CHARACTERIZING THE ACCURACY OF DoD OPERATING AND SUPPORT COST ESTIMATES

Erin Ryan, Major, PhD
Air Force Institute of Technology
Life Cycle Cost

O&S Costs comprise 60-75% of Life Cycle Costs
# DoD Cost Growth Studies (1972-08)

**The AFIT of Today is the Air Force of Tomorrow**

## ACQUISITION COST GROWTH

<table>
<thead>
<tr>
<th>Year</th>
<th>Study Title</th>
<th>Author/Institution</th>
</tr>
</thead>
<tbody>
<tr>
<td>1972</td>
<td>Choice Among Strategies for System Acquisition</td>
<td>RAND</td>
</tr>
<tr>
<td>1972</td>
<td>Should Cost/Will Cost/Must Cost--A Theory on the Cause of Cost Growth</td>
<td>U.S. Army SSO</td>
</tr>
<tr>
<td>1973</td>
<td>Cost Growth in Major Weapon Systems</td>
<td>GAO</td>
</tr>
<tr>
<td>1973</td>
<td>Cost Overruns in Defense Procurement: A Problem of...</td>
<td>Northwestern University</td>
</tr>
<tr>
<td>1974</td>
<td>The Study of Cost Growth of a Major Weapon System</td>
<td>NPGS</td>
</tr>
<tr>
<td>1974</td>
<td>Bias in Initial Coat Estimates: How Low Estimates Can Increase the Cost...</td>
<td>RAND</td>
</tr>
<tr>
<td>1974</td>
<td>A Cost Growth Model for Weapon System Development Programs</td>
<td>AFIT</td>
</tr>
<tr>
<td>1974</td>
<td>A Model to Predict Final Cost Growth in a Weapon System Development Program</td>
<td>AFIT</td>
</tr>
<tr>
<td>1976</td>
<td>Study of Weapon System Cost Growth</td>
<td>OSD</td>
</tr>
<tr>
<td>1976</td>
<td>Statistical Analysis of the Effectiveness of Program Initial Conditions as Predictors...</td>
<td>NPGS</td>
</tr>
<tr>
<td>1977</td>
<td>A General Technique for R&amp;D Cost Forecasting</td>
<td>USAF Academy</td>
</tr>
<tr>
<td>1977</td>
<td>Study of Factors Leading to Changes in Cost Estimates...</td>
<td>George Washington University</td>
</tr>
<tr>
<td>1978</td>
<td>Methodology for Developing Total Risk Assessing Cost Estimates</td>
<td>U.S. Army MRDC</td>
</tr>
<tr>
<td>1978</td>
<td>A Range of Cost Measuring Risk and Uncertainty in Major Programs</td>
<td>GAO</td>
</tr>
<tr>
<td>1978</td>
<td>Financial Status of Major Federal Acquisitions</td>
<td>GAO</td>
</tr>
<tr>
<td>1979</td>
<td>Inaccuracy of DoD Weapons Acquisition Cost Estimates</td>
<td>Committee on Govt Operations</td>
</tr>
<tr>
<td>1979</td>
<td>An Overview of Acquisition Policy Effectiveness in the 1970s</td>
<td>RAND</td>
</tr>
<tr>
<td>1986</td>
<td>Improving the Military Acquisition Process, Lessons from RAND Research</td>
<td>RAND</td>
</tr>
<tr>
<td>1988</td>
<td>Weapons Cost: Analysis of Major Weapon Systems Cost and Quantity Changes</td>
<td>GAO</td>
</tr>
<tr>
<td>1989</td>
<td>Acquiring Major Systems: Cost and Schedule Trends and Acquisition...</td>
<td>IDA</td>
</tr>
<tr>
<td>1993</td>
<td>Analysis of Weapon System Cost Growth; Pitfalls in Calculating Cost Growth...</td>
<td>RAND</td>
</tr>
<tr>
<td>1996</td>
<td>The Defense System Cost Performance Database: Cost Growth Analysis Using SARs</td>
<td>RAND</td>
</tr>
<tr>
<td>1999</td>
<td>The Impact of the Packard Commission's Recommendations on Reducing...</td>
<td>Air Force</td>
</tr>
<tr>
<td>2000</td>
<td>Acquisition Trend Metrics in the Department of Defense</td>
<td>DAU</td>
</tr>
<tr>
<td>2003</td>
<td>Estimating Procurement Cost Growth Using Logistic and Multiple Regression</td>
<td>AFIT</td>
</tr>
<tr>
<td>2003</td>
<td>The Relationship Between Cost Growth and Schedule Growth</td>
<td>DAU</td>
</tr>
<tr>
<td>2004</td>
<td>Surveying Cost Growth</td>
<td>OSD/AT&amp;L</td>
</tr>
<tr>
<td>2004</td>
<td>An Analysis of Aircraft Weapon Systems Cost Growth and Implementation...</td>
<td>AFIT</td>
</tr>
<tr>
<td>2006</td>
<td>Historical Cost Growth of Completed Weapon System Programs</td>
<td>RAND</td>
</tr>
<tr>
<td>2007</td>
<td>Is Weapon System Cost Growth Increasing</td>
<td>RAND</td>
</tr>
<tr>
<td>2008</td>
<td>Sources of Weapon System Growth</td>
<td>RAND</td>
</tr>
</tbody>
</table>

## O&S COST GROWTH

None
Agenda

The AFIT of Today is the Air Force of Tomorrow

- Study Motivation
- Methodology
- Results
  - Annualized Unit O&S Cost (AUC)
  - Life Cycle Cost (LCC)
  - Total O&S Cost
- Implications
- Validity
- Now What?
Study Motivation

The AFIT of Today is the Air Force of Tomorrow

- Premise: DoD needs to be able to characterize accuracy of O&S/LCC estimates for its programs
  - Affordability!
- Greater emphasis from OSD on this topic—

“The strategic intent is to emphasize how O&S cost estimates will actively support key decisions throughout the system life cycle, rather than calling for O&S cost estimates simply for the sake of having an estimate.”

-- FY2010 Annual Report on Cost Assessment Activities, Director CAPE, Feb-2011
Recent O&S Cost Growth Studies

And then came WSARA…
- 4 O&S “cost growth” studies since 2009

<table>
<thead>
<tr>
<th>#</th>
<th>Source</th>
<th>Year</th>
<th># of Systems</th>
<th>Method</th>
<th>Quant. Results?</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>OSD</td>
<td>2009</td>
<td>34</td>
<td>Cost Growth in O&amp;S Actuals</td>
<td>n/a</td>
</tr>
<tr>
<td>2</td>
<td>CNA</td>
<td>2009</td>
<td>23</td>
<td>Cost Growth in O&amp;S Estimates</td>
<td>n/a</td>
</tr>
<tr>
<td></td>
<td>CNA</td>
<td>2009</td>
<td>3</td>
<td>O&amp;S Estimates vs. O&amp;S Actuals</td>
<td>No</td>
</tr>
<tr>
<td>3</td>
<td>IDA</td>
<td>2010</td>
<td>1</td>
<td>Cost Growth in O&amp;S Estimates</td>
<td>n/a</td>
</tr>
<tr>
<td>4</td>
<td>GAO</td>
<td>2010</td>
<td>5</td>
<td>Cost Growth in O&amp;S Estimates</td>
<td>n/a</td>
</tr>
<tr>
<td></td>
<td>GAO</td>
<td>2010</td>
<td>2</td>
<td>O&amp;S Estimates vs. O&amp;S Actuals</td>
<td>Yes</td>
</tr>
</tbody>
</table>
Methodology

The AFIT of Today is the Air Force of Tomorrow

Three elements required
- Estimates, Actuals, and Elapsed Time
- Predictions vs. “ground truth”
Definitions

- **O&S Cost**
  - Total cost to sustain weapon system after fielding

- **Annual Unit O&S Cost (AUC)**
  - Yearly cost to maintain per unit

- **Life Cycle Cost (LCC)**
  - Total cost to govt spanning all phases of the program’s life
  - Essentially LCC = Total Acq Costs + Total O&S Costs
Summary Statistics

The AFIT of Today is the Air Force of Tomorrow

- 470 SARs (observations)
- 36 MDAPs (24 Navy, 12 Air Force)
- 53% Aviation, 35% Maritime

SARs By Service Component

- 317 Air Force
- 153 Navy

SARs By System Type

- 249 Aviation
- 166 Maritime
- 55 Munition
AUC Estimate Errors

- Overall estimate accuracy (392 cases)
  - 84% of estimates had error >15%; 68% of estimates > 25%

---

**Mean Error**
- Navy: 5.6%
- Air Force: 17.7%

**Mean Abs. Error**
- Navy: 41.2%
- Air Force: 42.5%

**Data**
- Navy: 24 cases
- Air Force: 11 cases
AUC Estimate Errors

- Accuracy trends over time (35 programs)
  - For 15/35 programs, estimate accuracy did not improve over time

![Graph showing accuracy trends over time for different programs.]

- Maritime
- Aviation
- Munitions

Mean Slope of Abs Error:
- Navy: -1.7% / year
- AF: 1.3% / year

Navy (24), Air Force (11)
**LCC Estimate Errors**

- **Overall estimate accuracy (317 cases)**
  - 56% of estimates had error >15%; 38% of estimates ≤ 25%

![Graph showing error in LCC estimate over time](image)

<table>
<thead>
<tr>
<th></th>
<th>Mean Error</th>
<th>Mean Abs. Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Navy</td>
<td>-4.7%</td>
<td>22.4%</td>
</tr>
<tr>
<td>Navy Mean Error</td>
<td>3.2%</td>
<td>20.2%</td>
</tr>
<tr>
<td>Navy Mean Abs. Error</td>
<td>20.2%</td>
<td></td>
</tr>
<tr>
<td>AF Mean Error</td>
<td>-21.6%</td>
<td>27.1%</td>
</tr>
<tr>
<td>AF Mean Abs. Error</td>
<td>27.1%</td>
<td></td>
</tr>
</tbody>
</table>
LCC Estimate Errors

- Trends Over Time (31 Programs)
  - For 10/31 programs, estimate accuracy did not improve over time

![Graph showing trends over time for LCC estimate errors with data points for different categories such as Maritime, Aviation, Munitions, Navy (21), and Air Force (10). The graph includes mean slopes of absolute error with values:
  - Navy Mean Slope of Abs. Error: -1.8% / year
  - AF Mean Abs. Error: -1.3% / year
  - Mean Slope of Abs. Error: -1.7% / year]
Overall estimate accuracy (317 cases)
- 79% of estimates had error >15%; 62% of estimates ≤ 25%

<table>
<thead>
<tr>
<th></th>
<th>Mean Error</th>
<th>Mean Abs. Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Navy (21)</td>
<td>28.2%</td>
<td>46.1%</td>
</tr>
<tr>
<td>Air Force (10)</td>
<td>-25.6%</td>
<td>36.6%</td>
</tr>
</tbody>
</table>
Total O&S Cost Estimate Errors

- Trends Over Time (31 Programs)
  - For 15/31 programs, estimate accuracy did not improve over time

![Graph showing trends over time with errors for different programs, including Navy (21) and Air Force (10). Mean slopes and error rates are also displayed.]

- Navy Mean Slope of Abs. Error: -1.9% / year
- AF Mean Abs. Error: 0.1% / year

The AFIT of Today is the Air Force of Tomorrow
Key Findings

Accuracy of O&S-based cost estimates is poor and improves little over time

- AUC Estimates
  - Magnitude of mean errors \(~40\%\); reduces \(~1\%\) per year on average
- LCC Estimates
  - Magnitude of mean errors \(~20\%\); reduces \(~1.5\%\) per year on average
- Total O&S Cost Estimates
  - Magnitude of mean errors \(~40\%\); reduces \(~1\%\) per year on average

O&S cost estimates behave differently than acquisition cost estimates

- Consistently greater levels of inaccuracy
- Do not converge in the time spans of consideration
More Key Findings

- Error biases extend in both directions
- Estimate accuracy and trends vary significantly between the Navy and the Air Force
- Many other program elements exhibit significant relationships with estimate accuracy
  - Type of system
  - Size of acquisition effort
  - Procurement Quantity
  - Cost Variance Trends

Opportunity to improve cost estimating...
Implications

- **Future Studies**
  - Invalidates premise of extant O&S characterization studies

- **Lack of Accuracy Impacts Funding**
  - Underestimating cost creates challenges for entire portfolio
  - Overestimating cost liability creates opportunity loss

- **Lack of Convergence Affects Budgeting Strategy**
  - Decision-makers may be under the (apparently) false impression that later cost estimates are more reliable
More Implications

The AFIT of Today is the Air Force of Tomorrow

● Decision Analysis
  ● AUC frequently used to differentiate competing designs
  ● LCC is discriminator between competing programs or comparing cost-effectiveness of modifying vs. initiating a new acquisition
    ● Absolute accuracy less important than relative accuracy
    ● Relative accuracy appears no better than absolute

Calls into question validity of value decisions based on AUC/LCC estimates
Key Question

- Is it valid to compare O&S cost estimates to actuals?
  - Assumptions used to construct the estimates are often fundamentally different from what occurred in reality
    - Peacetime vs. war, commodity prices, healthcare costs, quantities, etc.
    - Uncertainty and long time horizon greatly complicate estimate
  - Corollary: Should programs be held accountable for O&S cost estimates?
Validity?

Response

● Is it valid to compare O&S cost estimates to actuals?
  ● Acq phase also lengthy and characterized by uncertainty
  ● Why demand accuracy/accountability for estimates that apply to acquisition phase but not sustainment?
  ● Is goal to have best estimate assuming current baseline is fixed or do we want best estimate in the real world of changing baselines?

● Should programs be held accountable?
  ● Absolutely!
  ● If estimates can’t be accurate—
    ● Why go to all the effort of building them?
    ● Why base key budgetary/programmatic decisions upon them?
Now What?

Explore the “why”
- Not the fault of cost estimators! Process is flawed

Build cost estimating models based on findings here

New model “corrects” original LCC estimates to achieve greatly improved accuracy
- “Macro-Stochastic” cost estimating
Summary

- Importance of O&S cost estimating accuracy
- DoD lacks insight into current accuracy levels
- O&S cost estimates are very poor and improve little
- Patterns exist in estimate accuracy
- Opportunities exist to improve estimates
- Embrace Uncertainty!
  - An otherwise “perfect” cost estimate constrained by today’s baseline is bound to be wrong tomorrow
  - Decision makers need an estimate that accounts for uncertainty
More Information

“A Proposed Methodology to Characterize the Accuracy of Life Cycle Cost Estimates for DoD Programs”
- *Procedia Computer Science*

“Characterizing the Accuracy of DoD Operating and Support Cost Estimates”
- *Journal of Public Procurement*

“A Macro-Stochastic Model for Improving the Accuracy of DoD Life Cycle Cost Estimates”
- *Journal of Cost Analysis and Parametrics*
BACKUPS
Methodological Concerns

- Inherent dataset biases
  - Programs that provided "good" data
  - Programs that "succeed"
  - Phasing of actuals (majority of actual costs incurred during war)
Methodological Concerns

- Incomplete Data
  - Must infer LCC from partial lifecycle actuals
  - Op Service Life is held constant
  - No escalation factor applied (i.e., CGAI)
    - May skew results, but not likely to change general findings
Methodological Concerns

- Data Integrity
  - Errors in Prediction: Reliability of SARs
  - Errors in Actuals: Reliability of VAMOSC

- Scope of Applicability
  - MDAPs
  - Air Force and Navy programs

- Precursor paper details full methodology
Mixed Models

- **Mixed models compensate for correlated errors**
  - Can account for subject observations not independent
  - Allow data to exhibit inherent correlations and non-constant variability that arise from the data hierarchy
  - Some regression parameters are population-specific (fixed-effects)
  - Other parameters are subject-specific (random-effects)

\[ y = X\beta + Z\gamma + \varepsilon \]

- \( y \) = Observed data vector
- \( X \) = Fixed-Effect Design Matrix
- \( \beta \) = Vector of Fixed-Effect Parameter Estimates (same for all subjects)
- \( Z \) = Random-Effect Design Matrix
- \( \gamma \) = Vector of Random-Effect Parameter Estimates (varies by subject)
- \( \varepsilon \) = Vector of Residual Errors