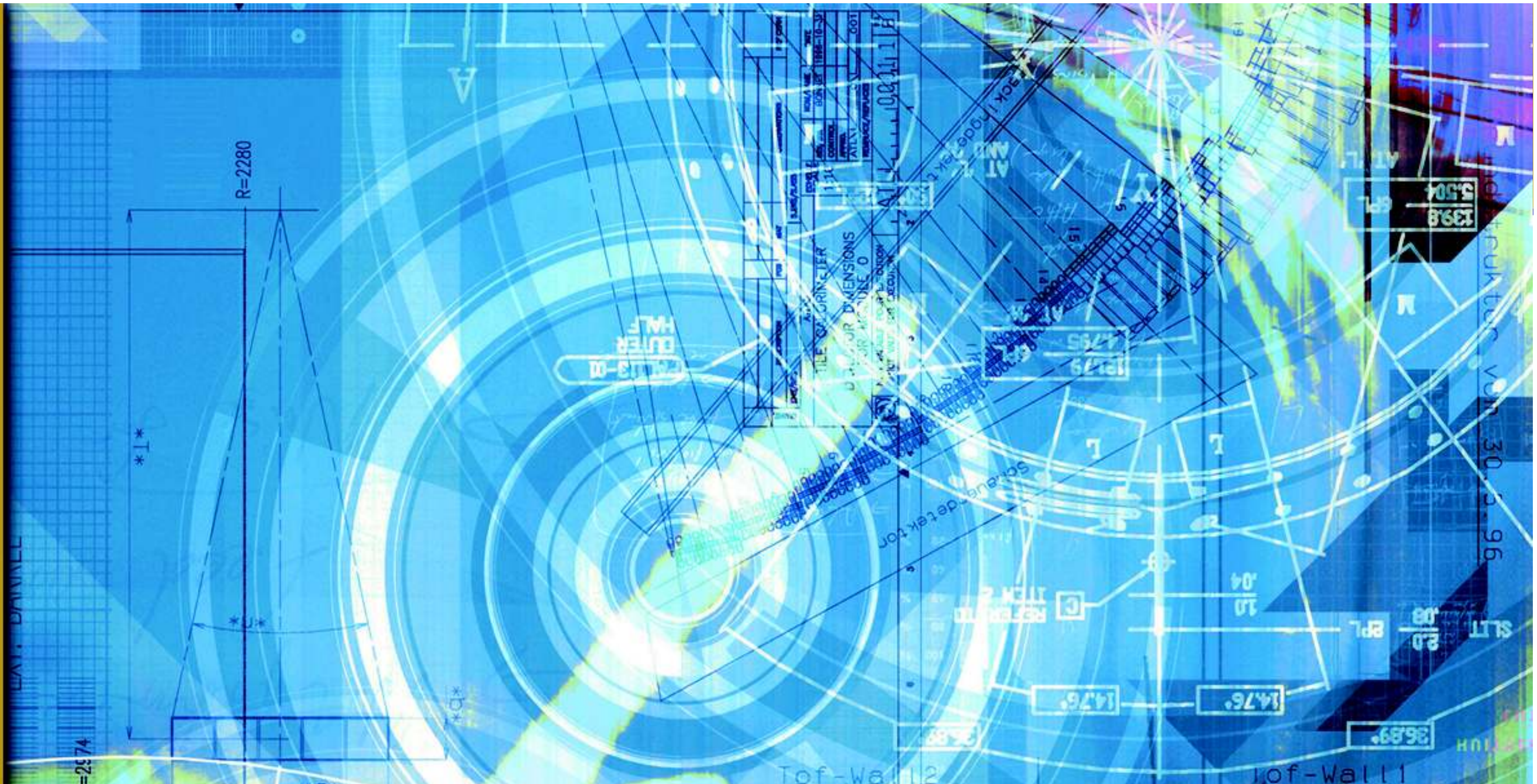


How to Develop Parametric Cost Estimating Relationships (CER)

presented by:

Jason A. Dechoretz

Senior Vice President, MCR LLC

Chairman, ISPA Board of Directors

Critical Thinking. Solutions Delivered.

MCR Critical Thinking. Solutions Delivered.

Topics

- Presentation Objectives
- Recall Why a Parametric CER is Helpful
- Overall Process
  - Requirements & Define WBS
  - Determine Cost Drivers and Heuristic CERs
  - Collect and Normalize Data
  - Develop and Validate Statistical CERs
- Candidate DACE CERs
- Example CER (Handout)
- Summary
- About MCR
- Acronyms

MCR Critical Thinking. Solutions Delivered.

Presentation Objectives

- Outline the general steps to generate a parametric CER that useful and valid
  - Identify key issues
  - Highlight lesson’s learned
- Identify important topics which take more presentation time
- Help DACE members learn the utility of parametric CERs
- What it is not:
  - Hands on CER development instruction
  - Tutorial on regression analysis (basic knowledge assumed)

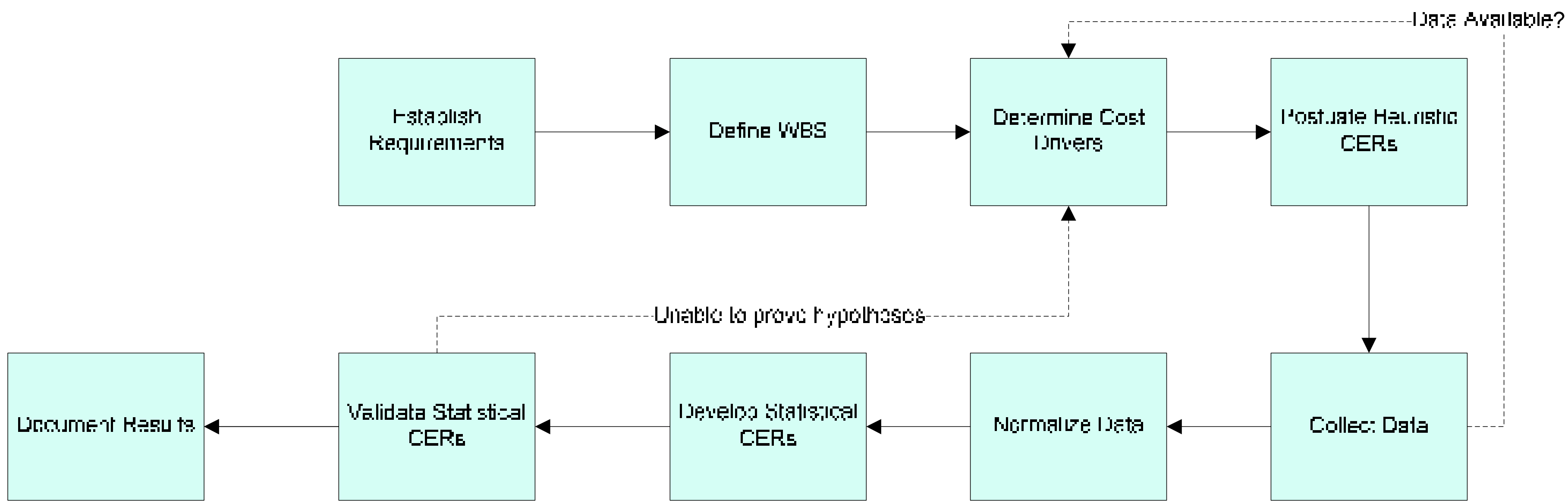
MCR CRITICAL THINKING. SOLUTIONS DELIVERED.

Recall Why a Parametric CER is Helpful

- The right environment
  - Doing something (technology, development process, operating environment, etc.) which is related but different than in the past
  - Historical data available from past, related efforts
  - Limited time to create initial estimate
  - Potential of multiple estimates over time: updates or additional efforts
- Desire to show a causal and tested relationship between estimate of future requirements to known historical experience
  - Not affected by Subject Matter Expert (SME) bias or experience base
  - Statistical tests available from widely available software
- It is fun!

CRITICAL THINKING. SOLUTIONS DELIVERED.

Overall Process



Requirements & Define WBS

- What needs to be estimated?
  - Normally it is a capital investment which has known design traits and significant variable costs
  - Scope of the estimates
    - Entire program (multiple CERs)
    - Single part family (single CER)
    - Product only or all below the line (BTL) costs: Program Management (PM), System Engineering (SE), Installation & Checkout (I&C), etc.
    - Range of operating environments (material, locations)
- Create a product based Work Breakdown Structure (WBS)
  - Level of detail set to support decision making, track high cost items, and support reporting requirements (e.g., bids)
  - Functional enablers are either blended into product WBS elements or maintained as BTL cost categories

Requirements & Define WBS

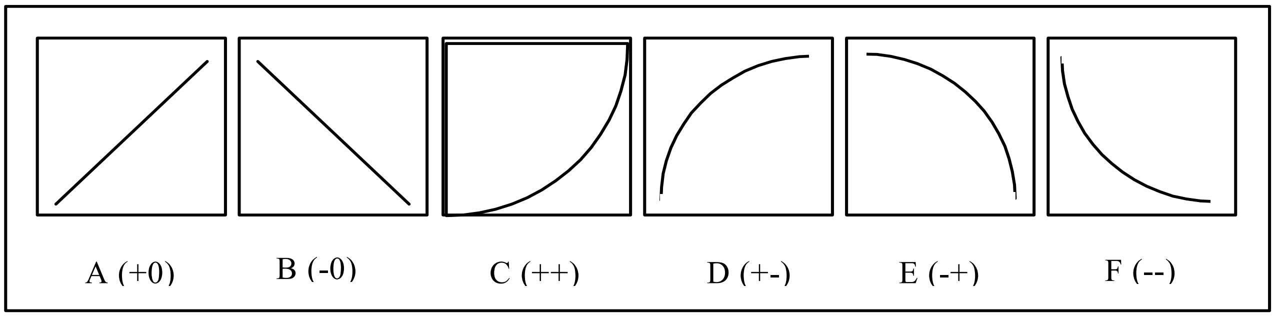
- Lesson's Learned & Helpful Hints
  - Requirements not just Today's but also near future (some CERs are relevant for years)
  - One CER (or model) does not fit all (even if the title looks like it)
    - Program Management:
      - Cost at only the program level
      - All people with the job code)
    - Plumbing:
      - Piping only
      - Super structure and/or pumps
    - Sensors:
      - Active: IR, RFID
      - Passive: flow meter
  - BLT enablers accommodate different and/or changing business models
    - Size and allocation of PM and SE resources

Determine Cost Drivers & Heuristic CERs

- Cost Drivers: parameters which have a direct impact on cost
  - Usually categorized by type
    - Scope: how big (surface or weight), how many (unique or total)
    - Complexity: precision (movement or interface), number of processes
    - Performance: Mean Time Before Failure (MTBF), accuracy of reading
    - Other: Cost to Cost, Time to Cost (using base year for technology)
  - Identified through experience and expected availability of data
  - Used in regression analysis to define impact trend with cost
  - Sometimes a proxy variable is used for a collection of known (but less important or harder to define) variable

Determine Cost Drivers & Heuristic CERs

- Heuristic CERs
  - Hypothesize the trend (direction and rate of change)!

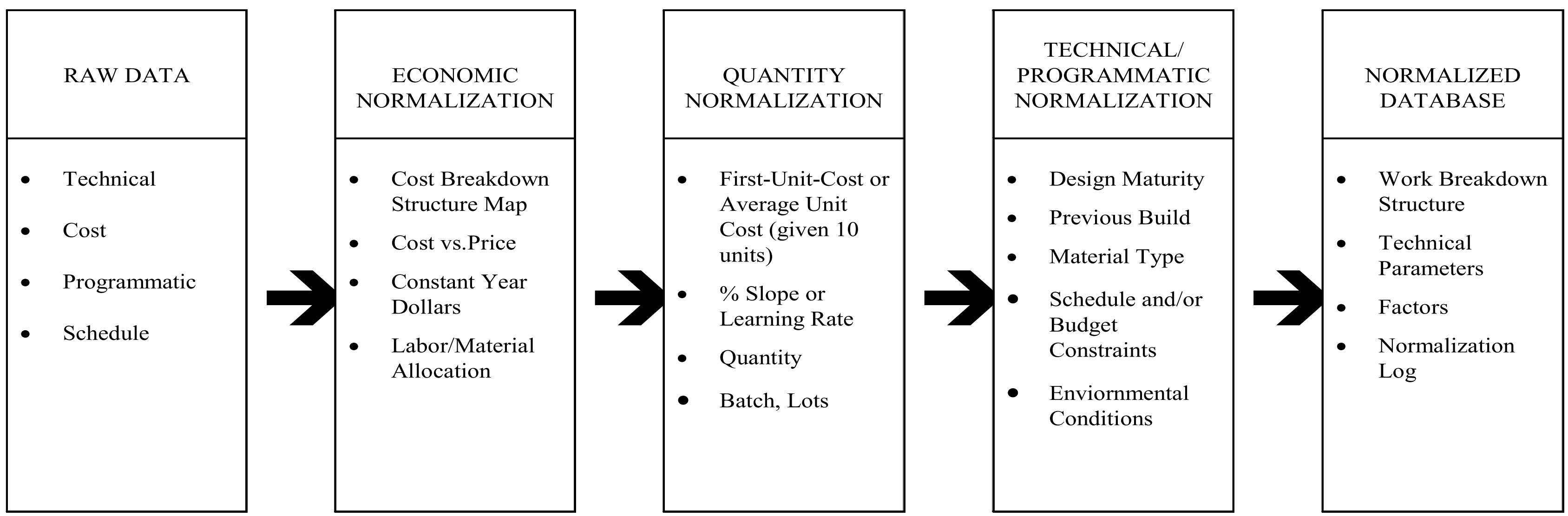


  - Identify which cost drivers are complementary (for multivariate equations) and which are redundant (avoid collinearity)
- Lesson’s Learned and Helpful Hints
  - Select cost drivers found in historical, factual documents not based on judgment
  - Don’t select cost drivers which are not available in the requirements or early design phase!

CRITICAL THINKING. SOLUTIONS DELIVERED.

Collect and Normalize Data

- Data Collection
  - Identify relevant programs and their data sources; each becomes a data point
  - Types of historical data: cost, programmatic/biographical, technical, schedule
- Data Normalization



Collect and Normalize Data

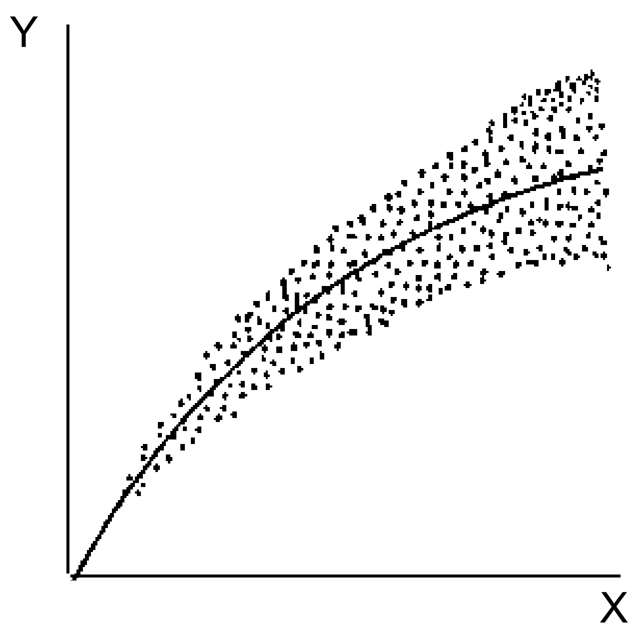
- Lesson's Learned and Helpful Hints
  - Collect copy of source documents and translate into data collection guide
  - Establish Ground Rules and Assumptions before normalization
    - Log exceptions by data point
    - Review normalized data with data source
  - Keep every data point at all costs!
    - Remove only if normalization requires judgment than remaining fact (e.g., extensive breakout across CBS elements)
    - Remember you will test for outliers
  - Mark and protect proprietary data
    - Note when a CER is developed only with proprietary data

CRITICAL THINKING. SOLUTIONS DELIVERED.

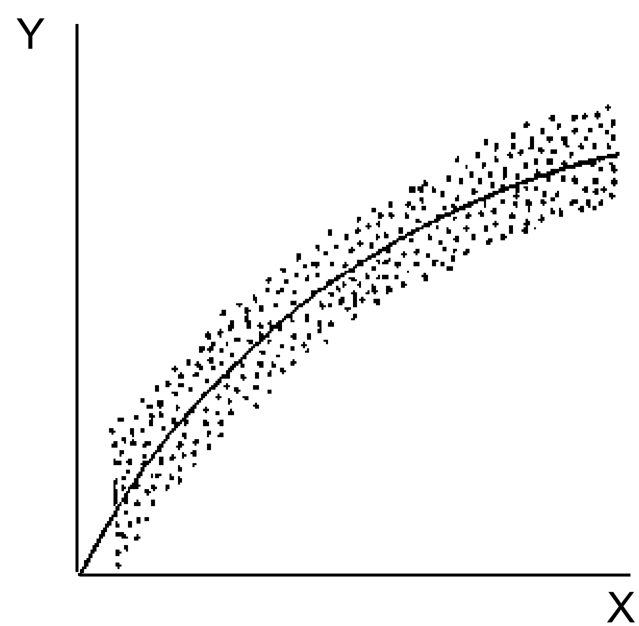
Develop and Validate Statistical CERs

- Select the functional equation form
  - Linear vs non-linear
  - Additive vs relative error term

Multiplicative Error



Additive Error



Reference: H.L. Eskew and K.S. Lawler, "Correct and Incorrect Error Specifications in Statistical Cost Models," *Journal of Cost Analysis*, Spring 1994, page 107.

MCR CRITICAL THINKING. SOLUTIONS DELIVERED.

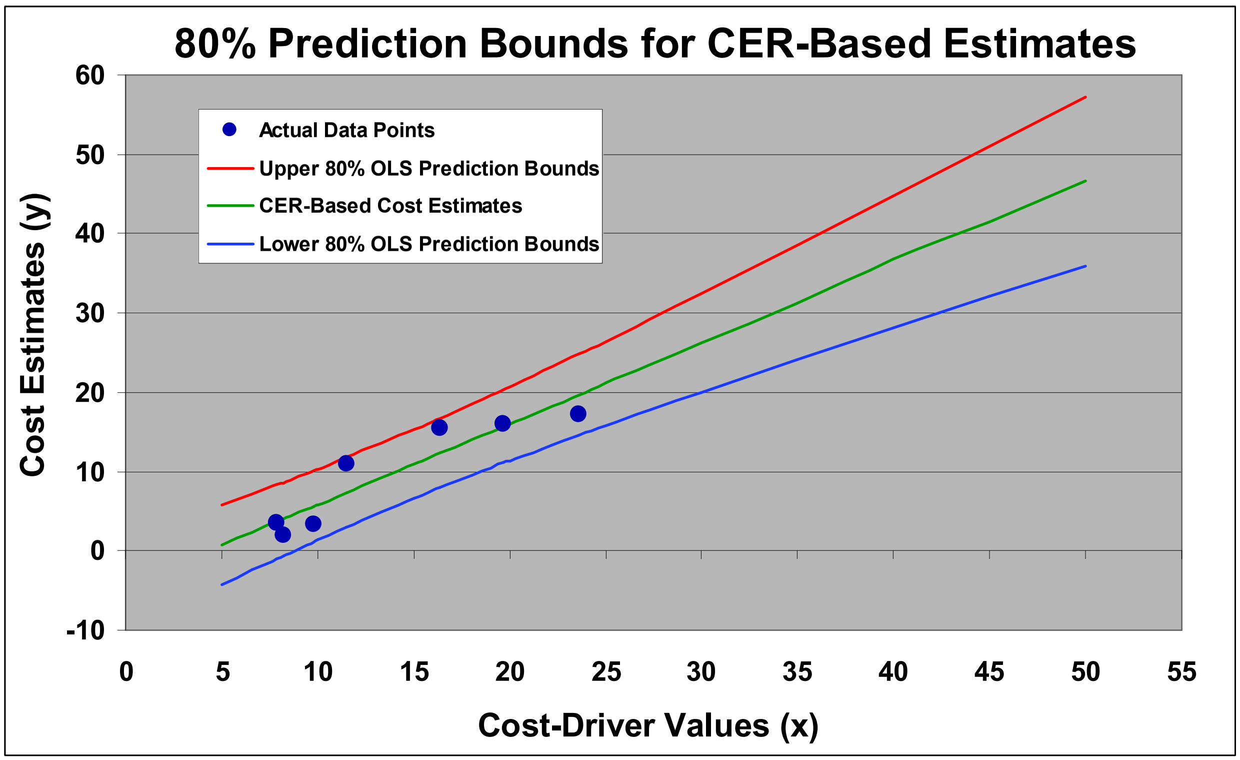
Develop and Validate Statistical CERs

- Confirm Heuristic CERs with scatter plots
- CER Selection Criteria

MEASURE	CRITERIA	EXPLANATION
COEFFICIENT OF VARIANCE	CV < 20%	CER is a reliable predictor of costs (variance sufficiently minimized, relative to the sample mean)
ADJ R^2	R2 > 0.80	Good correlation between cost and cost drivers
F-RATIO	F-RATIO > F-TABLE @ 90% Confidence	Regression equation is a better predictor of cost than the sample mean (average cost)
T-STAT	T-STAT > T-TABLE @ 90% Confidence	Correlation between cost and the independent variable likely does not occur by chance
DATA SET ROBUSTNESS	DF/N >0.6	Not every point is automatically influential
OUTLIERS	No statistical outliers	No obvious potential problem with data homogeneity
PARSIMONY	Minimize parameters	Over-specification could lead to spurious curve fitting
LOGIC	Make engineering sense	Valid estimator of cost because of causality

Develop and Validate Statistical CERs

- Validation
  - Ideally withhold 20% of data from regression analysis (minimum of two randomly selected data points)
  - Estimate using parametric values and compare historical costs to prediction intervals



MCR Critical Thinking. Solutions Delivered.

Develop and Validate Statistical CERs

- Lesson's Learned and Helpful Hints
  - Select the right equation form
    - Cost driver may indicate non-linear but need enough data to find the global minimum!
    - Data that spans an order of magnitude in cost should be modeled with relative error term
  - Tests: know the math but don't be a slave to the math
    - Statistical significance gives confidence the estimated coefficient is consistent and relevant to the sample
    - Sometimes engineering knowledge or programmatic knowledge need to over-rule the statistics for selections

Candidate DACE CERs

- Traditional CERs
  - Facility development
    - Piping design cost = $A * \text{Pipe Length}^{Z} + B * (\text{Fittings} + \text{Valves})$
  - Facility Modernization
    - Construction labor cost = $A * M^2_{(new)} + B*[M^2_{(modified)} / M^2_{(existing)}]$
- Cost to Cost and Year to Cost Factors
  - BTL costs for construction and operations: benchmarks
    - PM costs = A * Construction labor cost
  - Processing volume or yields relative to a base year
    - Delivery costs = $A * [\text{Max Volume}_{(Y)} - \text{Max Volume}_{2001)}]^{Z}$

* Regression analysis estimates values for A, B, Z

Summary

- Parametric CER development
  - Requires an upfront investment of data collection and analysis
  - Results in lower time and effort to create initial and subsequent new or modified estimates
  - Credibility derived from relationship to multiple historical programs and statistical tests
  - Allows for the sharing of equations without exposure of proprietary data
- Advances in mathematics and statistics continue to improve data utilization estimating accuracy
- Wanting more? It is normal!
  - Typical CER development class takes 1 – 3 days

Critical Thinking. Solutions Delivered.

About MCR

MCR, LLC continues its more than 30-year tradition of providing high quality, value-added support to government and industry customers in life cycle Integrated Program Management (IPM). MCR's approach utilizes our four core competencies defining program requirements through **Strategic Planning**; establishing program baselines by conducting **Cost and Schedule Analysis**; ensuring the Program Manager properly utilizes their resources and schedule with day to day **Acquisition Management**, and achieving program performance requirements by addressing risk and contractual execution issues through ongoing **Program Assessment**. With over 600 professional staff, MCR is considered the "Thought Leader" in IPM providing the expertise, knowledge, and tools to the Federal marketplace. MCR has supported literally hundreds of organizations and acquisition programs in numerous domains, including defense, energy, environmental, information management, construction, transportation, and space. Our people have been successful in providing the right solution for their needs and ultimately the country and its citizens. In other words, our *Critical Thinking* leads to the best *Solutions Delivered*...

MCR Critical Thinking. Solutions Delivered.

Acryonyms

BTL – Below-The-Line

CBS – Cost Breakdown Structure

CER – Cost Estimating Relationship

I&C – Installation & Checkout

IPM – Integrated Program Management

IR – Infra-Red

MTBF – Mean-Time Before Failure

PM – Program Management

RFID – Radio Frequency Identification

SE – System Engineering

SME – Subject Matter Experts

WBS – Work Breakdown Structure

1.2.2.1.1 GROUP A KIT PRODUCTION
Multiple System Modifications
Primary Detail Level Equation

GRP_AT10 = 25.2 + 16.2 * SYS_ADD + 21.0 * LRU_MOD

Where: GRP_AT10 = Theoretical average unit cost of the Group A Kit based on the production of ten units, in CY97K$.
and: SYS_ADD = Total number of systems added to the host platform during the modification.
LRU_MOD = Total number of Line Replaceable Units (LRUs), already located on platform, which are modified during the effort.

CORRELATION MATRIX:

	GRP_AT10	SYS_ADD	LRU_MOD
GRP_AT10	1.00	0.60	0.81
SYS_ADD	0.60	1.00	0.05
LRU_MOD	0.81	0.05	1.00

CER FIT STATISTICS:

# of OBS	MEAN	SEE	CV	ADJ R^2	R^2
10	277.0	33.4	12.1%	0.95	0.96

INDEPENDENT VARIABLE INFORMATION:

VARIABLE	T-STAT	STD ERR	SIG-T	BETA COEFF	RANGE	MEAN
CONSTANT	1.2	21.4	0.3	N/A	N/A	N/A
SYS_ADD	7.8	2.1	0.0	0.56	2.0 - 12	6.20
LRU_MOD	10.8	1.9	0.0	0.78	0.0 - 17	7.20

ANOVA TABLE:

SOURCE	Sum of Squares	DF	Mean Square Error	F Ratio/SIG (F)
Regression	207544.2	2	103772.1	92/0.0
Residual	7825.5	7	1117.9	
TOTAL	215369.7			

DATA POINTS:

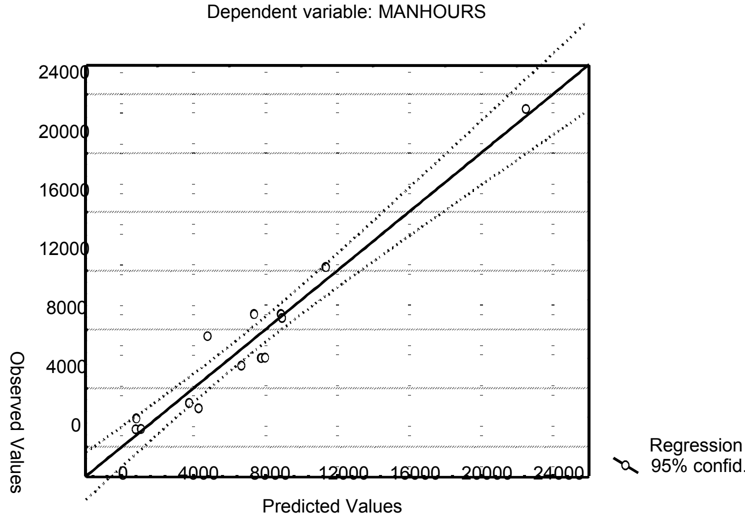
A-1
A-2
A-3
A-4
A-6
A-7
A-8
A-9
A-10
A-64

Critical Thinking. Solutions Delivered.

Sample CER

1.2.2.1.1 GROUP A KIT PRODUCTION
Multiple System Modifications

Primary Detail Level Equation



OUTLIER ANALYSIS:

- Record 5 was removed from the modeling dataset because due to evidence of cost allocations with another modification.
- Record 42 was removed from the modeling dataset because the Group A Kit includes a significant amount of equipment to modify the engines.
- Record 51 was removed from the modeling dataset due to incomplete cost data.

COMMENTS:

- None.

Sample CER (2)

CRITICAL THINKING. SOLUTIONS DELIVERED.