IMPROVING SPACE PROJECT COST ESTIMATING WITH ENGINEERING MANAGEMENT VARIABLES

Joseph W. Hamaker, Engineering Cost Office, NASA Marshall Space Flight Center

---

Abstract

Current space project cost models attempt to predict space flight project cost via regression equations, which relate the cost of projects to technical performance metrics (e.g. weight, thrust, power, pointing accuracy etc.). This paper examines the introduction of engineering management parameters to the set of explanatory variables. A number of specific engineering management variables are considered and exploratory regression analysis is performed to determine if there is statistical evidence for cost effects apart from technical aspects of the projects. It is concluded that there are other non-technical effects at work and that further research is warranted to determine if it can be shown that these cost effects are definitely related to engineering management.

Introduction

Predicting the cost of future projects is not an exact science. There are far too many variables involved and many of these are not predictable with any precision before the project begins. This is especially true in the venue of NASA space projects that are almost always on the cutting edge of technology. Nevertheless, the requirement remains for approaches which can give some indication of the cost of a space project before full commitment is made. Current space cost models employ cost estimating relationships (CERs) based on historical projects, which regress technical parameters of these past projects against the known cost of the past projects. This approach, the current state of the art, works marginally well. But the CERs typically have large variance, which leads to wide confidence intervals around any estimate.

Most conventional thinking on this subject has followed the line that this variance in the regressions has been due to the technical parameters not being sufficiently addressed in terms of fully describing the complexities of the projects. However, it is the thesis of this paper that some significant part of the variance in the historical cost of NASA projects could be caused by engineering management differences between the projects that is not being captured in the traditional cost models. Using previous work examining NASA's engineering management history (Hamaker, 1999) this paper will outline what engineering management culture changes have occurred over NASA's history and perform an exploratory analysis to attempt to determine if there is statistical evidence that improvements in management culture could be having an effect on project costs.

Background

The referenced previous work qualitatively assessed various engineering management initiatives throughout NASA's history in terms of their potential for improving productivity. The initiatives were categorized and examined in terms of their effectiveness on aerospace programs. The work surfaced three distinct eras of NASA engineering management philosophy as depicted in Exhibit 1: (1) The newly formed NASA of the 1960's (termed the "Apollo Era"), (2) the maturing NASA of the 1970's through the 1980's (termed the "Shuttle Era") and finally (3) the NASA of the 1990's attempting to restructure itself under the stewardship of the NASA Administrator Dan Goldin (termed the "Goldin Era"). For each of these eras, the previous study examined the 13 separate engineering management criteria (listed in exhibit 1) and characterized the NASA organization against these criteria.

Extension of Previous Work

This paper extends the work summarized in Exhibit 1. Here we will search for statistical evidence that the engineering management improvements claimed have provided any measurable reductions in NASA project costs. Eventually, this must be done at a very detailed level—i.e. at the individual project level by researching and rating each of the above 13 engineering management criteria for each project in the data base—see *Recommendations for Further Study* below. However, before undertaking the massive job of researching and rating such a large number of variables for a number of past projects, this paper performs the exploratory phase of the work by taking the shortcut of capturing the effects of engineering management improvement trends via the introduction of a time variable into the regression models used to predict NASA project cost. If there is any underlying improvement occurring in management over the 3 eras as discussed above, then it should be demonstrable with a negative slope on cost versus time. Because any cost decreases over time could just as well be attributable to a loosening of requirements on technical specifications for NASA projects (e.g. perhaps projects are being reduced in scale or complexity) the analysis will also include the necessary technical variables as well. Thus the inclusion of the technical variables can be used as a control and the analysis will search for cost improvements over time holding the technical requirements constant.

Exhibit 1: Engineering Management Criteria

Engineering Management Criteria	NASA (1960-1970)—Apollo Era	NASA (1970-1990)—Shuttle Era	NASA (1990-2001)--Goldin Era
1. Business Environment	Rapidly changing technology, short cycles; Technology push during development project; response time fast but only with heavy use of overtime; major projects took 5-7 years	Stable technologies (due to declining budgets) and long cycles; Less technology push and technology harvesting; response time slow; major projects took 8-10 years	Return to rapidly changing technology, short cycles; Use of precursor ground based and X vehicle technology maturation; response time more rapid; major projects take 3-5 years; small projects in <3 years
2. Organizational Form	Functional, specialized, mechanistic and centralized	Functional, specialized, mechanistic and centralized; matrix approach attempted to increase responsiveness	Functional fortresses much reduced, much less use of matrix; reorganized into product oriented teams
Use of Teams	Teams used but inefficient implementation	Team efficiency increased	Product development teams initiated
• Team classification	• Working groups	• Pseudo teams	• Potential/real teams
• Size	• Much larger than 2 to 12	• Much larger than 2 to 12	• ~5 to ~20 typical
• Diversity	• Little diversity	• Token diversity	• Noticeably more diversity
• Volunteer or Draft	• Appointed	• Appointed	• Appointees and volunteers
• Team Leader	• Project Manager typically	• Project Manager typically	• Team lead (non supervisor)
• Training	• No team training	• No team training	• Team training
• Performance Evaluation	• Project level evaluation, not team	• Project level, not team	• Some team performance evaluation.
3. Management Process, Decision Making, Vision and Values	Top management and project managers make decisions; nationally mandated goal used as vision	Top management and project managers make decisions and dictate direction with little emphasis on vision and values	Teams and team leads generate many solutions; management tends to policy issues much more
4. Chain Of Command and Communications	Formal, vertical, top down; freedom of information often restrictive, proprietary; need to know only;	Mostly vertical, top down but with some bottom to top, some horizontal cross functional communications; freedom of information more open but still a need to know mentality	Significant autonomy granted; Vertical channels augmented by horizontal cross functional communications, more communication with customer; heavy use of intranet and internet; freedom of information much more open but still pockets of need to know mentality
5. Job Descriptions	Detailed prescriptions	Detailed but with some flexibility	Less detail, more responsibility and authority given to employee
6. Span of Control/Support	7 employees for each supervisor (span of control)	10 employees for each supervisor (span of control)	14 employees for each supervisor (more span of support approach)
7. Valued Skills	Specialists	Specialists	More fostering of generalists
8. Training	Moderate quantity but mostly technical	High quantity, technical and management subjects	High quantity, adding diversity, safety, team building
9. Motivation and Awards	Formal, by quota, very delayed	Formal, by quota, very delayed	More team awards and more use of on the spot awards
10. Performance Appraisal	Tied to job description	Complicated system; tied to job description	Relatively simple paperwork; performance tied to organizational strategic plan
11. Policies, Procedures and Specifications	Inherited military systems which were extensive	Extensive policies and procedures tailored to NASA	Significantly rolled back
12. Supplier & Contractors	To be controlled	To be controlled	Frequently a partner
13. Customer Focus	Early customer focus on the Executive and Legislative branches of government and effective public relations with the taxpayer; success criteria includes performance, schedule (r variables) and reliability (p variable)	Same customer focus continued: success criteria include performance, cost control (r variables) and reliability (p variable); quality Circles attempted.	Consistent with TQM, more importance put on internal, intermediate customers; success criteria include performance, cost control (r variables), reliability but also customer satisfaction and retention, morale, rewards (p variables); product teams with QA providing support, training; emphasis on ISO 9000.

Methodology

The methodology for this initial exploratory analysis is to first develop a regression equation for predicting NASA space project cost using only technical parameters. Once a suitable technical parameter cost estimating (CER) equation has been derived, a time variable (as a proxy variable for engineering management improvement trends) will be introduced to determine if additional predictive power is observed and if the trend is negative with respect to time (i.e. a negative coefficient on time). If so, this will indicate that further research is merited to more precisely quantify the effect of the 13 engineering management variables on individual specific projects in the database.

Regression Analysis Using Technical Variables Only

The NASA NAFCOM (NASA-Air Force Cost Model) database (NAFCOM, 1999) was examined to obtain project level cost and technical metrics for a substantial number of historical NASA projects. The NAFCOM database is NASA's main repository of normalized cost and technical information on historical projects for use by the Agency's cost estimating community. Exhibit 2 provides the data that was extracted from the NAFCOM database for this analysis:

The data table includes variables that that have been observed to yield good predictor equations in the past and the additional variable for time (launch year). The variables are:

- Project cost, the dependent variable, expressed in millions of 1999 dollars and transformed to natural logs due to the wide variance) and abbreviated as *LnCost*).
- Dry spacecraft weight in pounds (transformed to natural logs due to the wide variance) and abbreviated as *LnWt.*
- Number of structural materials utilized (e.g. aluminum only =1, aluminum + titanium = 2, aluminum + titanium + composites = 3, etc.) and abbreviated as *NumMatls.*
- Number of deployable structures, using a dummy variable, where each *type* of deployable is counted as 1 (; e.g. 1 antenna + 1 solar array = 2, 3 antennas + 2 solar arrays = 2, 1 solar array + 1 antenna + 1 sensor boom = 3, etc.) and abbreviated as *NumDeploy.*
- Type of power generation, abbreviated as *Generate*, using a dummy variable where
  - None =1
  - Silicon solar arrays = 2
  - Gallium Arsenide solar arrays = 3
  - Fuel cells = 3
  - Radioactive Thermal Generators (RTGs) = 4
- Battery type, abbreviated as *Battery*, using a dummy variable, where:
  - None = 0
  - Nickel cadmium = 1
  - Silver zinc = 2
  - Nickel hydrogen = 3
  - Lithium Ion = 4
- Communications and Data Handling maximum data rate capability in kilobits per second (transformed to natural logs due to the wide variance) and abbreviated as *LnDataRate*
- Type of thermal control, abbreviated *Thermal* where
  - Passive = 1
  - Active = 2
- Type of Attitude Control, abbreviated *Control*, using a dummy variable, where
  - None = 0
  - Spin stabilized = 1
  - Despun section = 2
  - Gravity gradient = 2
  - 3 axis controlled = 3
- Type of Guidance, Navigation and Control sensors, abbreviated *Sensors*, using a dummy variable, where
  - None = 0
  - Sun sensors = 2
  - Earth horizon sensors = 3
  - Star trackers = 4
- Type of Reaction Control, abbreviated *Reaction*, using a dummy variable, where
  - None = 0
  - Monopropellant = 1
  - Bi-propellant = 2
  - Dual mode (mono-prop and bi-prop) = 3
- Human rated, abbreviated *Human*, using a dummy variable where
  - Not human rated = 1
  - Human Rated = 2
- Launch year, transformed to Launch Year less 1960 to convert to 2 a digit metric and abbreviated as *Year-1960.*

First, Best subsets regression was used to as an efficient way to select a group of promising CERs. The results from the best subsets regression are shown in Exhibit 3. The Cp statistic is used as a criterion where we look for models where Cp is small and is also close to p (where p is the number of parameters in the model including the intercept). If the model is adequate (i.e., fits the data well), then the expected value of Cp is approximately equal to p.

Exhibit 2. Database for Analysis

Project	Project Cost (1999$M) Cost	LnCost	Dry Wt. (lbs) Weight	LnWt	Number Struct. Malls *NumMalls*	Number Deploy *Struct.* NumDep	Type Power Generate *0=none* 1=Si 2=GaAs 3=Fuel cell *4=RTGs* Generate	Type of Battery *0=none* 1=NiCd 2=AgZn 3=NiH *4=Li Ion* Battery	Comm Data Rate *(kbps)* Data	LnData	*Type of* Thermal Control 1=Passive *2=Active* Thermal	Type of Attitude Control *0=None* 1=Spin 2=Despun 2=GG *3=3 Axis* Control	Type of *G&N* 0=None 1=Sun 2=Horizon *3=Star* Sensors	Type of *Reaction* Control 1=Mono 2=BiProp *3=Dual* Reaction	Human Rated 1=No *2=Yes* Human	*Launch* Year	Yr-1960
ACTS	535	6.283	2684	7.895	1	8	1	1	220000	12.301	2	3	2	1	1	1993	33
AE-3	69	4.228	780	6.659	1	0	1	1	131	4.875	2	2	2	1	1	1973	13
AEM-HCMM	18	2.889	185	5.220	1	2	1	1	8	2.079	2	3	2	1	1	1978	18
AMPTE	39	3.673	280	5.635	2	1	1	1	211	5.352	2	1	1	2	1	1984	24
ALEXIS	4	1.479	70	4.248	2	1	1	1	750	6.620	1	1	2	0	1	1993	33
Apollo CSM	9497	9.159	31280	10.351	2	3	3	1	51	3.932	2	3	0	1	2	1968	8
Apollo LM	6511	8.781	8071	8.996	2	4	0	2	51	3.932	1	3	0	1	2	1968	8
ATS-1	154	5.034	526	6.265	1	0	1	1	1	0.000	1	1	0	1	1	1966	6
ATS-5	198	5.288	755	6.627	1	1	1	1	1	0.000	2	1	2	1	1	1969	9
ATS-6	361	5.889	2535	7.838	3	5	1	1	1	0.000	2	3	2	1	1	1974	14
Centaur-D	1962	7.581	7674	8.946	1	0	0	2	1	0.000	1	3	1	1	1	1966	6
Centaur-G'	678	6.519	12920	9.467	2	1	0	2	1	0.000	1	3	3	1	1	1985	25
COBE	121	4.793	4320	8.371	1	6	1	1	4	1.386	2	3	2	1	1	1989	29
CRRES	81	4.393	2876	7.964	1	2	1	1	256	5.545	2	1	2	1	1	1990	30
DE-1	34	3.536	569	6.344	2	7	1	1	131	4.875	2	1	2	0	1	1981	21
ERBS	86	4.452	4493	8.410	1	3	1	1	128	4.852	2	3	2	1	1	1984	24
FAST	29	3.382	268	5.591	1	8	2	1	2250	7.719	1	1	2	0	1	1996	36
External Tank	985	6.893	74453	11.218	2	0	0	0	1	0.000	1	0	0	2	2	1981	21
Galileo Orbiter	1170	7.065	2755	7.921	2	6	4	4	134	4.898	2	2	3	2	1	1989	29
GEMINI	3219	8.077	7344	8.902	4	1	3	2	1	0.000	1	3	2	2	2	1965	5
GRO	413	6.022	13448	9.507	1	3	1	1	512	6.238	2	3	3	1	1	1991	31
HEAO-1	193	5.263	2593	7.861	1	1	1	1	6	1.792	2	3	3	1	1	1977	17
Hawkeye	9	2.177	58	4.060	3	2	1	2	0.2	-1.609	2	1	2	0	1	1974	14
IUS	1085	6.990	7227	8.886	3	0	0	2	64	4.159	1	3	2	3	1	1982	22
LANDSAT-1	185	5.219	1375	7.226	2	2	1	1	15000	9.616	2	3	3	1	1	1972	12
Lewis	30	3.395	463	6.138	2	5	1	3	2000	7.601	1	3	3	1	1	1997	37
Lunar Orbiter	411	6.019	394	5.976	3	7	1	1	50	3.912	1	3	3	2	1	1966	6
Lunar Prospector	30	3.399	327	5.790	2	3	1	3	3	1.099	1	1	1	1	1	1997	37
Lunar Rover	135	4.905	509	6.232	1	5	0	2	1	0.000	1	0	2	1	2	1971	11
Magsat	20	3.003	168.0	5.124	1	6	1	1	320	5.768	2	3	1	0	1	1979	19
Magellan	615	6.422	2554	7.845	1	2	1	1	268	5.591	2	3	3	1	1	1989	29
Mariner-10	294	5.682	936	6.842	2	11	1	1	117	4.762	1	3	3	1	1	1973	13
Mariner-4	329	5.796	516	6.246	2	4	1	2	16	2.773	1	3	1	1	1	1964	4
Mariner-6	563	6.334	705	6.558	2	4	1	2	16	2.773	2	3	3	1	1	1969	9
Mariner-8	425	6.052	1069	6.974	2	5	1	1	132	4.883	2	3	3	3	1	1971	11
Mars Global Surv.	117	4.766	1275	7.151	2	3	2	3	85	4.443	2	3	1	3	1	1996	36
Mars Observer	324	5.780	851	6.746	1	9	1	1	85	4.443	2	3	3	3	1	1992	32
Mars Pathfinder	184	5.214	1502	7.315	1	0	1	2	6	1.792	2	1	3	1	1	1996	36
NEAR	116	4.757	1480	7.300	1	5	2	1	26	3.258	2	3	3	3	1	1996	36
OMV	701	6.553	12811	9.458	1	6	2	2	972	6.879	2	3	2	3	1	1985	25
OSO-8	185	5.220	1037	6.944	3	1	1	1	128	4.852	2	2	3	1	1	1975	15
Pioneer Venus	114	4.735	760	6.633	5	5	1	1	2	0.693	2	1	3	1	1	1978	18
Pioneer-10	303	5.714	422	6.045	2	5	4	2	16	2.773	2	1	2	1	1	1972	12
SAMPEX	74	4.298	125.4	4.832	1	2	1	1	3000	8.006	2	3	1	0	1	1992	32
Shuttle Orbiter	12343	9.421	153522	11.942	4	8	3	0	192	5.257	2	3	3	3	2	1981	21
S-IC	3492	8.158	197363	12.193	1	0	0	2	1	0.000	2	3	0	2	2	1968	8
S-II	3723	8.222	71255	11.174	1	0	0	2	1	0.000	2	3	0	2	2	1968	8
S-IVB	1893	7.546	32419	10.386	1	0	0	2	1	0.000	2	3	0	2	2	1968	8
Skylab OWS	2507	7.827	68001	11.127	1	2	1	0	50000	10.820	2	0	0	1	2	1973	13
SMS-1	140	4.939	406	6.006	3	0	1	1	1	0.000	2	1	2	1	1	1974	14
SRB	691	6.538	37698	10.537	1	2	0	0	1	0.000	1	3	0	0	2	1981	21
SSM	800	6.685	11233	9.327	3	4	1	3	1024	6.931	2	3	1	2	2	1990	30
Surveyor	1287	7.160	647	6.472	1	5	1	2	1	0.000	2	3	1	2	1	1966	6
SWAS	35	3.554	373	5.922	1	5	2	1	1800	7.496	1	3	3	2	1	1995	35
TDRSS	667	6.503	3391.1	8.129	4	4	1	1	300000	12.612	2	3	2	1	1	1983	23
TIROS-M	128	4.852	435.1	6.076	1	3	1	1	2	0.693	2	2	2	0	1	1970	10
TIROS-N	64	4.152	836.0	6.729	3	3	1	1	2662	7.887	2	3	2	0	1	1978	18
TOMSEP	56	4.017	432	6.068	1	2	1	1.5	202	5.308	1	3	2	1	1	1996	36
TOPEX	201	5.304	3154	8.056	1	3	1	1	1024	6.931	2	3	3	1	1	1992	32
UARS	303	5.714	7385	8.907	3	3	1	1	512	6.238	2	3	3	3	1	1991	31
Viking Lander	1315	7.181	1908	7.554	2	6	4	1	16	2.773	2	3	0	1	1	1975	15
Viking Orbiter	886	6.787	1941	7.571	2	5	1	1	2116	7.657	2	3	3	3	1	1975	15
VOYAGER	502	6.218	1410	7.251	2	5	4	1	115	4.745	2	3	3	1	1	1977	17
X-34	177	5.176	13925	9.541	2	3	0	1	1	0.000	1	3	0	0	1	1999	39

A small value of Cp indicates that the model is relatively precise (has small variance) in estimating the true regression coefficients and predicting future responses. Models with considerable lack of fit have values of Cp larger than p. Using the C_p criteria, all regressions below with $C_p \leq$ (Vars +1) are acceptable (Vars + 1 because the *Vars* value shown in the statistical output does not count the constant and thus 1 must be added to Vars to equal p). All equations with 4 or more variables are indicated by the C_p criterion to be promising.

Exhibit 3: Best Subsets Regression.

```
Response is LnCost

                                          N N G   L       R
                                          u u e B n T C S e
                                          m m n a D h o e a
                                          M D e t a e n n c H
                                        L a e r t t r t s t u
                                        n t p a e a m r o i m
                Adj.                    W l l t r R a o r o a
Vars    R-Sq    R-Sq    C-p         s   t s o e y a l l s n n

   1    63.4    62.8   13.9    1.0168   X
   1    36.3    35.2   68.8    1.3420                       X
   2    67.5    66.4    7.7   0.96620   X     X
   2    66.2    65.1   10.2   0.98453   X                 X
   3    69.2    67.7    6.1   0.94763   X     X           X
   3    69.0    67.5    6.6   0.95127   X     X       X
   4    71.2    69.3    4.1   0.92396   X     X       X     X
   4    70.7    68.7    5.3   0.93348   X     X           X X
   5    72.7    70.3    3.2   0.90862   X     X       X   X X
   5    71.9    69.4    4.8   0.92180   X X   X       X     X
   6    73.3    70.5    4.0   0.90609   X     X   X   X   X X
   6    73.1    70.3    4.4   0.90948   X X   X       X   X X
   7    73.7    70.4    5.1   0.90700   X X   X   X   X   X X
   7    73.5    70.2    5.5   0.91033   X   X X   X   X   X X
   8    73.9    70.1    6.7   0.91174   X X X X   X   X   X X
   8    73.9    70.1    6.8   0.91218   X X   X   X   X X X X
   9    74.2    69.8    8.2   0.91571   X X X X   X   X X X X
   9    74.0    69.7    8.5   0.91809   X X X X X X   X   X X
  10    74.2    69.4   10.0   0.92264   X X X X X X   X X X X
  10    74.2    69.3   10.2   0.92412   X X X X   X X X X X X
  11    74.3    68.8   12.0   0.93110   X X X X X X X X X X X
```

Because C_p indicated that all equations with 4 or more variables are promising, a regression was first performed with all 10 variables. The results of that analysis, shown in Exhibit 4, indicate that the CC&DH data rate (abbreviated as *lnDataRate*) and the type of GN&C sensors (abbreviated as *Sensors*) have negative slopes (i.e. increasing the requirement results in lower cost). This result is nonsensical from an engineering point of view. Thus these two variables were dropped form the regression. Also, from Exhibit 4, it can be observed from the value of the p coefficients that four other variables have poor significance. These include the variables *Thermal, NumDeploy, Battery and NumMatls*. These variables were actually carefully excluded one at a time while observing the impact on the overall R squared statistic and the value of the other individual predictor p values. As each variable was dropped, previous variables that had been dropped up to that step were re-introduced. This re-introduction included *LnDataRate* and *Sensors* to see if the sign reversal problem corrected itself but this was not the case. In no cases were the re-introduced variables sustained in terms of improving the CER.

In the end, the best subset CER selected, shown in Exhibit 5, included 5 variables: *Dry LnWt, Generate, Control, Reaction* and *Human*. This is actually an adequate CER for space missions with an R squared

Exhibit 4. Full Variable Regression Analysis.

```
The regression equation is
LnCost = - 0.888 + 0.549 LnWt + 0.135 NumMatls + 0.0445 NumDeploys
+ 0.265 Generate + 0.077 Battery - 0.0418 LnDataRate + 0.057 Thermal +
0.265 Control - 0.101 Sensors + 0.230 Reaction + 0.695 Human

Predictor        Coef       StDev          T        P        VIF
Constant      -0.8877      0.9047      -0.98    0.331
LnWt           0.5485      0.1014       5.41    0.000        2.6
NumMatls       0.1348      0.1301       1.04    0.305        1.1
NumDeplo      0.04449     0.05509       0.81    0.423        1.5
Generate       0.2654      0.1379       1.92    0.060        1.4
Battery        0.0768      0.1679       0.46    0.649        1.2
LnDataRa     -0.04177     0.03960      -1.05    0.296        1.3
Thermal        0.0573      0.2800       0.20    0.839        1.2
Control        0.2645      0.1390       1.90    0.063        1.3
Sensors       -0.1006      0.1455      -0.69    0.492        1.8
Reaction       0.2298      0.1557       1.48    0.146        1.4
Human          0.6954      0.4782       1.45    0.152        2.6

S = 0.9311      R-Sq = 74.3%     R-Sq(adj) = 68.8%

Analysis of Variance

Source            DF          SS          MS         F         P
Regression        11     130.104      11.828     13.64     0.000
Residual Error    52      45.081       0.867
Total             63     175.185
```

Exhibit 5. Final Best Subset of Technical Variables Regression Equation.

```
The regression equation is
LnCost = - 0.834 + 0.526 LnWt + 0.302 Generate + 0.269 Control + 0.244
Reaction + 0.943 Human

Predictor        Coef       StDev          T        P        VIF
Constant      -0.8344      0.5618      -1.49    0.143
LnWt          0.52603     0.09383       5.61    0.000        2.3
Generate       0.3024      0.1189       2.54    0.014        1.1
Control        0.2688      0.1300       2.07    0.043        1.2
Reaction       0.2441      0.1407       1.73    0.088        1.2
Human          0.9427      0.4153       2.27    0.027        2.0

S = 0.9086      R-Sq = 72.7%     R-Sq(adj) = 70.3%

Analysis of Variance

Source            DF          SS          MS         F         P
Regression         5     127.300      25.460     30.84     0.000
Residual Error    58      47.884       0.826
Total             63     175.185
```

correlation of 70.3% (acceptably good for space cost CERs), with an F test p value of 0.000 indicating that the overall equation is acceptable and with all independent variable p values less than 0.10 (indicating a confidence level of better than 90% that these variables do drive cost).

Introduction of a Time Variable

At this point, armed with a CER that fairly adequately predicts the cost of space missions, the time variable, abbreviated as *Years-1960*, is introduced into the regression to see if the predictive power of the resulting CER is an improvement. The results of that regression are shown in Exhibit 6.

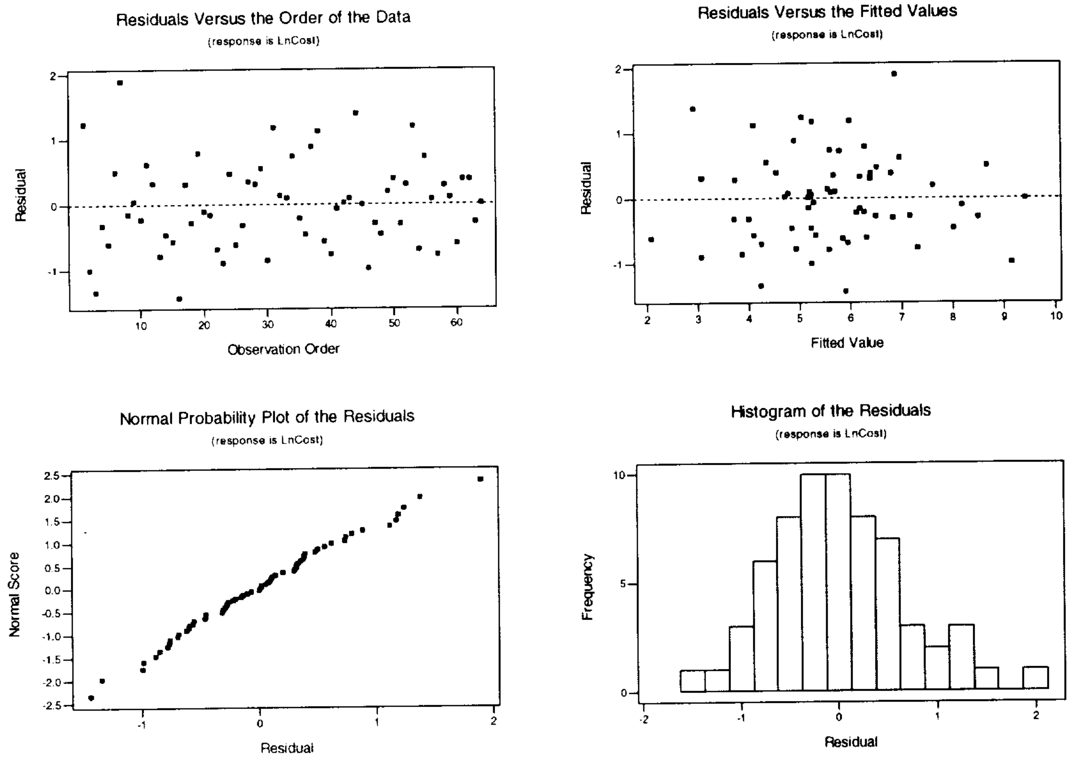
Exhibit 6. Regression Analysis With Technical Variables and Time.

```
The regression equation is
LnCost = 0.693 + 0.592 LnWt + 0.311 Generate + 0.228 Control + 0.261 Reaction
+ 0.247 Human - 0.0559 Year-1960

Predictor         Coef        StDev          T         P        VIF
Constant        0.6927       0.5053       1.37     0.176
LnWt           0.59244      0.07424       7.98     0.000        2.4
Generate       0.31121      0.09309       3.34     0.001        1.1
Control         0.2276       0.1020       2.23     0.030        1.2
Reaction        0.2609       0.1102       2.37     0.021        1.2
Human           0.2467       0.3443       0.72     0.477        2.3
Year-196     -0.055879     0.009108      -6.14     0.000        1.1

S = 0.7113        R-Sq = 83.5%      R-Sq(adj) = 81.8%

Analysis of Variance

Source            DF          SS          MS          F         P
Regression         6     146.346      24.391      48.21     0.000
Residual Error    57      28.839       0.506
Total             63     175.185
```

Exhibit 7. Regression Analysis With Technical Variables (Less Human) and Time.

```
The regression equation is
LnCost = 0.808 + 0.629 LnWt + 0.309 Generate + 0.205 Control + 0.258 Reaction
- 0.0580 Year-1960

Predictor         Coef        StDev          T         P        VIF
Constant        0.8081       0.4769       1.69     0.096
LnWt           0.62932      0.05327      11.81     0.000        1.2
Generate       0.30949      0.09267       3.34     0.001        1.1
Control        0.20531      0.09676       2.12     0.038        1.1
Reaction        0.2576       0.1096       2.35     0.022        1.2
Year-196     -0.058029     0.008563      -6.78     0.000        1.0

S = 0.7083        R-Sq = 83.4%      R-Sq(adj) = 82.0%

Analysis of Variance

Source            DF          SS          MS          F         P
Regression         5     146.086      29.217      58.24     0.000
Residual Error    58      29.099       0.502
Total             63     175.185
```

Exhibit 8. Analysis of Residuals



Interestingly enough, this regression yields a CER with an improved R squared value of 81.8% (significantly better than the CER with technical variables only which had an R squared of 70.3%). The variable *Year-1960* is seen to be very significant with a p value of 0.000 and with a negative slope indicating improvements in cost over time. Though very heartening, a more careful examination of Exhibit 6 also indicates that one variable, Human, now has a troubling p value of 0.477. The regression was thus run again, dropping Human to see if the CER is improved. The results of this are shown in Exhibit 7.

Omitting *Human* results in a better CER since it has fewer variables, a slightly higher R squared and essentially unchanged p values for the independent variables. It can also be argued that several of the remaining variables are explaining the contribution of *Human* including weight (*LnWt*) because human rated projects tend to have larger weights, *Generate* (via the fuel cells selection because only human missions have used fuel cells, and *Control* (since all Human missions have been 3 axis controlled. So it could well be that the variable *Human* is not offering any additional explanatory power once the above variables are included. As a final check of this regression, an analysis of residuals was performed resulting in the graphs in Exhibit 8.

As can be observed, all of the residual plots are acceptable. The residuals versus order of the data shows no non-random time related effects in the residuals. The residual versus fitted values shows the desired random pattern on either side of zero with no reason to suspect that the error is not random. The normal probability plot of the residuals generally forms a straight line indicating normally distributed residuals. Finally, the histogram of the residuals has essentially the desired bell curve shape around a mean of zero.

Conclusions and Observations

The final regression equation of

```
LnCost = 0.808 + 0.629 LnWt + 0.309
Generate + 0.205 Control + 0.258
Reaction - 0.0580 Year-1960
```

indicates, through the negative coefficient on the variable *Year-1960*, that cost is decreasing each year for reasons in addition to the weight, type of power generation, type of attitude control or type of reaction control. Because the previous regression without *Year-1960* did a creditable job of predicting cost, the hypothesis that there are some other effects beyond these major technical variables that are contributing to a decreasing cost trend cannot be rejected. This analysis does not, of course, prove that these other effects are due to engineering management factors. However, the initial qualitative analysis presented in this paper at least suggests that engineering management improvements and could be the reason for improved cost effectiveness.

Recommendations For Further Study

It is recommended that further study be undertaken to collect additional engineering management data on specific past NASA projects in the database used here. This data could then be used to rank the individual projects leading to a quantitative scale for individual projects over time. Further regression analysis could then be accomplished to determine if the engineering management factors are a statistically significant driver in the cost of space projects.

The major obstacle to such a continued study will be the research required to obtain credible and documented information on engineering management culture sufficient for a NASA confident quantitative ranking. Preliminary research by the author into the NASA archives has turned up considerable documentation on the management practices of individual projects. It is for this reason that the author has some confidence that such a study is feasible.

The benefits of introducing engineering management variables into the mix of independent variables used for space systems cost analysis would be significant. These additional variables help reduce residual variance in the CERs and cost models. It would also serve to highlight the importance of engineering management practices on the cost outcome of projects and serve to give project managers quantitative information about the likely outcome of choices they have in how to manage their projects.

References

Hamaker, Joseph W., *The Faster, Better, Cheaper Approach To Space Missions: An Engineering Management Assessment* (1999)

NASA/Air Force Cost Model 99 (NAFCOM99), Version 6.1, Science Applications International Corporation, Huntsville, Alabama, (June 2000)

About the Author

Joseph W. Hamaker received his M.S. in Engineering Management from the University of Alabama in Huntsville (UAH). He holds a B.A. degree in Economics from UAH and a B.S. degree in Industrial Engineering from Tennessee Technological University. He has completed course work for the Ph.D. program in Engineering Management at UAH and is currently working on his dissertation. He is employed by the NASA Marshall Space Flight Center in Huntsville, AL where he is Manager of the Engineering Cost Office.