. Moreover, contract duration and price should be estimated using the model results. Therefore, floats and budgets should be extended to account for the risk of change which is estimated by using the proposed regression model.

The research goal is to provide investors in the construction business with a tool to predict and quantify substantial risk associated to change. A regression model is developed and validated in this research to predict project delay and cost escalation using real project data of different construction types and different sizes for each type. Additionally, the contract scope variation is also considered, i.e., civil, electromechanical or both.

## LITERATURE BACKGROUND

Bromilow established the parameters of cost/time performance predictability using the contract time performance of 329 projects constructed during 1964-1969 (Bromilow, 1969).

Leishman presented the legal consequences of delays in construction (Leishman, 1991). Yates developed a decision support system for delay analysis (Yates, 1993). Herbsman *et al.* studied the effect of delays on cost and quality (Herbsman *et al.*, 1995). Assaf *et al.* surveyed the causes of delay in large building-construction projects in Saudi Arabia as seen by contractors, consultants and owners (Assaf *et al.*, 1995). Kaming *et al.* mentioned main causes of cost overruns: material cost increase due to inflation, inaccurate material estimating and project complexity. Kaming found that the predominant causes of delay are design changes, poor labor productivity and inadequate planning (Kaming *et al.*, 1997).

Al-Momani developed a quantitative regression model for estimating the actual time using the data of 130 public building projects constructed during 1990-1997 in Jordan. Al-Momani concluded that the main causes of delay in construction projects are caused by: designers, owner changes, weather, differing site conditions, delays in material deliveries, economic conditions and increase in quantities. The research reported frequencies of time extensions for the different causes of delays (Al-Momani, 2000).

Odeh *et al.* identified the most important causes of delay through a survey targeted contractors and consultants in Jordan (Odeh *et al.*, 2002). Hsieh *et al.* used statistical correlation and variance analysis to find the connection among layers of events or causes for change orders. The statistical analysis covered a data base of 90 public projects completed before 2000 in Taiwan (Hsieh *et al.*, 2004).

Lyer *et al.* investigated factors adversely affecting the cost performances of projects in India. The causes were: conflict among project participants, ignorance and lack of knowledge, presence of poor project specific attributes and non existence of cooperation, hostile socio-economic and climatic conditions, reluctance in timely decision, aggressive competition at tender stage and short bid preparation time (Lyer *et al.*, 2005). Causes of delay and cost overrun are analyzed using 450 randomly selected projects (Koushki, 2005).

Project duration forecasting methods using earned value analysis are used during the project to highlight the need for eventual corrective action (Vandevoorde, 2006).

Assaf *et al.* surveyed opinions of contractors, consultants and owners on delay causes. The study identified 73 causes of delay (Assaf *et al.*, 2006). Faridi *et al.* ranked the causes of delay based on their relative importance index. The most significant delay causes in the UAE construction market are: approval of drawings, inadequate early planning and slow decision making by owners (Faridi *et al.*, 2006). Finally, Sweis *et al.* identified the most common causes of residential project delay. Financial difficulties and owner required change orders are at the forefront of other causes of delay. Sweis concluded that severe weather conditions and changes in government regulations ranked among the least important causes for delay (Sweis, 2007).

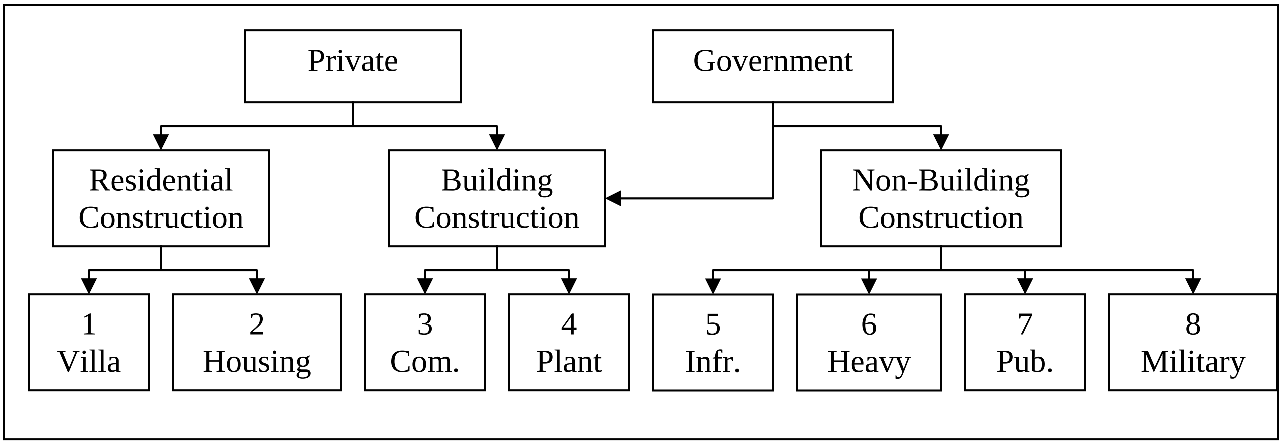


**Figure (1): Project Classification Flow Chart.**

**Table (1): Project Classification (A).**

| | Classification A: Residential - Building Construction- Heavy- Industrial | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **Classes** | **Private** | | | | **Government** | | | |
| **Category** | **Residential** | | **Building Const.- Non-Resid.** | | **Non-Building Const.** | | **Buil.Const. Proj.** | |
| **Sub-Category** | Villa (1) | Appart. (2) | Commer. (3) | Plant (4) | Infrastructure (5) | Heavy (6) | Public (7) | Military (8) |
| **Code#** | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| **Sample Count** | 17 | 15 | 43 | 8 | 2 | 1 | 68 | 14 |
| **Class A** | A: Solo | A: Building | A: Stall Shops | A: Workshop | A: Transportation | A: Dams | A: Adminis. | A: Adminis. |
| **Class B** | B: Twin | B: Complex | B: Mall | B: Complex Zone | B: Bridges | B: Environ. | B: Parks | B: Camps |
| **Class C** | C: Complex | C: Tower | C: Offices | C: Manuf. Zone | C: Sewage | | C: Religious | |
| **Class D** | D: Palace | | D: Recreation & Tour. | D: Transform Plants | D: Communication | | D: Education | |
| **Class E** | | | E: Hotels | E: Mining | E: Power | | E: Health | |

## DATA COLLECTION

A pilot study is carried out in this research on real data of 140 projects selected via a systematic random sampling procedure in order to develop generic and custom regression models. The projects were constructed during 1994-2002. This period witnessed a stable inflation rate in the Jordan construction market. Project information is collected that are built in the same period from owners, i.e., private and public governmental institutions, and from contractors. For the critical financial-managerial aspect of this research, data providers are not disclosed for confidentiality reasons.

Sample projects are classified per project type and per job type. Construction project types are sampled for the assumption that the project type impacts the cost escalation and project delay variables. The type breakdown follows the Engineering News Record bulletin of project breakdown into four major categories: residential, building construction, heavy and industrial projects.

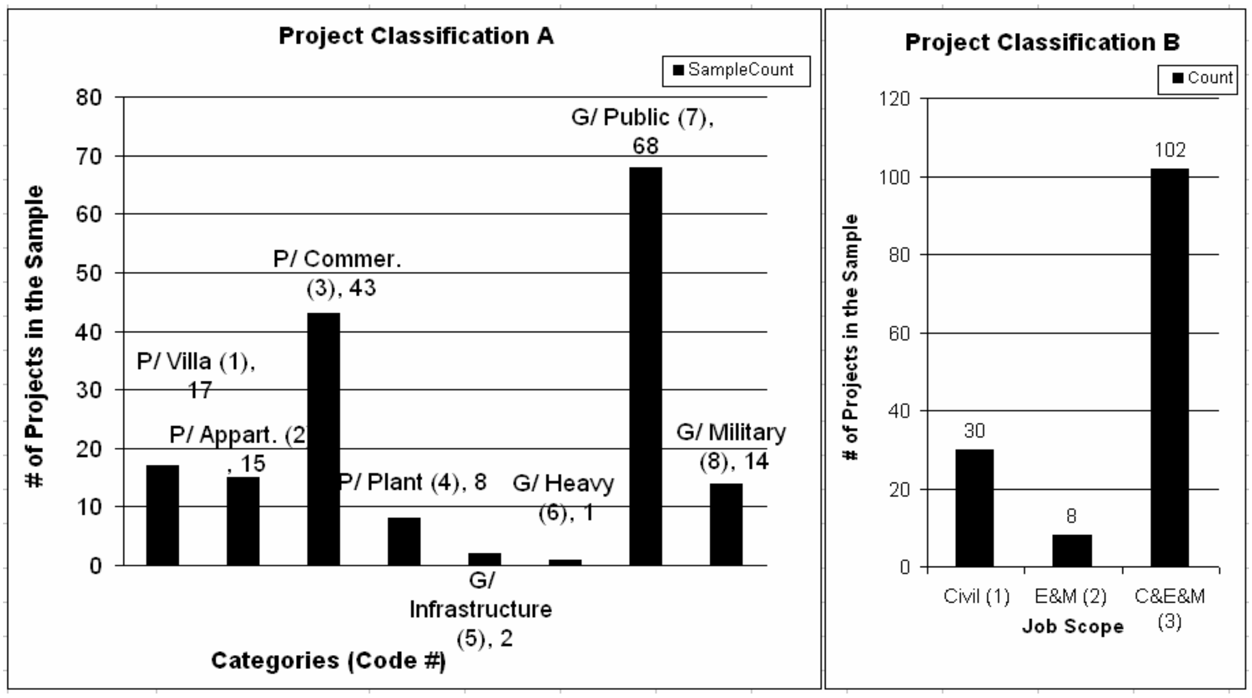
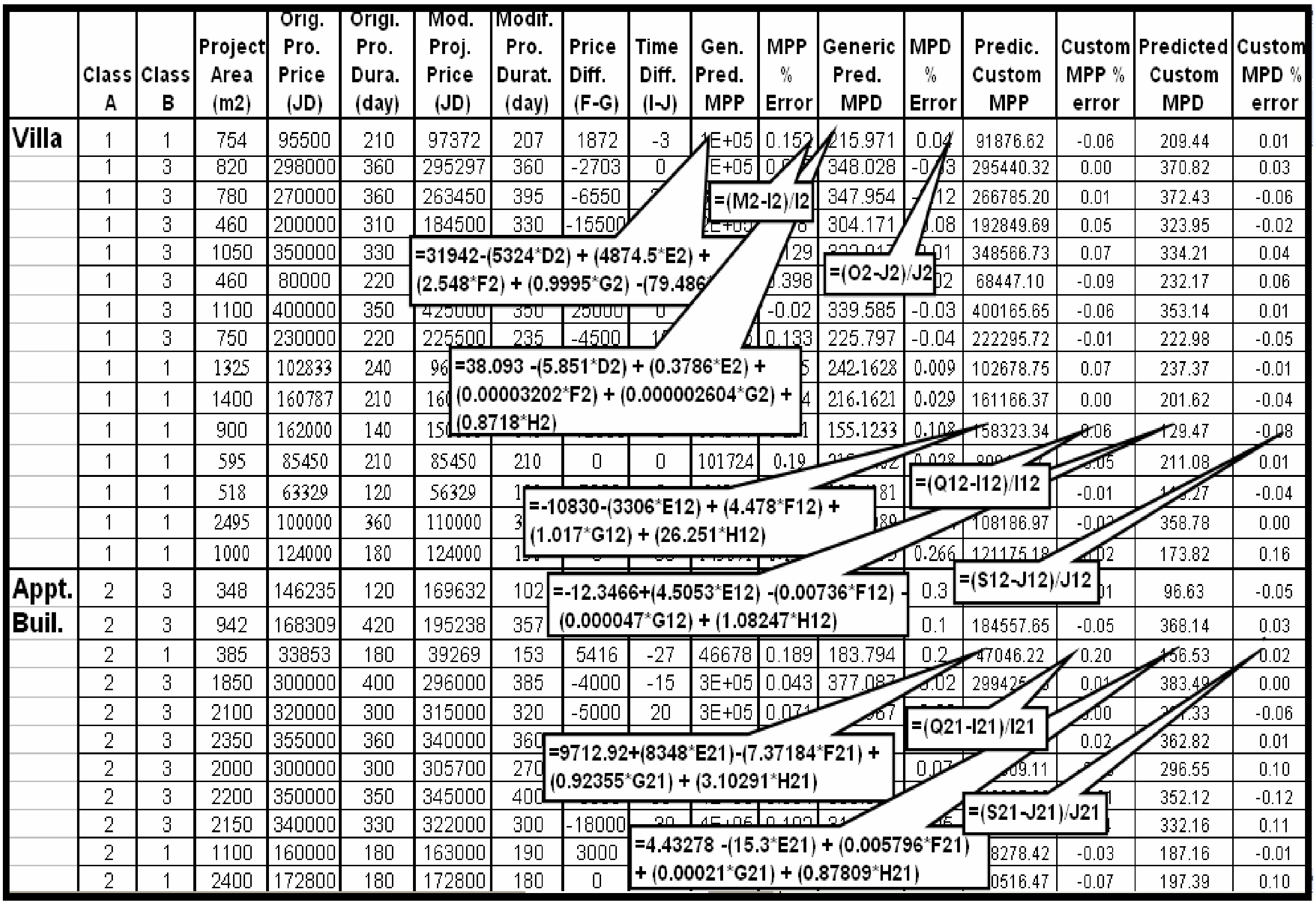


**Figure (2): Sample Project Classification (A) and (B).**

**Table (2): Summary Statistics for Data Variables.**

| Col. Title | Area (m$^2$) | Orig. PP (JD) | MPP (JD) | price diff. (JD) | Orig. PD (day) | MPD (day) | time Diff. (day) |
|---|---|---|---|---|---|---|---|
| Mean | 3492.73 | 725246.3 | 733959.4 | -9349.9 | 244.835 | 228.2747 | 16.56 |
| SD | 8965 | 3.49E+06 | 3489685 | 104015 | 130.68 | 128.14 | 40.172 |
| Sample Size (n) | 88 | 88 | 88 | 88 | 86 | 86 | 86 |
| SEM | 939.78 | 365680 | 365818 | 10720 | 13.699 | 13.433 | 4.211 |
| LL 95%CI | 1622.6 | -2457 | 5980.7 | -30045 | 217.57 | 201.54 | 8.18 |
| UL 95%CI | 5362.9 | 1452950 | 1461938 | 12619 | 272.1 | 255.01 | 24.941 |
| Min. | 75 | 3578 | 4150.5 | -875000 | 30 | 24 | -90 |
| Median | 1325 | 160000 | 150000 | 0 | 240 | 210 | 10 |
| Maximum | 75391 | 3.18E+07 | 3.16E+07 | 268750 | 900 | 857 | 185 |
| Normality Test KS | 0.3419 | 0.4356 | 0.4404 | 0.3187 | 0.113 | 0.07893 | 0.1602 |
| Normality Test P-value | <0.0001 | <0.0001 | <0.0001 | <0.0001 | 0.006 | >0.1 | <0.0001 |
| Normal? | No | No | No | No | No | Yes | No |



=31942-(5324*D2) + (4874.5*E2) + (2.548*F2) + (0.9995*G2) -(79.486*

=(M2-I2)/I2

=(O2-J2)/J2

=38.093 -(5.851*D2) + (0.3786*E2) + (0.00003202*F2) + (0.000002604*G2) + (0.8718*H2)

=-10830-(3306*E12) + (4.478*F12) + (1.017*G12) + (26.251*H12)

=(Q12-I12)/I12

=-12.3466+(4.5053*E12) -(0.00736*F12) - (0.000047*G12) + (1.08247*H12)

=(S12-J12)/J12

=9712.92+(8348*E21)-(7.37184*F21) + (0.92355*G21) + (3.10291*H21)

=(Q21-I21)/I21

=(S21-J21)/J21

=4.43278 -(15.3*E21) + (0.005796*F21) + (0.00021*G21) + (0.87809*H21)

| | Class A | Class B | Project Area (m2) | Orig. Pro. Price (JD) | Origi. Pro. Dura. (day) | Mod. Proj. Price (JD) | Modif. Pro. Durat. (day) | Price Diff. (F-G) | Time Diff. (I-J) | Gen. Pred. MPP | MPP % Error | Generic Pred. MPD | MPD % Error | Predic. Custom MPP | Custom MPP % error | Predicted Custom MPD | Custom MPD % error |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Villa** | 1 | 1 | 754 | 95500 | 210 | 97372 | 207 | 1872 | -3 | 1E+05 | 0.152 | 215.971 | 0.04 | 91876.62 | -0.06 | 209.44 | 0.01 |
| | 1 | 3 | 820 | 298000 | 360 | 295297 | 360 | -2703 | 0 | [illegible] | [illegible] | 348.028 | [illegible] | 295440.32 | 0.00 | 370.82 | 0.03 |
| | 1 | 3 | 780 | 270000 | 360 | 263450 | 395 | -6550 | [illegible] | [illegible] | [illegible] | 347.954 | [illegible] | 266785.20 | 0.01 | 372.43 | -0.06 |
| | 1 | 3 | 460 | 200000 | 310 | 184500 | 330 | -15500 | [illegible] | [illegible] | [illegible] | 304.171 | [illegible] | 192849.69 | 0.05 | 323.95 | -0.02 |
| | 1 | 3 | 1050 | 350000 | 330 | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | 348566.73 | 0.07 | 334.21 | 0.04 |
| | 1 | 3 | 460 | 80000 | 220 | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | 68447.10 | -0.09 | 232.17 | 0.06 |
| | 1 | 3 | 1100 | 400000 | 350 | 425000 | 350 | 25000 | 0 | [illegible] | -0.02 | 339.585 | -0.03 | 400165.65 | -0.06 | 353.14 | 0.01 |
| | 1 | 3 | 750 | 230000 | 220 | 225500 | 235 | -4500 | [illegible] | [illegible] | 0.133 | 225.797 | -0.04 | 222295.72 | -0.01 | 222.98 | -0.05 |
| | 1 | 1 | 1325 | 102833 | 240 | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | 242.1628 | 0.009 | 102678.75 | 0.07 | 237.37 | -0.01 |
| | 1 | 1 | 1400 | 160787 | 210 | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | 216.1621 | 0.029 | 161166.37 | 0.00 | 201.62 | -0.04 |
| | 1 | 1 | 900 | 162000 | 140 | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | 155.1233 | 0.108 | 158323.34 | [illegible] | 129.47 | -0.08 |
| | 1 | 1 | 595 | 85450 | 210 | 85450 | 210 | 0 | 0 | 101724 | 0.19 | [illegible] | [illegible] | [illegible] | [illegible] | 211.08 | 0.01 |
| | 1 | 1 | 518 | 63329 | 120 | 56329 | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | -0.01 | [illegible] | -0.04 |
| | 1 | 1 | 2495 | 100000 | 360 | 110000 | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | 108186.97 | [illegible] | 358.78 | 0.00 |
| | 1 | 1 | 1000 | 124000 | 180 | 124000 | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | 0.266 | 121175.18 | [illegible] | 173.82 | 0.16 |
| **Appt. Buil.** | 2 | 3 | 348 | 146235 | 120 | 169632 | 102 | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | 0.3 | [illegible] | [illegible] | 96.63 | -0.05 |
| | 2 | 3 | 942 | 168309 | 420 | 195238 | 357 | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | 0.1 | 184557.65 | -0.05 | 368.14 | 0.03 |
| | 2 | 1 | 385 | 33853 | 180 | 39269 | 153 | 5416 | -27 | 46678 | 0.189 | 183.794 | 0.2 | 47046.22 | 0.20 | [illegible] | 0.02 |
| | 2 | 3 | 1850 | 300000 | 400 | 296000 | 385 | -4000 | -15 | 3E+05 | 0.043 | [illegible] | [illegible] | [illegible] | 0.01 | 383.49 | 0.00 |
| | 2 | 3 | 2100 | 320000 | 300 | 315000 | 320 | -5000 | 20 | 3E+05 | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | -0.06 |
| | 2 | 3 | 2350 | 355000 | 360 | 340000 | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | 0.02 | 362.82 | 0.01 |
| | 2 | 3 | 2000 | 300000 | 300 | 305700 | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | 296.55 | 0.10 |
| | 2 | 3 | 2200 | 350000 | 350 | 345000 | 400 | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | 352.12 | -0.12 |
| | 2 | 3 | 2150 | 340000 | 330 | 322000 | 300 | -18000 | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | 332.16 | 0.11 |
| | 2 | 1 | 1100 | 160000 | 180 | 163000 | 190 | 3000 | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | -0.03 | 187.16 | -0.01 |
| | 2 | 1 | 2400 | 172800 | 180 | 172800 | 180 | 0 | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | -0.07 | 197.39 | 0.10 |

**Figure (3): Computation of the Generic versus Custom MPP and MPD %Prediction Error.**

The project type classification (depicted in Figure 1 and Table 1) designates the project with a code number from 1-8 indicative of the classification type. Private project categories are residential: villa and housing type projects; and non-residential or building construction: commercial and plant projects. Government project categories are building construction projects in particular; i.e., hospitals, schools, malls, hotels and non-building construction projects; i.e., infrastructure, highway construction, heavy, public and military projects.

The job type classification (classification B) designates the sample projects with a number 1-3. Number one designates civil job. Number two designates electromechanical jobs and number three designates projects of civil and electromechanical jobs. The objective of classification B is to track the effect of contract type on the variables in concern; i.e., cost and duration.

The data are tabulated in rows for each project type. The columns are designated the following variables: project type, job type, project area, original project price, modified project price, original project duration and modified project duration. The project cost difference and project time difference are calculated in two additional columns.

Figure 2 depicts the frequencies of project types and job types, classification A and B, respectively. Unfortunately, plant, infrastructure and heavy projects of project classification A have few number of data points; thus, cannot be used in developing custom regression formulas.

Table 2 shows the summary statistics of the data variables: the mean, standard deviation, standard error of the mean and the 95% confidence interval. In addition, the Kolmogorov-Smirnov (KS) test of normality for the random variables is calculated. The KS test indicates that only the *modified project duration* data follow the normal distribution; however, the other variables are not normally distributed.

Henceforth, a paired t-test is performed for the mean differences among the modified project price MPP and the original project price (cost escalation). Additionally, the t-test is performed for the mean differences among the Modified Project Duration MPD and the original project duration (delay). The two tests are performed to analyze the statistical significance of cost escalation and project delay based on the sample project data.

## DATA ANALYSIS

### Nonparametric Analysis for the Mean Project Cost Escalation

The two-tail P-value is calculated for the mean of the differences among the modified and original project prices of the data. The mean differences are considered not significantly different from zero because the two-tailed P-value is 0.40 >>0.05, for a t-statistic= 0.84 with 87 degrees of freedom (number of data points less one).

The pairing of the data appears to be effective because the correlation coefficient r = 0.99 with a two-tailed P-value <0.0001, considered extremely significant. Therefore, effective pairing results in a significant correlation between the two data columns. Table 2 shows that the differences are not sampled from a normal distribution. Thus, the normality assumption is tested using the Kolmogorov-Smirnov (KS) test of 0.31 with a P-value < 0.0001; that is less than 0.05. Therefore, a nonparametric test, namely Wilcoxon Matched-Pairs Signed-Ranks test (WMPSR), is performed using the sample data of the project cost escalation. The median of the differences between the original and modified project price does not differ significantly from zero. The P-value is 0.89 and exceeds the 0.05 level of significance, thus considered not significant.

Table 3 shows the WMPSR calculations. The analysis is performed for 76 pairs of data; 12 data pairs are excluded from the calculations for the reason of equality.

**Table (3): WMPSR Calculation for the Mean Project Cost Escalation.**

| | |
|---|---|
| Sum of all signed ranks (W) | -51.000 |
| Sum of positive ranks (T+) | 1437.5 |
| Sum of negative ranks (T-) | -1488.5 |
| Number of pairs | 76 |

### Nonparametric Analysis for the Mean Project Delay

The two-tail P-value is calculated for the mean of the differences between the modified and original project duration. Contrary to the project cost analysis in the previous section, the mean differences are considered extremely significant and different from zero. This is evident because the two-tailed P-value is 0.0006 <<0.05, for a t-statistic= 3.57 with 85 degrees of freedom (86 data points less one).

The pairing of the data appears to be effective because the value of the correlation coefficient r = 0.95 and the value of the two-tailed P-value <0.0001 (considered extremely significant). Therefore, effective pairing results in a significant correlation between the data columns. The last three columns of Table 2 show that the time differences are not sampled from a normal distribution. The normality assumption is further tested using the Kolmogorov-Smirnov (KS) test and found to fail the normality test with a KS value= 0.18 and a P-value < 0.0001; thus less than 0.05. Therefore, the WMPSR nonparametric test is performed using the sample data of the project delay. The median of the differences between the original and modified project duration are found to differ significantly form zero. The P-value is 0.0002<<0.05 level of significance; thus considered significant.

Table 4 shows the WMPSR calculations. The analysis is performed for a 65 pairs of data; whereas 21 pairs are excluded from the calculations, since both values are equal.

**Table (4): WMPSR Calculation for the Mean Project Delay.**

| | |
|---|---|
| Sum of all signed ranks (W) | 1147 |
| Sum of positive ranks (T+) | 1646 |
| Sum of negative ranks (T-) | -499 |
| Number of pairs | 65 |

### Multiple Linear Regression Analysis for Estimating the Modified Project Price (MPP) and the Modified Project Duration (MPD)

The regression analysis tool performs linear regression analysis by using the "least squares" method to fit a line through a set of observations. Regression analysis provides inference about how a single dependent variable is affected by the values of one or more independent variables.

Dependent and independent variables are defined at the beginning of the regression analysis. Two multiple regression steps are performed on the data for: i) the Modified Project Price MPP, and ii)) the Modified Project Duration MPD. Thus, two generic regression equations are developed utilizing the project data of different classes.

**Table (5): The Correlation Matrix.**

| | | **Dependent Variable (Y)** | **Independent Variables** | | | | |
|---|---|---|---|---|---|---|---|
| | | E: MPP | A | B | C | D | G |
| Var. 1: A | Project Type | 0.0631 | 1 | 0.4127 | -0.0455 | 0.0666 | -0.1652 |
| Var. 2: B | Job Type | 0.1023 | 0.4127 | 1 | 0.0587 | 0.103 | 0.1917 |
| Var. 3: C | Project Area | 0.3661 | -0.0455 | 0.0587 | 1 | 0.3602 | 0.1927 |
| Var. 4: D | Original Bid Price | 0.9996 | 0.0666 | 0.103 | 0.3602 | 1 | 0.5896 |
| Var. 5: G | Orig. Proj. Duration | 0.5882 | -0.1652 | 0.1917 | 0.1927 | 0.5896 | 1 |

Each Correlation Coefficient is calculated independently, without considering the other variables.

### Regression Model for the Modified Project Price (MPP)

At the regression analysis, the MPP is set as the (Y) dependent variable. The assigned dependent variables (X) are: project and job types, project area, estimated project price and estimated project duration. All independent variables are known and estimated based on the project blue prints and estimated Bill of Quantities (BOQ).

The degrees of freedom are calculated as equal to the number of data points minus number of independent variables-1. That is 91-5-1=85.

The correlation coefficient matrix of Table 5 depicts the linear relationship between each two variables. The coefficient of correlation is a value between negative one and positive one. Additionally, a value closer to +ve 1 indicates a strong relationship, a value of zero indicates no relationship among the two variables; however, a value close to –ve 1 indicates a reverse relationship.

#### *Equation of the Multiple Regression Analysis for the MPP*

The regression analysis returned the following equation (that statistically fits the data best) accompanied with statistical inference.

**[F: MPP]= 31942 – 5324 * [A: Project Type]+ 4874.5 * [B: Job Type] + 2.548 * [C: Project Area] + 0.9995 * [D: Original Project Price] – 79.486 * [G: Original Project Duration]**

#### *Regression Model Goodness of fit to Real Project Data*

Table 6 depicts the 95% confidence interval of the regression model coefficients. The 95% confidence interval for the constant means that there are 95%

confidence that the true population mean of the equation constant, -51550, lies in the interval of [lower limit =mean-2*SE, and upper limit =mean+2*SE] = [-51550 - 45733, -51550 + 45733]= [-142636,39535]. Of course, with 99.7% confidence the CI expands over 6-sigma (standard error) around the mean coefficient -51550 instead of 4-sigma in the case of 95.4% CI.

**Table (6): Standard Error and 95% Confidence Intervals for the Regression Equation Coefficients.**

| The Goodness of fit of the Regression Equation to the Real Data | | | | |
|---|---|---|---|---|
| **Variable** | **Coefficient** | **SE** | **LL- 95%CI** | **UL- 95%CI** |
| **Constant** | 31942 | 38809 | -142636 | 39535 |
| **A: Proj. Type** | -5324 | 4745 | -14774 | 4126.5 |
| **B: Job Type** | 4874.5 | 13799 | -22609 | 32358 |
| **C: Proj. Area** | 2.548 | 1.291 | -0.02376 | 5.119 |
| **D: Orig. Proj. Price** | 0.9995 | 0.004148 | 0.9912 | 1.008 |
| **G: Orig. Proj. Duration** | -79.486 | 110.45 | -299.46 | 140.49 |

**Table (7): Variables of Significant Contribution to the MPP Regression Model.**

| Variables of Significance to the Regression Model | | | |
|---|---|---|---|
| **Variable** | **t-ratio** | **P-value** | **Significance** |
| **Constant** | 0.823 | 0.4128 | Not Signif. |
| **A: Proj. Type** | 1.122 | 0.265 | Not Signif. |
| **B: Job Type** | 0.3532 | 0.7248 | Not Signif. |
| **C: Proj. Area** | 1.973 | 0.0517 | Not Signif. |
| **D: Orig. Proj. Price** | 240.95 | < 0.0001 | **Significant** |
| **G: Orig. Proj. Duration** | 0.7197 | 0.4737 | Not Signif. |

The goodness of the fit for the above equation is explained by the calculated R-squared value of 99.92%. This means that 99.92% of the variance in the variable (E: MPP) is explained by the model. The obtained P-value of <0.0001 (considered extremely significant) is the probability for obtaining an R squared value of 99.92% by chance assuming no linear relationship is established among the variables.

*Significant Variables of the Regression Model*

Each P-value of Table 7 compares the regression model with a simpler model deleting one of the variables. Therefore, the P-value tests the effect of one variable, after accounting for the other variables.

*Multicollinearity Assessment*

The R squared values depicted in Table 8 quantify how well that x-variable is predicted from the other x-variables (ignoring Y). Since all r squared values are low; i.e., less than 0.75, it is concluded that the x-variables are independent of each other. Therefore, multicollinearity poses no problem to the analysis.

**Table (8): R-Squared Values for X-Independent Variables.**

| **Variable** | **$R^2$ with other x** |
|---|---|
| A: project type | 0.2915 |
| B: scope | 0.251 |
| C: area | 0.143 |
| D: original price | 0.4454 |
| E: original duration | 0.4484 |

### Regression Model for the Modified Project Duration (MPD)

At the regression model, the MPD is set as (Y) dependent variable. The dependent variables (X) are:

project and job types, project area, estimated project price and estimated project duration. All independent variables are known and estimated, as explained before, based on the project blue prints and estimated Bill of Quantities (BOQ).

The correlation coefficient matrix shown in Table 9 depicts the linear relationship between each two variables of the regression analysis for the MPD.

**Table (9): The Correlation Matrix of the MPD Regression Analysis.**

| | | **Dependent Variable (Y)** | **Independent Variables** | | | | |
|---|---|---|---|---|---|---|---|
| | | H: MPD | A | B | C | D | G |
| Var. 1: A | Project Type | 0.3085 | 1 | -0.3973 | 0.2735 | -0.0274 | -0.1547 |
| Var. 2: B | Job Type | 0.1 | -0.3973 | 1 | 0.0573 | 0.0974 | 0.1672 |
| Var. 3: C | Project area | 0.2083 | 0.2735 | 0.0573 | 1 | 0.3608 | 0.0344 |
| Var. 4: D | Original Bid Price | 0.5855 | -0.0274 | 0.0974 | 0.3608 | 1 | 0.026 |
| Var. 5: E | Orig. Proj. Duration | -0.0831 | -0.1547 | 0.1672 | 0.0344 | 0.026 | 1 |

Each Correlation Coefficient is calculated independently, without considering the other variables.

**Table (10): 95% Confidence Intervals for the MPD Regression Equation Coefficients.**

| **The Goodness of Fit of the Regression Equation to the Real Data** | | | | |
|---|---|---|---|---|
| **Variable** | **Coefficient** | **SE** | **LL- 95%CI** | **UL- 95%CI** |
| **Constant** | 38.093 | 14.133 | 9.944 | 66.241 |
| **A: Proj. Type** | -5.851 | 1.728 | -9.293 | -2.41 |
| **B: Job Type** | 0.3786 | 5.025 | -9.63 | 10.387 |
| **C: Proj. Area** | 3.20E-05 | 0.0004702 | -0.0009044 | 0.0009685 |
| **D: Orig. Proj. Price** | 2.60E-06 | 1.51E-06 | -4.05E-07 | 5.61E-06 |
| **E: Orig. Proj. Duration** | 0.8718 | 0.04022 | 0.7917 | 0.9519 |

**Table (11): Variables of Significant Contribution to the MPD Model.**

| **Variables of Significance to the Regression Model** | | | |
|---|---|---|---|
| **Variable** | **t-ratio** | **P-value** | **Significance** |
| **Constant** | 2.695 | 0.0085 | Significant |
| **A: Proj. Type** | 3.386 | 0.0011 | Significant |
| **B: Job Type** | 0.07535 | 0.99401 | Not Signif. |
| **C: Proj. Area** | 0.0681 | 0.9459 | Not Signif. |
| **D: Orig. Proj. Price** | 1.724 | 0.0884 | Not Signif. |
| **G: Orig. Proj. Duration** | 21.676 | <.0001 | Significant |

*Equation of Multiple Regression Analysis for the MPD*

The regression analysis returned the following equation (that statistically fits the data best).

**[G: Modified Project Duration]= 38.093 – 5.851 * [A: Project Type]+ 0.3786 * [B: Job Type] + 3.202E-05 *[C: Project Area] + 2.604E-06*[D: Original Project Price] + 0.8718 * [E: Original Project Duration]**

*Regression Model Goodness of fit to Real Project Data*

Table 10 depicts the 95% confidence intervals of the regression model coefficients.

The goodness of fit for the above equation is explained by the calculated R-squared value of 92.1%. This means that 92.1% of the variance in the variable (G: MPD) is explained by the model. The obtained P-value of <0.0001 (considered extremely significant) is the probability for obtaining an R squared value of 92.1% by chance assuming no linear relationship is established among the variables.

*Significant variables to the Regression Model*

Each P-value of Table 11 compares the regression model with a simpler model deleting one of the variables. Therefore, the P-value tests the effect of one variable on the dependent variable, after accounting for the other variables. The null hypothesis stating that there is no significant effect of the independent variable on the dependent variable is rejected when the P-value is less or equal 0.05 at the 95% level of confidence. Conversely, the alternative hypothesis stating that there is a statistical significant effect of the variable on the dependent variable is accepted.

*Multicollinearity Assessment*

The R squared values depicted in Table 12 quantify how well that x-variable is predicted from the other x-variables (ignoring Y). The VIF values are calculated from the $R^2$ values. Since all r squared values are low; i.e., less than 0.75, it is concluded that the x-variables are independent of each other. Therefore, multicollinearity poses no problem to the analysis.

**Table (12): R-Squared Values for X-Independent Variables.**

| Variable | VIF | $R^2$ with other x |
|---|---|---|
| A: project type | 1.41 | 0.2915 |
| B: scope | 1.34 | 0.251 |
| C: area | 1.17 | 0.143 |
| D: original price | 1.8 | 0.4454 |
| E: original duration | 1.81 | 0.4484 |

**Inference on the Regression Analysis Results**

It is observed in Table 7 that only the original project price variable significantly has a statistical impact on the result of the MPP regression model with a P-value <0.0001. This is evident because in practice the Lowest Price Contractor (LPC) is always selected. Underbidding, that is pricing construction tenders lower than estimated costs, became a habit by contractors in response to severe competition. Underbidding contractors rely on VOs to account for their losses due to underbidding. Thus, VOs leading to cost escalation is only attributed to original project price (underbidding price). However, the other variables such as: project area, project type or job type were marginal to model result, which is underlined statistically in Table 7. Table 7 results can be attributed to contractors' involvement in project cost escalation due to underbidding practices in most public competitive bids.

Controversially, it is observed in Table 11 that the variables of significant contribution to the MPD regression model are project type and the original project duration in addition to the constant. The statistical evidence of the significant contribution of the original project duration variable to the result of the MPD regression model can be substantiated from practice. The original project duration is a special condition in the contract and is planned by the consultant (owner representative) responsible for preparing the tender documents. Captivatingly, the project type, which has also a significant contribution to the MPD regression model (as shown in Table 11), is a basis for selecting the appropriate project delivery method; namely, Design-Bid-Build (DBB), Design-Build (DB), Design-Build-Operate-Transfer (DBOT) and Construction Management project delivery systems (CM). The project delivery systems are originally ascribed to control over project duration throughout changing the sequence of the project life-cycle phases by paralleling the construction phase to the design phase, or further, phasing and packaging the construction phase for a fast-track management objective. Thus, Table 11 results can be attributed to consultants' weakness in predicting appropriate duration specific to project nature, complexity and selected delivery system.

Knowing that it is the consultant responsibility to prepare the bid documents including the project special conditions; i.e., project duration in addition to particular stipulations of how the project will be administered.

**Table (13): Descriptive Statistics for the %Prediction Error of the Generic Regression Models.**

| Descriptive Ststistics | MPP (JD) | MPD (day) | Price Diff. (JD) | Duration Diff. (day) | Pred. MPP (JD) | Pred. MPD (day) | MPP % Error | MPD % Error |
|---|---|---|---|---|---|---|---|---|
| **Mean** | 761939.3 | 232.77 | -8713.02 | 16.56 | 733969.14 | 228.26626 | -0.00154 | 0.023297 |
| **SD** | 3546075 | 128.69 | 102258 | 40.172 | 3488328 | 122.97 | 0.5948 | 0.1892 |
| **Sample Size (n)** | 88 | 88 | 88 | 88 | 88 | 88 | 91 | 91 |
| **SEM** | 378013 | 13.877 | 10720 | 4.211 | 365676 | 12.891 | 0.06235 | 0.01984 |
| **LL 95%CI** | 9315.8 | 205.13 | -30045 | 8.18 | 6273.7 | 202.61 | -0.1256 | -0.01618 |
| **UL 95%CI** | 1514563 | 260.41 | 12619 | 24.941 | 1461665 | 253.92 | 0.1225 | 0.06277 |
| **Min.** | 7827.7 | 36 | -875000 | -90 | -15297 | 24.06 | -2.95 | -0.45 |
| **Median** | 166316 | 210 | 0 | 10 | 173158 | 221.06 | 0.06 | 0 |
| **Maximum** | 31575000 | 857 | 268750 | 185 | 3.18E+07 | 8.67E+02 | 1.99 | 0.77 |
| **Normality Test KS** | 0.4417 | 0.08933 | 0.3187 | 0.1602 | 0.4338 | 0.09914 | 0.2521 | 0.1692 |
| **Normality P-value** | <.0001 | 0.0868 | <0.0001 | <0.0001 | <0.0001 | <0.0276 | <0.0001 | <0.0001 |
| **Normal?** | No | Yes | No | No | No | No | No | No |

In summary, Table 11 results can be debated further in the argument that project delay is the sole responsibility of the owner and project consultant. Though, cost escalation is associated to the contractor and bidding practices in pricing hidden items highly at the same time controlling the overall bid price by underpricing other invaluable items of the bid. The contractor has a clear vision -compared to consultant- of items leading to escalated project price. Change clause is used extensively by most contractors to break even with the underbidding price in order to survive fierce competitive tendering.

The analysis results underline that the contractor is primarily responsible for cost escalation. Most contractors finance project activities via surety loans. Therefore, the contractor bears the responsibility in delegate financial troubles since the contractor is entitled to suspend or terminate the contract in case of owner default in paying promptly per the FIDIC general conditions (FIDIC, 2006). In practice, contractors try hard to keep good relations with owners despite owner financial default and are reluctant to claim their rights through litigation. Although contractors commit severe management mistakes by not strictly apply the contract general conditions, reasons are driven by greed in obtaining future jobs from owners, or simply to avoid litigation procedure which seldom ended in favor of contractors.

On the other hand, owner and consultant are responsible for project delays attributed to inappropriate predetermined time frames or even to bad selection of proper delivery system.

Delay and cost escalation has scalable impact according to project type and scope and are estimated using the regression models developed in this research for the MPP and the MPD.

The analysis follows in the next section using the percent error statistic to check the accuracy in predicting the project cost and duration using the new regression models.

**Table (14): Custom Regression Models for each Project Category.**

| | Regression Coefficients | | | | | | Regression Statistics | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Project Categ. | Intercept | Class A | Class B | Project Area ($m^2$) | Orig. PP (JD) | Orig. PD (day) | Multiple R | R Square | Adjusted R Square | Standard Error | N |
| **MPP** | | | | | | | | | | | |
| Generic | 31942 | -5324 | 4874.5 | 2.548 | 0.9995 | -79.486 | | 0.9992 | | | 91 |
| 1 | -10830 | 0 | -3305.6 | 4.478 | 1.017 | 26.251 | 0.9955 | 0.991 | 0.887387 | 12078.98 | 15 |
| 2 | 9712.923 | 0 | 8347.99 | -7.371841 | 0.9235499 | 3.102911 | 0.9986 | 0.99721 | 0.996094 | 8175.485 | 15 |
| 3 | 111303 | 0 | -24426 | -23.87 | 1.143 | 13.978 | 0.99986 | 0.99973 | 0.99952 | 55481.55 | 15 |
| 4 | 46089 | 0 | -26831 | 0.5157 | 0.9774 | 86.978 | 0.99986 | 0.99986 | 0.999413 | 56737.59 | 8 |
| 7 | 69970 | | -19580 | 5.158 | 0.9907 | -124.11 | 1 | 1 | 0.9999 | 46396 | 25 |
| 8 | 7913.698 | 0 | 0 | -1.8913 | 0.9171782 | -4.60282 | 0.98446 | 0.96916 | 0.859909 | 32804.19 | 14 |
| **MPD** | | | | | | | | | | | |
| Generic | 38.093 | -5.851 | 0.3786 | 3.202E-05 | 2.604E-06 | 0.8718 | | 0.921 | | | 91 |
| 1 | -12.3466 | 0 | 4.50526 | -0.007358 | -4.68E-05 | 1.082469 | 0.99117 | 0.98243 | 0.975398 | 13.97708 | 15 |
| 2 | 4.432784 | 0 | -15.302 | 0.0057964 | 0.0002104 | 0.878085 | 0.98297 | 0.96622 | 0.952709 | 23.99193 | 15 |
| 3 | 48.788 | 0 | -14.12 | 0.007931 | 2.394E-05 | 0.963647 | 0.91665 | 0.84025 | 0.776355 | 52.45802 | 15 |
| 4 | -77.565 | 0 | -3.61 | 0.0005968 | 3.74E-05 | 1.188 | 0.9978 | 0.9957 | 0.9784 | 0 | 6 |
| 7 | 2.558 | | 3.535 | -0.005083 | 7.69E-06 | 0.8057 | 0.9826 | 0.9655 | 0.9587 | 25065 | 25 |
| 8 | 38.74327 | 0 | 0 | -0.00265 | -0.000106 | 0.785878 | 0.89759 | 0.80567 | 0.647369 | 41.85683 | 14 |

## Percent Error Statistic in Predicting Project Cost and Time via Regression Formulas

The percent errors for the predicted MPP and predicted MPD using the generic regression models are calculated in two additional columns of the data sheet. The percent error is estimated by calculating the difference between the predicted value by the model and the measured MPP and MPD, then the result is divided by the MPP and MPD value in each case. The generic regression models developed so far are not accurate. Inaccuracy is evident from the means, standard deviations and confidence intervals of the percent error statistics depicted in Table 13.

## Inference on the Percent Error Analysis of the Generic Regression Models

Table 13 indicates the following:

1) The 95% CI for the percent error of the regression model is (-0.125, 0.12) and (-0.01, 0.06) for the MPP and MPD, respectively. The 95% CI values indicate a low error margin for the MPD regression model estimate and a higher error margin for the MPP. Therefore, the duration estimate via regression formulas is more reliable than the price estimate.
2) In support of the previous point, the standard deviation for the MPP % error is high compared to the MPD % error.
3) The standard deviation is very high for the price diff. and duration diff.
4) The median of the price difference equals zero and the median of the time difference = 10 days.

## Custom Regression Models for each Project Type

Although the generic regression models provide a prediction of project cost and time overrun, custom regression models are further analyzed to pursue higher accuracy and reliability. Higher accuracy and reliability of the model are dictated through obtaining lower standard deviation of the distribution of the mean % error statistic; thus, narrower CI of the % error statistic. Custom regression models for each project type would result in a lower prediction error compared to the generic regression models.

**Table (15): Comparative Descriptive Statistics for the %Prediction Error of the Generic versus All Custom Regression Models.**

| *Descriptive Statistics* | *Price Diff. (JD)* | *Duration Diff. (day)* | *Generic Predict. MPP* | *MPP % Error* | *Generic Predict. MPD* | *MPD % Error* | *Cust. Predict. MPP* | *Cust. MPP % Error* | *Cust. Predict. MPD* | *Cust. MPD % Error* |
|---|---|---|---|---|---|---|---|---|---|---|
| **Mean** | -8713.02 | 16.56 | 733969.14 | -0.0015 | 228.26626 | 0.023297 | 723378.29 | 0.082701 | 161.559 | 0.078222 |
| **SD** | 102258 | 40.172 | 3488328 | 0.5948 | 122.97 | 0.1892 | 2872206 | 0.8213 | 170.15 | 0.2376 |
| **Sample Size (n)** | 91 | 91 | 91 | 91 | 91 | 91 | 137 | 137 | 137 | 91 |
| | | | | | | | | | | |
| **SEM** | 10720 | 4.211 | 365676 | 0.06235 | 12.891 | 0.01984 | 245389 | 0.07017 | 14.537 | 0.02505 |
| **LL 95%CI** | -30045 | 8.18 | 6273.7 | -0.1256 | 202.61 | -0.01618 | 242416 | -0.05483 | 133.07 | 0.02837 |
| **UL 95%CI** | 12619 | 24.941 | 1461665 | 0.1225 | 253.92 | 0.06277 | 1204341 | 0.2202 | 190.05 | 0.1281 |
| | | | | | | | | | | |
| **Min.** | -875000 | -90 | -15297 | -2.95 | 24.06 | -0.45 | -12663 | -2.38 | -147.59 | -0.31 |
| **Median** | 0 | 10 | 173158 | 0.06 | 221.06 | 0 | 190673 | 0 | 133.88 | 0.01 |
| **Maximum** | 268750 | 185 | 3.18E+07 | 1.99 | 866.61 | 0.77 | 31576080 | 6.54 | 856.17 | 0.97 |
| | | | | | | | | | | |
| **Normality Test KS** | 0.3187 | 0.1602 | 0.4338 | 0.2521 | 0.09914 | 0.1692 | 0.3929 | 0.346 | 0.1525 | 0.2028 |
| **Normality P-value** | <0.0001 | <0.0001 | <0.0001 | <0.0001 | <0.0276 | <0.0001 | <0.0001 | <0.0001 | <0.0001 | <0.0001 |
| **Normal?** | No | No | No | No | No | No | No | No | No | No |

For the objective of shortening the CI of the error margin, multiple regression models are developed specific for the data relative to each project type; i.e., villas, apartment buildings, commercial buildings, plants, public building construction projects and military projects. The custom regression models are developed by using the SPSS software and found similar to the results obtained via the data analysis tool in EXCEL.

The generic prediction formulas (regression coefficients), new custom prediction formulas specific to each project type and relevant statistics are depicted in Table 14. Regression formulas for project categories of code numbers: 5 and 6 (infrastructure and heavy projects) are not developed for the reason of insufficient number of data points in the sample necessary for conducting the analysis.

### Percent Prediction Error Analysis in Estimating MPP and MPD

The regression model formulas shown in Table 14 are embedded in the data sheet depicted in Figure 2. The custom regression formulas are programmed in the columns (Q and S). The percent error for the custom regression models are calculated in columns (R and T) relative to the modified project price and duration (columns I and J).

The custom regression models for all project types are tabulated in Table 15 in addition to their % prediction error, and compared to the % prediction error of the generic regression models shown in Table 13.

Table 15 compares the % error statistic of the generic formula of all data to the % error estimation of all data

(91 to 137 data points) after embedding the new custom prediction formulas opposite to each project category data.

Table 16 compares the % error statistic of the generic models for the villa projects (15 data points) to the % error statistic of the custom models for the villa projects.

Only the percent prediction-error statistics are transferred from Tables 15 and 16 to Table 17 for further analysis and ease in discussing the results.

**Table (16): Comparative Descriptive Statistics for the %Prediction Error of the Generic Regression Model versus Custom Regression Model for Villa Project Type.**

| *Descriptive Statistics* | *Price Diff. (JD)* | *Duration Diff. (day)* | *Generic Predict. MPP* | *MPP % Error* | *Generic Predict. MPD* | *MPD % Error* | *Cust. Predict. MPP* | *Cust. MPP % Error* | *Cust. Predict. MPD* | *Cust. MPD % Error* |
|---|---|---|---|---|---|---|---|---|---|---|
| **Mean** | -3180.9 | 1.8 | 199616.07 | 0.1712 | 255.495 | 0.02158 | 178309.38 | -0.004 | 256.436 | 0.00133 |
| **SD** | 11188 | 14.168 | 105540 | 0.144 | 72.837 | 0.09485 | 107085 | 0.04968 | 88.312 | 0.05866 |
| **Sample Size (n)** | 15 | 15 | 15 | 15 | 15 | 15 | 15 | 15 | 15 | 15 |
| **SEM** | 2888.7 | 3.658 | 27250 | 0.0372 | 18.806 | 0.02449 | 27649 | 0.01283 | 22.802 | 0.01515 |
| **LL 95%CI** | -9377.2 | -6.047 | 141164 | 0.0914 | 215.16 | -0.0309 | 119002 | -0.058 | 244.53 | -0.0312 |
| **UL 95%CI** | 3015.4 | 9.647 | 258068 | 0.251 | 295.84 | 0.07412 | 237617 | 0.02352 | 305.35 | 0.03382 |
| **Min.** | -24500 | -30 | 86571 | -0.0211 | 137.42 | -0.1191 | 55739 | -0.09 | 115.27 | -0.08 |
| **Median** | -4500 | 0 | 179074 | 0.152 | 225.8 | 0.00901 | 158323 | -0.01 | 232.17 | 0 |
| **Maximum** | 25000 | 35 | 416024 | 0.5369 | 348.03 | 0.266 | 400166 | 0.07 | 372.43 | 0.16 |
| **Normality Test KS** | 0.1924 | 0.3505 | 0.1787 | 0.1952 | 0.1916 | 0.2093 | 0.1698 | 0.1346 | 0.1855 | 0.1746 |
| **Normality P-value** | >0.1 | <0.0001 | >0.1 | >0.1 | >0.1 | 0.076 | >0.1 | >0.1 | >0.1 | >0.1 |
| **Normal?** | Yes | No | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |

**Inference on the Percent Error Analysis of the Custom Regression Models**

The following are observed for the prediction error in percent computed for all sample projects (the left side of Table 17).

1) The mean and standard deviation of the % error for both custom MPP (column 3) and custom MPD (column 4) have increased compared to the % error of the generic MPP (column 1) and generic MPD (column 2).
2) The width of the 95% CI for the mean % error statistic did not differ much for the two cases.
3) None of the statistics is normally distributed because the P-value is less than 0.05 (level of significance).

Conversely, on the right side of Table 17 the following are observed:

1) The mean and standard deviation of the % error for both custom MPP (column 7) and custom MPD (column 8) have substantially decreased compared to the % error of the MPP (columns 5, 3 and 1) and MPD (columns 6, 4 and 2). Taking the MPP statistics for example: the mean error in column 5 (0.17) is greater than in columns 1 and 3, (0.0015 and 0.08), respectively. This result makes sense because of the small sample size effect of 15 projects. The mean error is the minimum in column 7 (0.004) proving a statistical inference that the error in predicting the MPP via the custom regression formula is minimum. This is also evident in the value of the standard deviation in column7 (0.049) which is the minimum of the relevant statistic in columns 1, 3 and 5 (0.59, 0.82 and 0.14, respectively). This statistical evidence indicates that a higher accuracy level is accrued

in predicting the real MPP and MPD by the custom regression model over that of the generic regression model; thus the custom models are more reliable than the generic models.

**Table (17): Comparative %Prediction Error of the Generic Regression Model versus Custom Regression Model.**

| | *Table (15): Percent Prediction Error for all Projects in the Sample (%)* | | | | *Table (16): Percent Prediction Error for Villa Projects in the Sample (%)* | | | |
|---|---|---|---|---|---|---|---|---|
| | *Generic Regression Model* | | *Custom Regression Model* | | *Generic Regression Model* | | *Custom Regression Model* | |
| | *1* | *2* | *3* | *4* | *5* | *6* | *7* | *8* |
| *Descriptive Statistics* | *Gen. MPP* | *MPD* | *MPP* | *MPD* | *MPP* | *MPD* | *MPP* | *MPD* |
| **Mean** | -0.00154 | 0.0233 | 0.082701 | 0.07822 | 0.17118 | 0.02158 | -0.004 | 0.001333 |
| **SD** | 0.5948 | 0.1892 | 0.8213 | 0.2376 | 0.144 | 0.09485 | 0.04968 | 0.05866 |
| **Sample Size (n)** | 91 | 91 | 137 | 91 | 15 | 15 | 15 | 15 |
| **SEM** | 0.06235 | 0.01984 | 0.07017 | 0.02505 | 0.03719 | 0.02449 | 0.01283 | 0.01515 |
| **LL 95%CI** | -0.1256 | -0.01618 | -0.05483 | 0.02837 | 0.09142 | -0.03094 | -0.05801 | -0.03115 |
| **UL 95%CI** | 0.1225 | 0.06277 | 0.2202 | 0.1281 | 0.251 | 0.07412 | 0.02352 | 0.03382 |
| **Min.** | -2.95 | -0.45 | -2.38 | -0.31 | -0.02112 | -0.1191 | -0.09 | -0.08 |
| **Median** | 0.06 | 0 | 0 | 0.01 | 0.152 | 0.009012 | -0.01 | 0 |
| **Maximum** | 1.99 | 0.77 | 6.54 | 0.97 | 0.5369 | 0.266 | 0.07 | 0.16 |
| **Normality Test KS** | 0.2521 | 0.1692 | 0.346 | 0.2028 | 0.1952 | 0.2093 | 0.1346 | 0.1746 |
| **Normality P-value** | <0.0001 | <0.0001 | <0.0001 | <0.0001 | >0.1 | 0.076 | >0.1 | >0.1 |
| **Normal?** | No | No | No | No | Yes | Yes | Yes | Yes |

2) The MPP 95% CI of the prediction error in percent for the custom regression model in column 7 is the shortest interval [-0.05, 0.02], with length = 0.07%, compared to columns 5, 3 and 1 of CI values: ([0.09, 0.25], [-0.05, 0.22], [-0.12, 0.12]), respectively; indicating a higher precision in predicting the real value of the MPP via the custom regression model with a 95% point estimate of error = ±0.035%; that is, the interval length (0.07%) divided by two. Similarly, The MPD 95% CI of the prediction error in percent for the custom regression model in column 8 is the shortest interval [-0.03, 0.03], with length = 0.06, compared to columns 6, 4 and 2 of CI values: ([-0.03, 0.07], [0.02, 0.12], [-0.01, 0.06]), respectively; indicating a higher precision in predicting the real value of the MPD via the custom regression model. The accrued precision of the MPD custom regression models is indicated by the 95% point estimation of the error margin = 0.06 divided by 2= ± 0.03.

In general, the width of the 95% CI of the MPP is shorter than the 95% CI of the MPD.

3) All the statistics are normally distributed since the P-value is greater than 0.05.

## CONCLUDING REMARKS

1) Nonparametric tests on the mean delay and mean cost overrun indicate that the mean delay is

significant and poses threat to project success. However, the mean cost overrun is not significant.

2) Although delay and cost escalation have scalable impact on project success according to project type and scope, it is imperative to allocate the party responsible for each. Future research is to investigate delay and cost escalation causes for sample projects and analyze the share of each contract counterpart; i.e, contractor, owner and consultant contributing to project failure (delay and cost overrun).
3) Delay and cost escalation are estimated using the regression models developed in this research for the generic MPP and MPD. However, the percent error analysis indicated that the duration estimate via generic MPD formula is more reliable than price estimate via generic MPP formula.
4) The percent error statistics showed increased reliability in using the custom regression formulas over the generic regression formulas. It is statistically proven that the custom prediction models are accurate with a 95% probability and will not exceed an error margin of 0.07%, 0.06% for the MPP and the MPD, respectively.
5) The prediction of the proposed models is accurate with a probability of 95%. At such accuracy level, the prediction precision of the terminal project cost and duration predicted via the custom regression models is at most within ±0.035% of the estimated project cost and duration. Typically, the project cost is estimated based on the BOQ items, and the project duration is estimated via the CPM schedule of project activities.

## REFERENCES

Al-Momani, H.A. 2000. Construction Delay: A Quantitative Analysis. *International Journal of Project Management*, 18: 51-59.

Assaf, S.A., Al-Khalil, M. and Al-Hazmi, M. 1995. Causes of Delay in Large Building Construction Projects. *Journal of Management in Engineering*, 11 (2): 45-50.

Assaf, S.A. and Al-Hejji, Sadiq. 2006. Causes of Delay in Large Construction Projects, *International Journal of Project Management*, 24: 349-357.

Bromilow, F.J. 1969. Contract Time Performance: Expectations and the Reality. *Building Forum*, 70-80.

Faridi, A. and El-Sayegh, Sameh. 2006. Significant Factors Causing Delay in the UAE Construction Industry. *Construction Management and Economics*, 24: 1167-1176.

FIDIC. 2006. Conditions of Contract for Construction for Building and Engineering Works, Designed by the Employer, Multilateral Development Bank, Harmonized Edition.

Herbsman, Z.J., Chen, W.T. and Epstein, W.C. 1995. Innovative Contracting Methods in Highway Construction. *ASCE Journal of Construction Engineering and Management*, 121 (3): 272-281.

Hsieh, T., Lu, S. and Wu, C. 2004. Statistical Analysis of Causes for Change Orders in Metropolitan Public Works. *International Journal of Project Management*, 22: 679-686.

Jordan Investment and Finance Bank. 2007. Jordan Economy Performance in 2006. http//www.zawya.com/story.cfm/sidpdf_230707072805.

Kaming, Peter F., Paul O. Olomolaiye, Gary D. Holt and Frank C. Harris. 1997. Factors Influencing Construction Time and Cost Overruns on High-rise Projects in Indonesia. *Construction Management and Economics*, 15: 83-94.

Koushki, P.A., Al-Rashid, K. and Kartam, N. 2005. Delays and Cost Increases in the Construction of Private Residential Projects in Kuwait. *Construction Management and Economics*, *Taylor and Francis Group*, 23: 285-294.

Leishman, D.M. 1991. Protecting Engineer against Construction Delay Claims. *Journal of Management in Engineering*, 7 (3): 319-333.

Lyer, A., K.C. and Jha, K.N. 2005. Factors Affecting Cost

Performance: Evidence from Indian Construction Projects. *International Journal of Project Management*, 23: 283-295

Odeh, A. M. and Battayneh, Hussein. 2002. Causes of Construction Delay: Traditional Contracts. *International Journal of Project Management*, 20: 67-73.

Sweis, G., Sweis, R., Abu Hammad, A. and Shboul, A. 2007. Delays in Construction Projects: The Case of Jordan. Accepted for Publication at the *International Journal of Project Management*, Manuscript #:JPMA-D-07-00094R1, Elsevier Editorial System.

Vandevoorde, S. and Vanhoucke, M. 2006. A Comparison of Different Project Duration Forecasting Methods Using Earned Value Metrics. *International Journal of Project Management*, 24: 289-302.

Yates, J. 1993. Construction Decision Support System for Delay Analysis. *ASCE Journal of Construction Engineering and Management*, 119 (2): 226-244.