



Parametric Cost Estimating Training Track For The

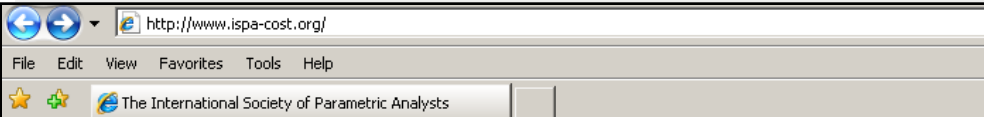
Parametric Estimating Handbook 4th Edition

Noordwijk The Netherlands 2008





www.ispa-cost.org



THE INTERNATIONAL SOCIETY
OF PARAMETRIC ANALYSTS

- ISPA Home
- Handbook
- Conferences
- Other Events (NEW)
- Chapters
 - Mid-Atlantic
 - Southern California
 - Southwest
- PCEI
- Membership
- Elections (NEW)
- Certification
- Publications
- Order Form
- Website Ads
- Website Job Postings
- Parametrics
- Acquisition
- Contact ISPA

Welcome to the ISPA

ISPA is a professional society dedicated to the improvement and promotion of:

- parametric cost modeling techniques and methodologies
- risk analysis
- econometrics
- design-to-cost
- technology forecasting
- cost management

ISPA provides a forum that encourages the professional development of its members and perspectives. ISPA members represent government agencies, universities, and countries. ISPA's membership range in experience from beginners to seasoned professionals with a strong interest in the practical application of parametric analysis.

ISPA Mission Statement

The purpose of ISPA is to improve our effectiveness and leverage the limited resources by providing information on techniques, tools, and methodologies for enhancing the process.

Updates to Website will be accomplished every Tuesday and up

Current Revisions to Website (NEW)

Chapter Workshop Announcements

- Latest information for Southwest Chapter's February 2009 Workshop. [February 2009](#)
- Southern California Announce their March 2009 Chapter Workshop. [Southern California](#)
- Southern California December 2008 Meeting Presentations Posted. [December 2008](#)

Meet us in St. Louis for the 2009 ISPA/SCEA Professional Development and Training Workshop 2-5 June 2009



The Place



St. Louis, Missouri - "Gateway to the West"

No matter what your interests may be, you'll be able to find just what you're looking for in St. Louis. Head to LaClede's Landing on the riverfront for live music and delicious food, or enjoy the French cuisine and jazz clubs of historic Soulard. Wander through the eclectic shops while admiring the stunning architecture of Union Station. Or how about a visit to the new Busch Stadium complex, or a tour of the Anheuser-Busch Brewery? From Victorian gardens to riverboat gaming, St. Louis has it all. Everything is just moments away from our hotel, the luxurious Renaissance Grand Hotel St. Louis, situated within walking distance of the Gateway Arch.

The Program

The Professional Development and Training Workshop is designed to appeal to all levels of experience. Cost community experts come together from government, industry, and academia to network and share ideas, present papers, attend panel discussions, and listen to the stimulating keynote addresses of leaders in the field. Training Workshops are offered in the Fundamentals, Practitioner, and Integration tracks, so whether you're attending to keep abreast of the latest techniques, or you're preparing for the CCE/A or CPP exam held Saturday, June 6, there will be plenty of opportunities!



For more information...



The 2009 ISPA/SCEA Professional Development and Training Workshop will be an event you won't want to miss! Look for future information about registration, exhibiting, and presenting at www.sceaonline.org (SCEA's website) or www.ispa-cost.org (ISPA's website), or call Erin at the SCEA & ISPA Joint Office at (703) 938-5090.



We look forward to seeing you there!

This event is sponsored by the Society of Cost Estimating and Analysis and the International Society of Parametric Analysts



You are the **15954** visitor!



Parametric Cost Estimating Training Track

Session 4

Cost Estimating Relationships

Dale Shermon

Handbook Chapter 3



Developing a CER – Definition & Context

- Definition: A Cost Estimating Relationship (CER) is a mathematical expression of cost as a function of one or more independent variables.
- Cost-to-cost relationship: One type of cost related to another cost.
- Cost-to-non-cost relationship: Cost is expressed as a function of an attribute or characteristic.
- Model: Integration of several CER's.
- Modeling System: Integration of several CER's where driving variables effect the relationships or slopes of the CER's and change the results of several of the CER's



Developing a CER – Overall Process

1. Identify opportunity
2. Collect data
3. Evaluate and normalize the data
4. Select CER variables
5. Test relationships
6. Perform regression and curve fitting
7. Conduct data analysis and correlation
8. Select CER's
9. Validation & approval of CER's
10. Institutionalize (establish CER in methods base)
11. Maintain CER (go back to step 2)



Developing a CER – Logic Testing

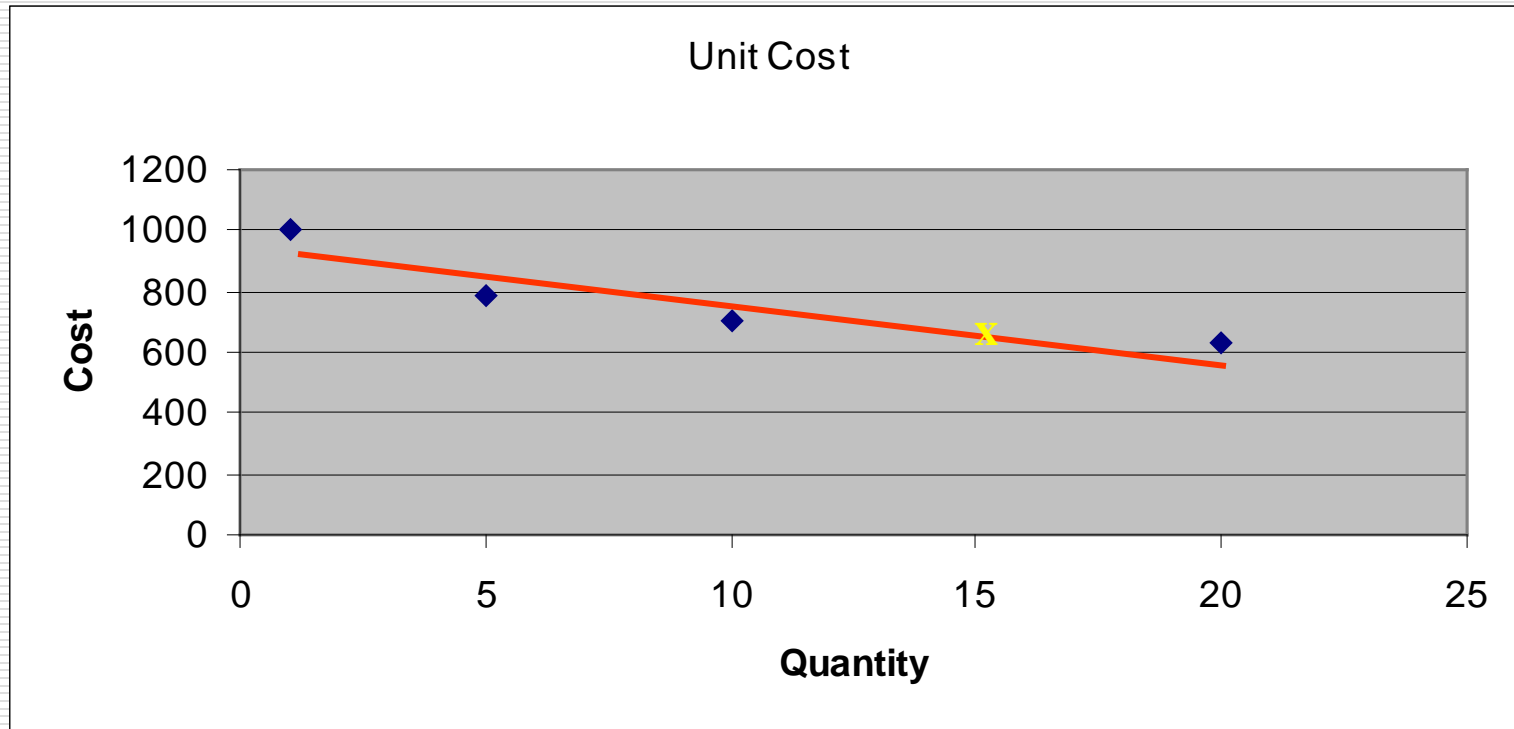
- Logic Testing
 - CER must be logical
 - Which comes first:
 - Hypothesis?
 - Database?
 - Subject Matter Experts (SME) are invaluable
 - Logic judgment is often statistically or results based

- The CER Model
 - Pronounced model: Based on the hypothesis of an SME; e.g. \$1M for every foot of documentation.
 - Regressed Model: Based upon statistical methods to fit an equation to data.



Curve Fitting- Graphical Method

- “Eyeball” fit of data – fitting by approximation

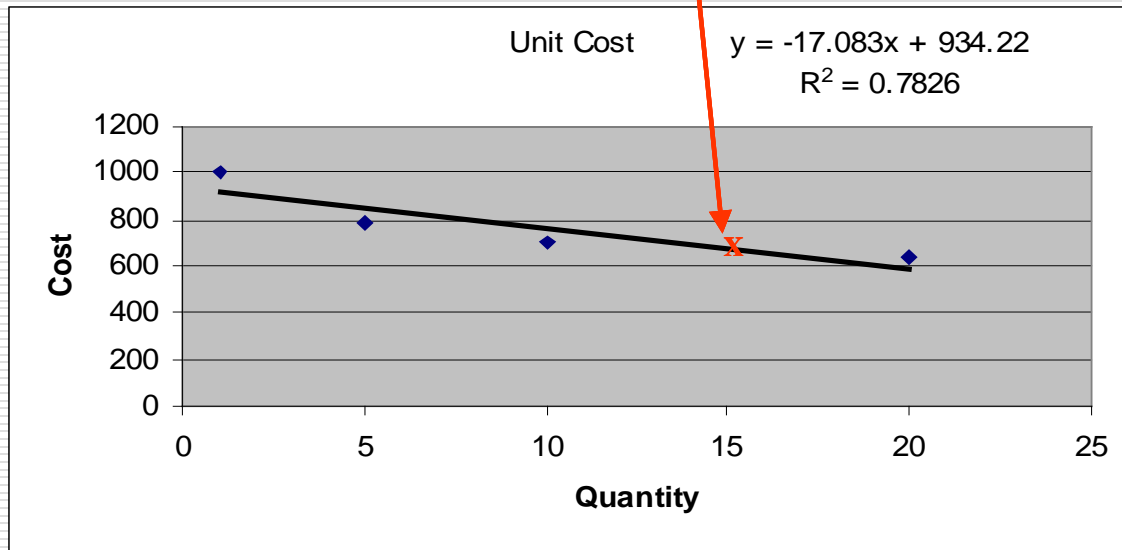




Curve Fitting - Least-Squares Best Fit (LSBF)

- The root of regression analysis
- Fits independent variables to a dependent variable or variables
- Resulting mathematical form may be linear or curvilinear
- Fit minimizes the square of errors between the specific equation and the data (least errors squared yields best fit)

$$-17.083 \times 15 + 934.22 = 678$$





Curve Fitting - Least-Squares Best Fit (LSBF)

Computational Example

	X	Y	X * Y	X ^ 2	Y ^ 2
	1	1000	1000	1	1000000
	5	783	3915	25	613068
	10	705	7047	100	496585
	20	634	12684	400	402234
Sum	36	3122	24646	526	2511888

$$Y = a + bX$$

$$b = \frac{\sum XY - n\bar{X}\bar{Y}}{\sum X^2 - n\bar{X}^2} = \frac{24645 - 4 * (36 / 4) * (3122 / 4)}{526 - 4 * (36 / 4)^2}$$

$$b = -17.083$$

$$a = \bar{Y} - b\bar{X} = (3122 / 4) - (-17.083) * (36 / 4)$$

$$a = 934.22$$



Curve Fitting - Least-Squares Best Fit (LSBF)

ASSUMPTIONS

- Mean value of distribution lies on the LSBF line; e.g. (9, 780) in example
- Dependent variable values normally distributed about the LSBF line
- Independent variables are error-free; i.e. not random
- Error term (e) is random; $Y = (A + b * X) + e$
- Extrapolation is theoretically invalid but practically accepted if
circumstances pass similarity tests.



Curve Fitting – Multiple and Curvilinear Regression

□ Multiple Regression

- 2 or more independent variables
- Step-wise regression most often used identify appropriate variables.
- Statistical “F-test” usually employed in decision to add variables.
- Multi-co linearity Problem
 - Independent variables may not really be so
 - Regression analysis doesn’t identify this
 - Must rely on pair-wise correlation to solve

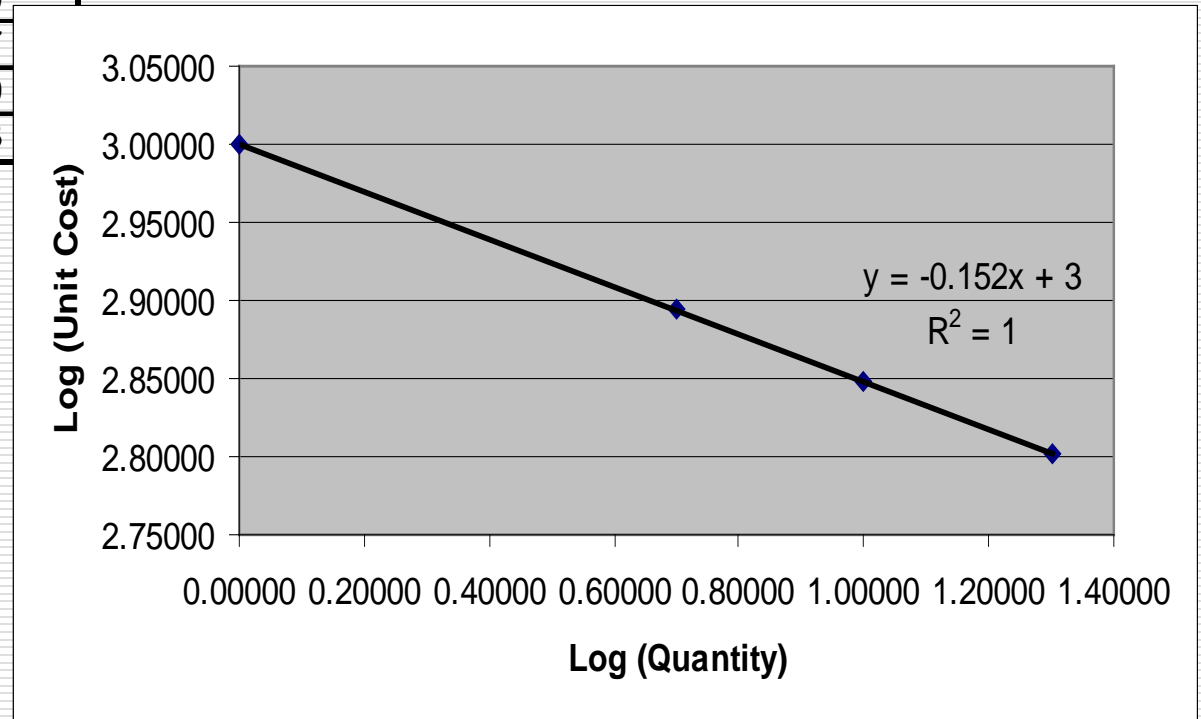
□ Curvilinear Regression

- Transformations used for exponential, logarithmic, and general power series functions
- e.g. $Y = A * X^b$ can be transformed to
 - $Y = A + b * X$;
 - $Y = \text{Log}(Y); A = \text{Log}(A); \&$
 - $X = \text{Log}(X)$



Curve Fitting – Curvilinear Regression

Log (Unit Cost)	Log (Quantity)
3.00000	0.00000
2.89375	0.69897
2.84800	1.00000
2.80224	1.30103



Regressed Equation: $\text{Log (Y)} = 3 - .152 * \text{Log (X)}$

or $Y = 1000 * X^{-.152}$

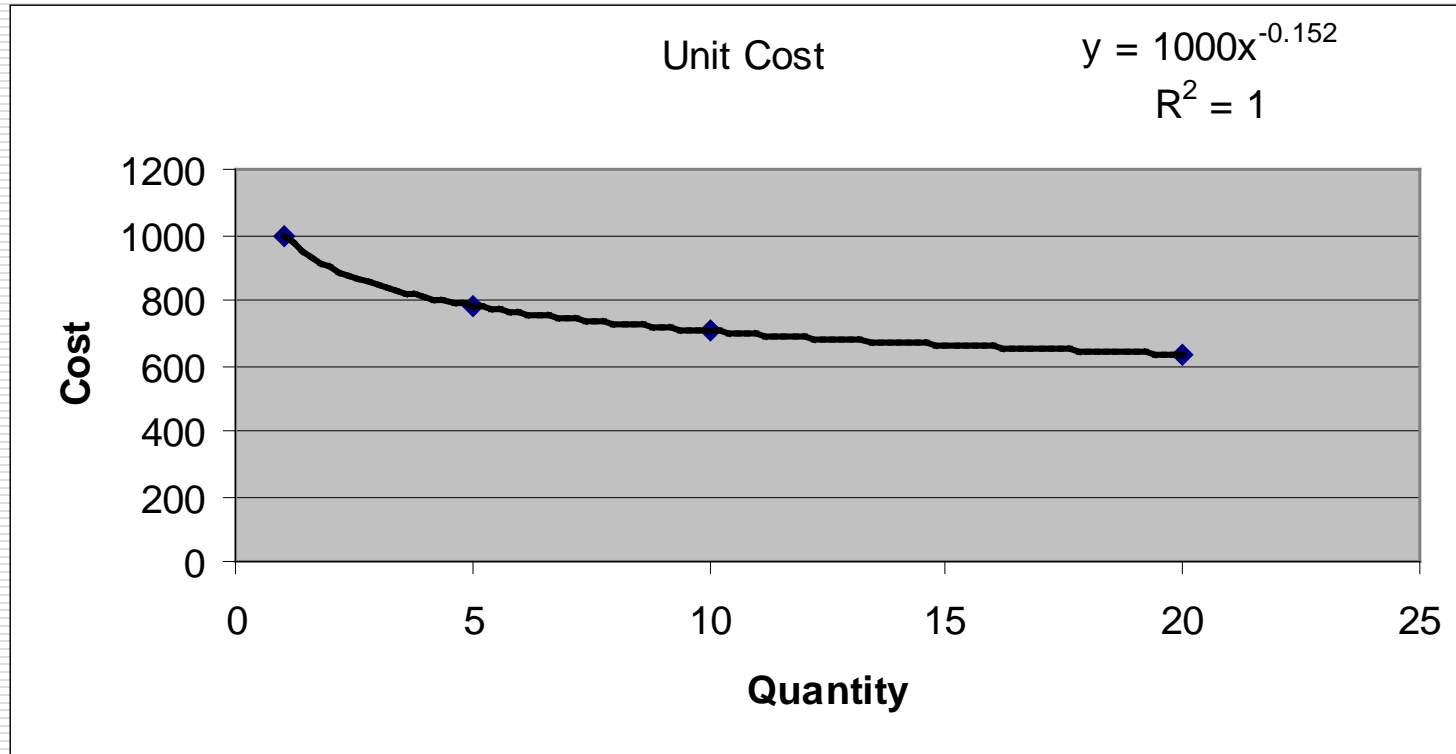


Curve Fitting – Cost Improvement Curve

- Also known as the Learning Curve
- Mathematical form: $Y = A * X^b$
 - Y is either unit cost or average unit cost
 - A is first unit cost (T-1)
 - X is unit number
 - b is improvement or learning factor
- Learning Curve Interpretation
 - Constant cost improvement with each doubling of quantity (e.g. 90% LC = 10% improvement with every doubling of quantity)
 - $b = \ln(\text{LC}) \div \ln(2)$



Curve Fitting – Cost Improvement Curve



$$-.152 = \ln(\text{LC}) \div \ln(2); \ln(\text{LC}) = -.152 * .692 = -.105$$

$$\therefore \text{LC} = e^{-.105} = .90 \text{ or a } 90\% \text{ Learning Curve}$$



Curve Fitting – Cost Improvement Curve

- Unit Learning Curve: Equation describes unit cost behavior
- Cumulative Average Learning Curve: Equation describes average cost behavior
- For a fixed data set, the difference is the T-1
- If A_U is the Unit LC T1 and A_C is the Cumulative Average LC T1, for unit N,

$$A_C \times N^b = (A_U \times \sum_{i=1}^N X_i^b) \div N$$



CER Significance Testing

- How good is the CER?
- How likely is it to estimate well?
- What is the confidence level?
- Is the database sufficiently large?
- Are there few independent variables?
- Is the range of applicability wide?
- Is the prediction error small?
- Quantitative & qualitative measures apply



CER Significance Testing – Validation Measures and Prediction Measures

- Validation Measures
 - R^2 : Y variation explained by X (.90 or higher is desired)
 - SE: Avg. estimating error
 - CV: $SE \div \text{Avg. } Y$ (.1 or less is good)
 - F-stat: Equation validity test (close to 0)
 - t-stat: X variable validity test (close to 0)
 - Data: Narrative description of amount and quality
 - Logical Relationships: Narrative description of rationale
- Prediction Measures
 - N: Number of observations (many)
 - d.f.: $N - \# \text{ Independent variables} + 1$ (close to N)
 - Outliers: Poorly predicted Y values (none is ideal)
 - Data Range: Large enough to cover expected future observations (complete coverage is not essential)



Appropriate Use of CER's – Strengths and Weaknesses

■ Strengths

- Excellent predictor if used appropriately
- Time Saving Method
 - Proposal cycle time
 - Evaluation & negotiation cycle time
- Easy to apply
- Functions with minimal information

■ Weaknesses

- Sometimes too simplistic
- Strong database dependencies
 - Developer should validate CER and database
 - Applier must validate use; i.e., database applicability, database age, data normalization
 - CER source documentation must be complete



Appropriate Use of CER's - General

- CER's targeted to specific forecasts invoke more confidence
- Generic CER's have wider applicability, but care must be exercised to:
 - Adapt to different assumptions through incorporation of new data and/or variables
 - Document the adaptation, no matter how slight
- Application domain is client driven
- Confidence in generalized CER's is gained by thorough testing and the wide applicability of the development database and method of construction



CER Examples – Electronics and Software

- **Electronics**
 - Variables & Constants
 - X: Number of ICs
 - Y: Electronic Item Unit Production Cost
 - A: Set-up cost (\$57)
 - b: Recurring Unit Cost per IC (\$1.10)
 - CER: $Y = A + b \times X$
 - Example: for a 30 IC unit,
 $Y = \$57 + \$1.10 \times 30 = \$90$



CER Examples – Construction/Reconstruction

- Variables & Constants
 - X_1 : Area (square feet)
 - a: area coefficient (\$75 per square foot)
 - X_2 : # Rooms
 - b: rooms coefficient (\$5,000 per room)
 - C: Home type (1 for tract, 1.33 for semi-custom, 1.5 for custom)
 - Y: Construction Cost
- CER: $Y = C \times (a \times X_1 + b \times X_2)$
- Example: A 12 room, 2500 square foot, semi-custom built home,
 $1.5 \times (\$75 \times 2500 + \$5000 \times 12) = \$371,250$.



CER Examples - Software

- Variables & Constants
 - X: Thousands of Delivered Source Instructions
 - A: Type coefficient; 3.2, 3.0, or 2.8
 - b: Type Exponent; 1.05, 1.12, or 1.2
 - Y: Labor months of development effort
- CER: $Y = A \times X^b$
- Example: For a 40,000 DSI, Type 3 program
 $2.8 \times 40^{1.2} = 234.2$ labor months



CER Examples

CER Title	Pool Description	Base Description	Application
Panstock Material	Allocated Panstock dollars charged.	Manufacturing Assembly “touch” direct labor hours charged.	Panstock is piece-part materials consumed in the assembly organization. The panstock CER is applied to 100% of direct labor hours for assembly effort.
F/A-18 Software Design Support	Allocated effort required to perform software tool development and support for computer & SW engineering.	Computer and software engineering direct labor hours charged.	F/A-18 computer and software engineering support direct labor hours estimated for tool development.
Design Hours	Design engineering including analysis and drafting direct labor hours charged.	Number of design drawings associated with the pool direct labor hours.	The design hours per drawing CER is applied to the engineering tree (an estimate of the drawings required for the proposed work).



CER Examples

CER Title	Pool Description	Base Description	Application
Systems Engineering	Systems engineering (including requirements analysis and specification development), direct labor hours charged.	Design engineering direct labor hours charged.	The systems engineering CER is applied to the estimated design engineering direct labor hours.
Tooling Material	Nonrecurring, in-house, tooling raw material dollar costs charged.	Tooling nonrecurring direct labor hours charged.	The tooling material CER is applied to the estimated nonrecurring tooling direct labor hours.
Test/Equipment Material (dollars for avionics)	Material dollars (<\$10k).	Total avionics engineering procurement support group direct labor hours charged.	The test/equipment material dollars CER is applied to the estimated avionics engineering procurement support group direct labor hours.



Evaluating CERs - Criteria

■ Criteria

- Customer compliance based
- Logical data relationships
- Adequate data used
- Evidence of strong relationships
- Consistent and valid application
- Cost drivers (variables) must make sense
- Direction of drive must make sense
- Driving data should be accessible for both CER development and implementation
- Driver list should be comprehensive
- Excluded outliers and rationale for exclusion should be identified.



Evaluating CERs – Criteria *Continued*

■ **Validation**

- Establishes the CER as a credible estimating method
- Validation = Documented Consistent Accuracy
 - Requires periodic testing
 - CER updated when testing indicates need
- Validation falls to CER developer
- Certification falls to CER applier



Evaluating CERs – Credible Data

- Credible Data
 - Assembling a quality database is the single most important and time consuming step
 - Database must be encompassing but not overwhelming
 - Historical data is preferred
 - Data normalization must be consistent
 - Formats for data types should be consistent (Cost, Technical, Programmatic)
 - Database must be fed and cared for
- Strength of Data Relationships
 - Can be judged with Statistical tests
 - Standard Error is most popular strength test
 - Apply development data to CER
 - Measure error (|predicted - actual|)
 - $\pm 10\%$ or better is desirable



CERs - Summary

A CER can be a single or multi-variable cost relationship.

- A model is a system of integrated CER's and possibly other relationships.
 - Building a quality database is the most intensive step in developing a valid CER.
 - LSBF is the most popular data regression method for CER development; it can be used for linear and non-linear relationships.
 - LSBF finds the plot with the least error to the regressed data; the mean value always lies on the LSBF plot.
 - The Cost Improvement or Learning Curve describes a relationship where either unit or average unit cost declines by a fixed amount for each doubling of quantity.
 - Statistics are frequently used to assess the significance of a CER developed by LSBF.
-
- Outliers and Extrapolation are discouraged by theory, but acceptable in practice.



- Questions?
- Discussion