Proceedings of the 2004 Winter Simulation Conference
R .G. Ingalls, M. D. Rossetti, J. S. Smith, and B. A. Peters, eds.

DEVELOPMENT OF A PARMETRIC ESTIMATING MODEL FOR TECHNOLOGY-DRIVEN DEPLOYMENT PROJECTS

Rudy J. Watson
Young Hoon Kwak

School of Business
The George Washington University
Washington, DC 20052, U.S.A

ABSTRACT

The purpose of this paper is to define the process used to develop a parametric estimating model and to explain the use of the model as it allows a non subject matter expert to predict the cost of deployment of a technology-driven project with improved accuracy.

The overall objective is to define criteria for organizations to use as a rule of thumb estimating model for determining potential resource requirements during the pre-concept and concept phase of a technology-driven project.

The use of the model will determine the complexity level of a project and then using a matrix simulating the task complexity knowledge of a deployment subject matter expert, the organization deployment resource required will be estimated based upon responses to certain inquiries. The user of the model will then have the capability to modify the output to derive their specific deployment estimates.

1 INTRODUCTION

The parametric estimating model has its basis in learning curve theory first described in 1936 by T.P Wright. He proposed that as personnel gained experience at performing manufacturing tasks, less time was required to perform these tasks. He observed that the learning effect was not linear but appeared to have a constant decrease in the time required to perform the task for each time the experience level doubled

$$T_n = T_1 * N^{(\ln\%/\ln 2)} .$$

Therefore the time involved to perform the task at n^{th} time is a function of the time it took to perform the task the first time and the amount of learning that took place as the result of a doubling of the experience. Taking this theory into account, the estimates derived from the parametric estimating model incorporates decreases in the time required to perform tasks if they are performed by experienced personnel. This phenomena is captured in the assessment and designation of project and task complexity by the user of the model and is embedded in the resulting project estimates.

2 BENEFIT

Parametric estimating is the use of a subset of independent variables to predict the cost of a project. These independent variables may be specifications, features, function, or some other high level descriptive element that is used to define the scope of the deliverables at an early stage when there is a lack of detailed information.

Parametric cost estimating is widely used for bidding on a contract, input into a cost benefit analysis for selection within a project portfolio, and as the initial stage of building a plan for project implementation. Extensive literature reviews suggest that effective parametric estimating methodology is becoming an essential tool for technology-driven organizations. The use of parametric estimating in budgeting, scheduling, and control of projects will enhance the ability of project management organizations to effectively and efficiently utilize valuable resources.

Information about the model criteria must be available at pre-concept or project concept phase time. Considerations for complexity criteria include the number of stakeholder organizations, the number of physical locations, numbers of local languages involved, the function being deployed and duration of the project. Although this information is limited, it can be used within the model by a user simulating subject matter experts to produce an acceptably accurate estimate at a reasonable cost.

The majority of the literature reflected similar benefits associated with parametric estimating. The primary benefit is the ability to provide a timely estimate based upon limited knowledge of the project as long as sufficient information is available. This facilitates quicker response to competitive business environments. This is possible because the Cost Estimating Relationships are based upon actual historical data from related projects which reflects the impacts of cost growth, schedule changes and design changes. Statistical techniques are used to validate and maintain the relationships. The result of going through an estimation process provides greater insight and understanding of the major cost drivers of a particular project.

3 METHODOLOGY

3.1 Selection of Cost Estimating Relationships

The first step in the approach was to determine the cost estimating relationship (CER) variables. Research was done to identify grounding theories on the prediction of project costs. This resulted in a number of independent variables that were identified as being potentially useful and captured in Table 1. The continuation of this step was the selection and validation of the CERs to be used in the model.

Table 1: Cost Estimating Relationship Variables

Number of different user organizations
Application
Process
Number of new interfaces
Project Duration
System Architecture

Table 2 reflects the premise of the model, that the cost of the deployment of a technology-driven project is a function of the influence of project complexity and task complexity as intervening variables on the independent variables represented as project characteristics defining the scope of the tasks to be accomplished.

Table 2: Parametric Estimating Model

C	= technology-driven deployment cost estimate
P_t	= project complexity at time t
$T_{Xi(t)}$	= task complexity of specific task X_i at time t
X_i	= specific task estimate for task category i
n	= total number of specific tasks

$$\mathbf{C = f\,(P_t * \textstyle\sum_{i=(1-n)} X_i * T_{Xi(t)}\,)}$$

Project managers with extensive experience on the deployment of technology-driven projects were interviewed. Initially they were interview individually and then in joint sessions. These interviews were held to gather their input on the applicability of the existing independent variables and the identification of additional ones.

3.2 Determination of Estimates

The second step in the approach was to determine the appropriate values for the dependent variable based upon the values of the CERs. This was primarily derived from historical data from previous similar projects. The project managers were asked to provide data on the following:

- Information on the number of resources used on previous deployment projects summarized by deployment milestones.
- Suggestions on criteria that should be used to categorize projects in to high/med/low complexity. This criteria must be available at pre-concept phase.
- Review and agreement on the model after it is developed.
- Names of anyone else that should be involved in the assessment and development of the model.

The information in Table 3 contains the descriptions of the tasks that were ultimately selected as dependent variables for the model, the sum of which provides input to determine the total deployment cost estimate. Table 4 identifies the project milestones at which time cost data was captured from historical projects as input into the model.

Table 3: Estimated Task Categories - X_i

TM	Transformation Management
PM	Project Management
DA	Data Conversion and Migration
PR	Process Management
CO	Communications Management
IT	Information Technology
JD	Job Design
MS	Measurement Systems
TST	Testing
ET	Education and Training
ASCA	Applications Systems Control Certification
MTP	Move to Production

4 LIMITATIONS

4.1 Data Availability

The reason for or use of the estimate may drive the method utilized; however, the method is also often driven by the availability of data. Since parametric estimating is done

Table 4: Deployment Milestones

TM$_1$	An organization readiness assessment is completed
PM$_1$	A detailed project plan has been developed
DA$_1$	The integrated data sourcing and migration plan is completed
TM$_2$	A transformation blueprint, with aligned/ integrated activities is completed
TST$_1$	Tests cases ready to start
DA$_3$	Conversion and migration of data is ready to commence
TST$_2$	User acceptance tests completed
ASCA$_2$	Application System Control and Auditability certification obtained
ET$_2$	A plan is ready for the development of organization specific materials
PM$_2$	Process and system deployment is completed

early in the life of a project, detailed information about design is normally not available. For transportation projects, Harbuck (2002) surmised that the available information is in the range of 5-20%. Although not explicitly stated, the literature gives the impression that this upper range is high for technology related projects. Even in situations where significant data is available, Staub-French et al. (2003) speculates that the quantity-based CER's are not completely adequate because the estimator's rationale for the relationship is not captured.

In the absence of complete information, Figure 1 shows the steps can be taken to improve the parametric estimating process with technological aids and thereby increase the accuracy of the estimates. It begins with the capture and retention of historical information. By improving the ability to access and manipulate the cost information stored in databases and facilitating the application of algorithms and statistical validation, the resulting reporting of information can be utilized to generate more accurate estimates.

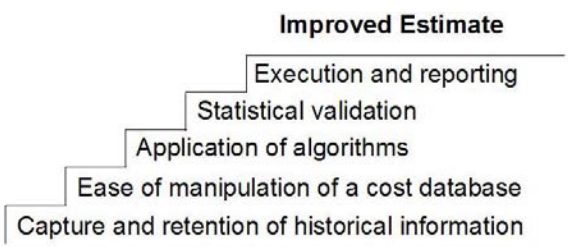

Figure 1

4.2 Accuracy

Oberlender and Trost (2001) identified four determinants of estimate accuracy:

- Who was involved in preparing the estimate
- How the estimate was prepared
- What was known about the project
- Other factors considered while preparing the estimate

He also describes the accuracy of parametric cost estimates as being affected by:

- Changes in project scope
- Changes in design standards
- Incorrect unit cost/quantity assumptions
- Unforeseen problems in implementation

Standardization and consistency are also important factors in deriving accurate estimates. Table 5 indicates three approaches put forth by Black, Bajaj and Koenigseker for providing consistency in the estimates derived from parametric estimating models.

Table 5: Approaches for Consistency

Black (1984)	Bajaj, Gransberg, and Grenz (2002)	Koenigseker (1982)
Definition of the problem	Effective cost accounting	Develop standards
Collection of the data	Historical database	Develop parametric sets
Normalization of the data	Normalize data - multipliers	Define the project
Interdependencies	Exclude outliers	Apply sets to standards
Derivation of the CER	Statistical analysis - CERs	
Establishment of limitations	Apply multipliers	
Documentation	Prepare a template	

5 USING THE MODEL

5.1 Determine Project Complexity

Each deploying organization completes a template for a specific project. The purpose is to determine which resource template the model will use for the estimates in conjunction with the task complexity assessments. The project complexity is designated as H (High), M (Med), or L (Low) in response to the individual questions / topics.

If three or more criteria are identified as high, the over all project is high complexity. If less than three are identified as high and three or more are identified as medium, the overall project is medium complexity. It two or less criteria are rated high, and two or less are medium and five or more are low, the overall project "may" be considered low complexity at the discretion of the deploying organization. The criteria definitions are listed below.

The number of organizations implies the number of unique stakeholders being deployed within this project. Although each deploying organization is completing their own estimate, this question should be answered if it is known that another user organization will be deploying as a part of the same project. That would increase the complexity and subsequent resource requirements for an individual organization due to the dependencies and coordination involved.

For application, answer high complexity if this is the first deployment of the application within this specific organization. Answer medium if this is the deployment of enhancements to an already installed application within this user organization. Answer low if it is a rollout of an existing application to a new organization or group of new users within an already deployed organization without any major enhancement to function.

The process question applies to the particular organization using the model. High is it the first deployment of this process in the organization. A Medium response reflects changes to the existing process already deployed. Low is the roll out of an existing (previously deployed to this organization) process to additional users with no enhancements to the process.

The interface question relates to the development of local interfaces (organization specific) outside of the common interfaces within the project deployment scope assuming the project is being deployed simultaneously to multiple user organizations.

Project duration is the timeframe inclusive of Concept through Rollout phase activities.

If the supporting architecture includes E-business, On-Demand, or external customer access, the project is high complexity. If the users are internal or certain business partners with direct access (not web), the architecture criteria can be considered medium or low based on assessment of the deploying organization.

5.2 Determine Task Complexity

Task complexity is also determined by responding to specific questions with H, M, or L. After each question is answered for a task category, determine the overall complexity for the category based on the responses to the individual questions. The definitions of the estimated task categories are below.

For Transformation Management, the task is high complexity if the organization has never deployed transformation management for this application for any users. The task is medium if an organization has previously deployed transformation management for users although some estimation of quality should apply. If a geo has deployed Transformation Management but it was not successful or lessons learned determined that it was not done with "good" quality, this question should be answered as if this is the first deployment which is high complexity. It is assumed that if an organization has deployed two or more user groups, even if they were not of high quality, the complexity of a third deployment is low.

Use of common centralized project resources across unique user organizations to perform transformation management activities for the deployment will lessen the complexity of the task. Using common centralized project resources for Transformation Management should lessen the risks and leverage the experience of the team, therefore answering yes to this question places the task at low complexity. Using individual user organization resources for this task (particularly inexperienced resources), results in this task being rated as high complexity.

The criteria for project management is binary. If the project manager is experienced on similar projects, the task is low on complexity. If the project manager is new or inexperienced on similar technology-driven projects, the task is rated high on complexity.

Related to Data Conversion and Migration, using a centralized set of skills is a factor for rating this task as low complexity. The is due the reuse of experienced skills. If a deployment project is not using a centralized set of skills and relying on the individual user organizations, the complexity of the task should be rated high. The implication is that more resources may be needed.

The volume of data to be converted is also a factor. The size of the data refers only to master data and operational data such as customer records and customer inventory. The model may be updated in the future to reflect a range of numbers of records to determine if the task is high or low complexity.

Process Management is related to whether an organizations has experienced a previous deployment. The task is high complexity if the organization has never deployed process management for this application for any user group. The task is medium if an organization previously deployed process management in a user group with some estimation of quality being applied. If a organization has deployed Process Management but it was not successful or lessons learned determined that it was not done with "good" quality, this question should be answered as if this is the first deployment which is high complexity. It is assumed that if an organization has deployed two or more user groups, even if they were not of high quality, the complexity of a third deployment is low. If the process documentation has to be converted to a national language, the complexity of the task is high.

Communications Management is determined to be a high complexity task if the organization has never deployed communications management for this application for any user group. The task is medium if an organization has previously deployed communications management in a user group with some estimation of quality applied. If the communications documentation has to be converted to a national language, the complexity of the task is high

Information Technology is represented by criteria related to workstation replacement, printers, scanners and other infrastructure. If the required desktop infrastructure exists and no changes are required, the task is low complexity. This refers to modifications to desktop hardware and software as well as printer hardware configurations. If some minor changes are required the task is medium If the infrastructure requires substantial change, the task is rated high complexity. The same rating structure is applied to the network infrastructure.

If the job design templates have already been rolled out prior to this project, the task complexity is low. If the job design templates are rolled out concurrent with this project, the complexity is high.

If the standard measurement system that "comes" with the application is sufficient, the task complexity is low. If minor changes/enhancements are required to meet organization or user unique requirements, the task complexity is medium. If significant change to the measurement system is required or a new measurement system is being implemented concurrent with this project, the task complexity is high.

The use of a central test team will make the complexity of the task rated as low because the central test team will bring experience and reuse of skills. If organization or user group is building its own test team, the complexity is high. A large number of organization or user "unique" test cases will make the task complexity high. If the test cases are common regression test cases, a large number of them would not necessarily make the task complexity high. It would be high only if the test cases are unique to this particular project. This model may be updated in the future to provide an example of the number of test cases to use as a reference point. If this is a new deployment, the influence on the complexity of the testing task ranking is high. If it is an upgrade to a previously installed application, the complexity is low. An "upgrade" could be considered as a new deployment if there was a tremendous change in function.

The range of users to be educated affects the task complexity. If 35 or more users are to be educated, the complexity is high. If 10 or less users are to be educated, the complexity is low. Otherwise the complexity is medium. If the education material has to be translated into a national language, the complexity is high.

Relative to an Application Systems Control and Audit ability (ASCA) review, the task complexity is high for a full ASCA review. The task complexity is low for a delta ASCA review (review of changes made since the last ASCA review) if the changes being reviewed do not include any changes to control points. The task complexity is medium if a delta review includes changes to control points. This is due to the additional testing and documentation that may be required. A task is also consider medium if the delta review is for rollout to additional user groups which have not been previously deployed. The first deployment in a organization normally requires a full ASCA review. Even if it a delta review is performed, this task should be considered high complexity for the first deployment.

Examples of size of move to production data may be included in updates to this model. If the data volume is large (eg., cannot be done in one weekend), the complexity is high. If this is the first deployment of this application by this organization, the complexity is high. If this move to production is adding a new user group (regardless of whether a previous deployment has occurred in that organization), this task criteria is rated high complexity. If this move to production is an upgrade to a previously deployed country, the task criteria is rated low.

5.3 Simulate Estimates of Subject Matter Experts

The combination of responses into the model results in an estimate determined by historical data and subject matter expert input gathered in interviews during the creation of the model.

The user of the model is provided the opportunity to over ride the estimates generated by the model. The rational and assumptions for the overrides should be documented. Figure 2 is an example of the task complexity output estimates with the overrides captured.

Deployment Resource Estimates

	# of hours	Adjustments to hours
TM - Transformation Management	100	0
PM - Project Management	800	100
DA - Data Conversion and Migration	500	300
PR - Process Management	300	-200
CO - Communications Management	100	0
IT - Information Technology	500	-100
JD - Job Design	100	-80
MS - Measurement Systems	50	0
TST - Testing	4200	-1500
ET - Education and Training	500	-150
ASCA - ASCA Certification	150	-100
MTP - Move to Production	800	100
Total Hours	**8100**	
Adjusted Hours		**-1630**
Total Resource Hours		**6470**

Figure 2

5.4 Compare Estimates to Acceptable Ranges

The final step in the use of the model is to compare the resulting estimates to an acceptable range for percentage of total deployment costs by category. It shows the model estimates and user overrides. Figure 3 is a representation of this output. The user of the model must go back to the previous step if additional overrides are determined to be required.

Deployment Resource Estimates

Category	Estimates	Adjusted Hours	Total Hours	Total %	Recommended % of total
TM	100	0	100	1.55%	>0 - 4
PM	800	100	900	13.91%	8-12
DA	500	300	800	12.36%	3-15
PR	300	-200	100	1.55%	5-15
CO	100	0	100	1.55%	>0 - 2
IT	500	-100	400	6.18%	4-10
JD	100	-80	20	0.31%	>0 - 4
MS	50	0	50	0.77%	>0 - 2
TST	4200	-1500	2700	41.73%	20-35
ET	500	-150	350	5.41%	1-6
ASCA	150	-100	50	0.77%	>0 - 2
MTP	800	100	900	13.91%	5-15

Figure 3

6 CONCLUSIONS

Competitive advantage is based upon the exploitation of the core competencies of an organization and its ability to optimize the effective and efficient use of its resources. The introduction and increased investment in technology must be a part of an overall strategic plan to ensure that it makes a sustained contribution to the competitive advantage of an organization. As increasingly more projects are integrated by technology components, the ability to estimate total costs and plan the efficient use of resources becomes more complex.

A literature review was conducted to assess the current thinking along these lines. From the review, we conclude that parametric estimating techniques are a way to facilitate effective estimates given limited but sufficient information. Overall, various literature suggest that organizations are finding great benefit in the investment in parametric estimation methodologies and tools. Much of the literature focused on the development and validation of Cost Estimating Relationships. There is a consensus on its importance however, there is still quite a bit of on-going research remaining to be done in this area as the introduction of new technology changes the characteristics of projects.

To summarize, the past few years has yielded greater understanding and use of parametric estimation methodologies and tools. This is leading to the ability to simplify the complexities of technology driven projects so that the major cost drivers can be identified, validated and maintained. The use of these estimates in budgeting, scheduling and control of projects will enhance the ability of project management organizations to effectively and efficiently utilize value resources.

The use of parametric estimating models which captures the knowledge of subject matter experts allows a non subject matter expert to simulate the generation of project estimates with acceptable accuracy at a reasonable cost.

REFERENCES

Bajaj, A., D.D. Gransberg, and M.D. Grenz (2002). Parametric Estimating for Design Costs. AACE International Transactions, EST.08, ES81.

Black, J.H. (1984). Application of Parametric Estimating to Cost Engineering. AACE Transactions, B-10, B.10.1.

Harbuck, R.H. (2002). Using Models in Parametric Estimating for Transportation Projects. AACE International Transactions, EST.05, ES51.

Koenigseker, N. (1982). Parametric Estimating of Buildings. Cost Engineering. Vol.24 Issue 6. 327-332.

Oberlender, G. D. and S. M. Trost (2001). Predicting Accuracy of Early Cost Estimates Based on Estimate Quality. Journal of Construction Engineering and Management. Vol. 127 Issue 3. 173-182.

Staub-French, S., M. Fischer., J. Kunz, and B. Paulson (2003). An Ontology for Relating Features with Activities to Calculate Costs. Journal of Computing in Civil Engineering. Vol. 17 Issue 4. 243-254.

AUTHOR BIOGRAPHIES

YOUNG HOON KWAK, Ph. D. is a faculty member of the Project Management Program at the Management Science Department at the George Washington University (GWU) in Washington, DC. He received his B.S (1991) in Civil Engineering from Yonsei University in Seoul, Korea, and M.S. (1992) and Ph.D (1997) in Engineering and Project Management from the University of California, Berkeley. Before joining GWU, he taught at the Florida International University in Miami and was a post doctoral scholar at the Massachusetts Institute of Technology. Dr. Kwak's research appears in over 40 papers, articles, book chapters and technical reports. He is currently serving as a member of the Editorial Review Board for the Project Management Journal. His major interests include project management and control, project risk management, construction management, technology management and international project management. His e-mail address is <kwak@gwu.edu> and his web address is <http://home.gwu.edu/~KWAK>.

RUDY J. WATSON is a doctoral candidate in Logistics, Technology, and Project Management at the Management Science department at The George Washington University in Washington, DC. He received his B.B.A (1975) in Information Processing, M.B.A (1979) in Information Systems Technology and M.S. (1999) in Project Management from The George Washington University. He has over thirty years of broad and diverse experience in information technology including twenty-seven years with IBM. He is a Project Management Institute Certified Project Management Professional. His research interests include educational institution effectiveness, knowledge management, project management, and project cost estimating. His e-mail address is <rudy.watson@gwu.edu>.