Time–Phasing Methods and Metrics

Erik Burgess
MCR, Inc.
(703) 633-2128
eburgess@mcri.com

37th Annual DoD Cost Analysis Symposium
Williamsburg, VA
10-13 February 2004

This research was initiated by the Intelligence Community Cost Analysis Improvement Group (IC CAIG) and jointly sponsored by the National Reconnaissance Office Cost Group (NCG). Views expressed herein do not reflect the official policy or position of the IC CAIG, NCG, or the U.S. government.
Time–Phasing Methods and Metrics

Erik Burgess
MCR, Inc.
(703) 633-2128
eburgess@mcri.com

37th Annual DoD Cost Analysis Symposium
Williamsburg, VA
10-13 February 2004

This research was initiated by the Intelligence Community Cost Analysis Improvement Group (IC CAIG) and jointly sponsored by the National Reconnaissance Office Cost Group (NCG). Views expressed herein do not reflect the official policy or position of the IC CAIG, NCG, or the U.S. government.
Contents

• Review

• Four common pitfalls
  – Lack of accuracy metrics
  – Trouble incorporating independent variables
  – Ambiguous start and end dates
  – Failure to convert cost to budget

• Space-system case study
Financial Management Terms That Matter

- **Budget Authority**: Authority by law to incur obligations

- **Obligation Authority, Total Obligation Authority (TOA)**: Budget authority plus other funds available for obligation

- **Obligations**: Contract award

- **Costs Incurred**: Actual cost of work performed

- **Expenditures**: Sending the check

- **Outlays**: Cashing the check
# Expenditures Lag Budgets by Years!

<table>
<thead>
<tr>
<th>Years</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>O&amp;M</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RDT&amp;E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Procurement</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MILCON</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Obligation Period

Expenditure Period

Source: DSMC (Tack 1997)

## Example: Outlay Rates (% of Year-1 TOA expended)

<table>
<thead>
<tr>
<th></th>
<th>Year 1</th>
<th>Year 2</th>
<th>Year 3</th>
<th>Year 4</th>
<th>Year 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>AF RDT&amp;E</td>
<td>60%</td>
<td>32%</td>
<td>4%</td>
<td>2%</td>
<td>2%</td>
</tr>
</tbody>
</table>

Program has Multiple Time-Phased Profiles
High-Level Phasing Models

- Rayleigh distribution provides good fit for development programs (Abernathy 1984; Elrod 1993; Lee et al. 1993; Watkins 1982)

\[ E(t) = d \left[ 1 - e^{-\alpha t^2} \right] \]

- Weibull distribution is better (Brown et al. 2002; Porter 2001; Unger 2001)

\[ E(t) = d \left[ 1 - e^{-\left(\frac{t-\gamma}{\delta}\right)^\beta} \right] \]
Rayleigh Model

- Underlying theory from Peter Norden (1970): Cumulative cost incurred, $E(t)$, is product of two functions
  - Percent of work remaining
  - Linear increase in skills/knowledge acquired

- General form:
  $$E(t) = d \left(1 - e^{-\alpha t^2}\right)$$

- Infinite tail must be truncated, giving
  $$d = D / E(t_c) = D / \left(1 - e^{-\alpha t_c^2}\right)$$

  \( t_c \) = time of completion

  \( D \) = total cost

- Parameter $\alpha$ determines time of peak expenditures (front/back loading)
How Much Front-Loading?

- Gallagher and Lee (1996) propose to fix completion time as when 97% of total funds are expended
  - Results in peak expenditures at 38% complete
  - Results in 60% spent at 50% time (a.k.a. 60/50 Rayleigh)
  - Truncation and resulting fixed Rayleigh now often referred to as “convention”
  - Rayleigh often mischaracterized as having fixed time of peak expenditures (not true)

- Case study of 14 single-satellite contracts gives average of 59% spent at 50% time

- Earlier study of 69 Air Force programs (ESC 1995) gives average of 64% spent at 50% time

60/50 Rayleigh Appears Appropriate When Commodity-Specific Models not Available
Contents

• Review

• Four common pitfalls
  – Lack of accuracy metrics
  – Trouble incorporating independent variables
  – Ambiguous start and end dates
  – Failure to convert cost to budget

• Space-system case study
Need for Accuracy Metrics

• Predictive accuracy metrics needed
  – To compare models
  – To communicate with budget analysts

• Goodness-of-fit statistics often reported
  – Specific curves against individual programs
  – Not useful as predictive metric

• Metrics should be:
  – Based on actual vs. estimated costs
  – Independent of functional form (e.g., Weibull)
  – Independent of regression approach
Cumulative Costs, Spend Rates

Cum-cost data:
As reported by contractor
Basis for regressions
Used for 2 metrics:
1. Std. dev. of all cum-cost residuals
2. Std. dev. of cum-cost residuals @ 40% complete

Expenditure-rate data:
Derived from cum-cost data
Used for 2 metrics:
3. Coefficient of Variation
4. Pearson’s $R^2$
1. SD of all cum-cost residuals = 6.3%
   *Indicates relative overall accuracy*

2. SD of residuals @ 40% complete = 9.8%
   *Indicates confidence range through critical early years*
Metric #3:
Pearson $R^2$ of Actual vs. Estimated Rate

3. Expenditure-Rate Pearson $R^2 = 0.73$

*Indicates amount of variation in annual costs explained by model*
Metric #4: Expenditure-Rate Coefficient of Variation

4. CV of Expenditure Rate = 32%
Indicates average confidence range of costs in any one year

Space-system case study
Contents

• Review

• Four common pitfalls
  – Lack of accuracy metrics
  – Trouble incorporating independent variables
  – Ambiguous start and end dates
  – Failure to convert cost to budget

• Space-system case study
All Data: Cumulative Cost vs. Time

26 Contracts
270 Individual Data Points

Costs Peak Earlier than Average
Costs Peak Later than Average

Average 65% spent at 50% Time (65/50)

Space-system case study

Modeling Objective: Improve Accuracy using Explanatory Variables
Results: Simple Curves

<table>
<thead>
<tr>
<th>Method</th>
<th>Result</th>
<th>SD of Cum. Residuals*</th>
<th>SD of Cum. at 40% complete</th>
<th>Pearson R^2 of Exp. Rate</th>
<th>CV of Exp. Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single Rayleigh Distribution</td>
<td>alpha = 4.17</td>
<td>9.6%</td>
<td>15.0%</td>
<td>0.55</td>
<td>44%</td>
</tr>
<tr>
<td>Single Weibull Distribution</td>
<td>alpha = 2.97 beta = 1.64</td>
<td>9.2%</td>
<td>14.5%</td>
<td>0.58</td>
<td>40%</td>
</tr>
<tr>
<td>Single Beta Distribution</td>
<td>alpha = 1.46 beta = 2.04</td>
<td>9.6%</td>
<td>14.7%</td>
<td>0.49</td>
<td>44%</td>
</tr>
</tbody>
</table>

Space-system case study

- Single curves can be fit to aggregation of all data (no independent variables)
- Shows no significant difference among curve forms
- Provides baseline for improvement
- Not a bad model

* Curve fit minimizes this parameter
Adding Drivers: Multi-Stage Regression

\[ j = 26 \text{ profiles with } i \text{ data points each} \]

\[ \% \text{cost, } E_i \text{ vs. } \% \text{time, } t_i \]

**STAGE 1**

- Run 26 regressions to estimate 26 pairs of parameters \( \alpha, \beta \)

**STAGE 2**

- Run one regression to estimate \( a, b, c \)

**STAGE 3**

- Run one regression to estimate \( d, e, f \)

Minimizing

\[ E(\alpha_j, \beta_j) = \sum_i \left( \frac{1 - e^{-\alpha_j t_i}}{1 - e^{-\alpha_j}} - E(t_i) \right)^2 \]

Results in \( j \) sets of parameters \( \alpha, \beta \)

**Multi-Stage Regression Shortfalls**

- Final model is not based on minimizing cost errors
- Parameters \( \alpha, \beta \) estimated independently
Adding Drivers: Single-Stage Regression

\[ j = 26 \text{ profiles with } i \text{ data points each} \]
\[ 270 \text{ total data points } (i^*j) \]
\[ \%\text{cost}, E_{i,j} \text{ vs. } \%\text{time}, t_{i,j} \]

Run one regression to estimate \( a, b, c, d, e, f \)

Minimizing

\[ E(a,b,c,d,e,f) = \sum_{i,j} \left( \frac{1 - e^{-\alpha t^*_j}}{1 - e^{-\alpha}} - E(t_{i,j}) \right)^2, \]

where

\[ \alpha = a + b \times \text{duration}_j + c \times \%\text{NR}_j \]
\[ \beta = d + e \times \text{duration}_j + f \times \%\text{NR}_j \]

Single-Stage Regression is More Accurate

- Directly minimizes cost errors
- Parameters \( \alpha, \beta \) estimated simultaneously
### Single vs. Multi-Stage Results

<table>
<thead>
<tr>
<th>Method</th>
<th>Result</th>
<th>SD of Cum. Residuals</th>
<th>SD of Cum. at 40% complete</th>
<th>Pearson $R^2$ of Exp. Rate</th>
<th>CV of Exp. Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weibull Model (Single-stage Regression)</td>
<td>$\alpha, \beta = f(%NR, Duration, #units)$</td>
<td>6.3%</td>
<td>9.9%</td>
<td>0.72</td>
<td>33%</td>
</tr>
<tr>
<td>Weibull Model (Multi-stage Regression)</td>
<td>$\alpha, \beta = f(%NR, Duration, #units)$</td>
<td>7.4%</td>
<td>11.9%</td>
<td>0.68</td>
<td>35%</td>
</tr>
</tbody>
</table>

Space-system case study

**Single-stage results are better in all four metrics**
Contents

• Review

• Four common pitfalls
  – Lack of accuracy metrics
  – Trouble incorporating independent variables
  – Ambiguous start and end dates
  – Failure to convert cost to budget

• Space-system case study
Defining Start, End Dates

- Must have precise definitions
  - Indexed to programmatic events
  - Common definition across programs, contractors

- Ambiguous definition of “time” results in baseless accuracy metrics

GOOD
Contract Award
Launch

WORSE
PDR, CDR
IOC, FOC
Milestone B
Need for Independent Schedule Estimate

Don’t phase costs to PM’s optimistic schedule
- Small difference requires large funding increase in early years
- Even worse if ICE is higher than PM estimate
Integrated Cost, Schedule, Phasing Result

Rate ($ per year)

Risk Dollars

Aggressive Schedule
Lower Cost Estimate

Expected Schedule
Higher Cost Estimate
Integrated Time-Phasing Process

1. Develop Integrated Cost-Risk Estimate
2. Determine which point estimate to phase (e.g., mean*)
3. Segregate into space, ground

Develop schedule estimate—based on weight, design life, # payloads

Develop time-phased expenditures—based on #units, %NR, duration

Develop time-phased budget—based on outlay rates

*May require fancy math if it's anything other than the mean
Contents

• Review

• Four common pitfalls
  – Lack of accuracy metrics
  – Trouble incorporating independent variables
  – Ambiguous start and end dates
  – Failure to convert cost to budget

• Space-system case study
Converting Cost to Budget

**WARNING:** Do not use cost profiles for budgeting

- Required Budget Authority exceeds expected costs incurred in early program years
  - Even if all budget is obligated each year
  - Difference published by DoD (outlay rates)

- Two analytical procedures for converting cost to budget proposed by Lee et al. (1997)
  1. Constrained nonlinear estimation
  2. Linear system with truncation
     - Easy to implement, distribute as spreadsheet tool
     - No iteration required
Cost-to-Budget Mechanics

**STEP 1:** Estimate Time-Phased Expenditures in BY$

![Bar chart showing annual expenditures (BY82$M) for FY82 to FY88.]

**STEP 2:** Convert to Real Dollars using raw indices

<table>
<thead>
<tr>
<th>Fiscal Year</th>
<th>Raw Inflation Index</th>
<th>Expenditure Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>FY82</td>
<td>1.000</td>
<td>56653</td>
</tr>
<tr>
<td>FY83</td>
<td>1.049</td>
<td>157796</td>
</tr>
<tr>
<td>FY84</td>
<td>1.089</td>
<td>166918</td>
</tr>
<tr>
<td>FY85</td>
<td>1.126</td>
<td>152075</td>
</tr>
<tr>
<td>FY86</td>
<td>1.157</td>
<td>122438</td>
</tr>
<tr>
<td>FY87</td>
<td>1.189</td>
<td>54209</td>
</tr>
<tr>
<td>FY88</td>
<td>1.224</td>
<td>20297</td>
</tr>
<tr>
<td>TOTAL</td>
<td></td>
<td>730387</td>
</tr>
</tbody>
</table>

**STEP 3:** Convert to Then-Year Budget Using Outlay Rates

Solve linear system:

\[ TOA_k = \left( e_k - s_2 TOA_{k-1} - s_3 TOA_{k-2} - \cdots - s_j TOA_{k-j+1} \right)/s_1 \]
Resulting Budget vs. Expenditure Profile

Often a big difference between cost and budget profiles.
Short-Cut Approach

• Converts costs directly to budget profile
  – No other calculations involved
  – Based on same underlying 60/50 Rayleigh model (38% peak)
  – Assumes AF 3600 RDT&E outlay rates

• Uses Beta distribution

<table>
<thead>
<tr>
<th>Number of Program Years</th>
<th>Alpha</th>
<th>Beta</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>2.65</td>
<td>5.69</td>
</tr>
<tr>
<td>4</td>
<td>2.47</td>
<td>4.95</td>
</tr>
<tr>
<td>5</td>
<td>2.11</td>
<td>3.87</td>
</tr>
<tr>
<td>6</td>
<td>1.97</td>
<td>3.38</td>
</tr>
<tr>
<td>7</td>
<td>1.90</td>
<td>3.10</td>
</tr>
<tr>
<td>8</td>
<td>1.88</td>
<td>3.00</td>
</tr>
<tr>
<td>9</td>
<td>1.86</td>
<td>2.88</td>
</tr>
<tr>
<td>10</td>
<td>1.84</td>
<td>2.80</td>
</tr>
</tbody>
</table>

Example

Use MS-Excel “BETADIST” Function
Contents

• Review

• Four common pitfalls
  – Lack of accuracy metrics
  – Trouble incorporating independent variables
  – Ambiguous start and end dates
  – Failure to convert cost to budget

• Space-system case study
Space-System Phasing Model: Ground Rules

• Space-segment costs
  – Generally includes space SEIT/PM
  – Does not include Systems Integrator, CAAS/SETA, etc.

• Expenditures, not budgets

• Based on actual costs and schedules of completed contracts (26 NRO and Air Force)
Weibull-Based Model Underestimates Launch-Year Costs

…especially for long, multi-unit programs
Solution: Modify Weibull Curve Form

• Add term that represents a constant expenditure rate ($R$):

Weibull Function

$$E(t) = d \left[1 - e^{-\alpha t^\beta}\right]$$

$$d = \frac{\text{total cost}}{1 - e^{-\alpha}}$$

Weibull + Constant Rate ($R$)

$$E(t) = d \left[Rt + 1 - e^{-\alpha t^\beta}\right]$$

$$d = \frac{\text{total cost}}{R + 1 - e^{-\alpha}}$$

• Rate, $R$, is a function of duration

• Improves launch-tail and overall accuracy of phasing model
Launch-Year Expenditures: New Model

Near-zero bias for short and long programs

[Graph showing Cost Residual at Launch vs. Duration (months) with Weibull + Constant Rate model (May 2003). The graph indicates that cost residuals are underestimated for short programs and overestimated for long programs.]
Result: Weibull + Constant-Rate Model

<table>
<thead>
<tr>
<th>Method</th>
<th>Drivers</th>
<th>SD of Cum. Residuals</th>
<th>SD of Cum. at 40% complete</th>
<th>Pearsons $R^2$ of Exp. Rate</th>
<th>CV of Exp. Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weibull Model C</td>
<td>$\alpha = f(\text{units}, \text{duration}, %\text{NR})$ $\beta = 1.71$</td>
<td>6.3%</td>
<td>9.9%</td>
<td>0.72</td>
<td>33%</td>
</tr>
<tr>
<td>Weibull + Constant Rate</td>
<td>$\alpha = f(\text{units}, \text{duration}, %\text{NR})$ $\beta = 1.71$ $R = f(\text{duration})$</td>
<td>6.3%</td>
<td>9.8%</td>
<td>0.73</td>
<td>32%</td>
</tr>
</tbody>
</table>

- **Weibull + Constant Rate** model is marginally better in 3 of 4 metrics, and solves launch-year issue
- Rate term is proportional to program duration

\[
E(t) = d \left[ Rt + 1 - e^{-\alpha t^\beta} \right]
\]

\[
d = \frac{\text{total cost}}{R + 1 - e^{-\alpha}}
\]

$t = \text{cumulative time/total time}$

$\alpha = -0.414 + 0.0729(\text{units}) + 0.0488(\text{months duration}) + 0.0145(\text{percent nonrecurring})$

$\beta = 1.71$

$R = 0.00148(\text{months duration})$
Typical Profile: Good or Bad Fit?

Model is Accurate for Cum Costs
Actual: 63% spent @ 44% time
Model: 63% spent @ 44% time

But no smooth curve hits every year

Pearson's R² = 0.88
Ranked 8 out of 26

Expenditure Rate (%Cost/%Time) vs. Time

0.000 0.040 0.120 0.200 0.280 0.360 0.440 0.520 0.600 0.680 0.760 0.840 0.920 1.000

0.000 0.050 0.100 0.150 0.200 0.250 0.300 0.350 0.400 0.450 0.500 0.550 0.600 0.650 0.700 0.750 0.800 0.850 0.900 0.950 1.000

0.000 0.040 0.120 0.200 0.280 0.360 0.440 0.520 0.600 0.680 0.760 0.840 0.920 1.000
Satellite Schedule Model: Ground Rules

- Schedule duration is from contract award to first-launch availability
  - Launch availability for first satellite in series
  - If not stored, then launch availability date = launch date
  - If stored, then launch availability = launch date minus factory storage time
  - Consistent with time-phasing model

- Multiple independent variables investigated
  - Dry weight
  - Power
  - Design Life
  - Orbital regime
  - Year of Award
  - NR/AUC Ratio
  - New vs. Replacement Capability
  - Qual/protoflight approach
  - Execution rate compared to phasing model
  - Number of distinct payloads
  - Time from award to Preliminary Design Review (PDR)
Satellite Schedule Model: Development Process

- Multiple databases assembled, investigated
  - All data
  - Govt only
  - Gov’t only, no Class C/D (experimental smallsats)
  - Gov’t only, new
  - Military only (NRO, AF, Navy)
  - NRO only

- Several stratification variables assessed
  - IMINT/Remote Sensor
  - SIGINT
  - Class C/D (experimental smallsats)
  - LEO Orbit
  - Commercial
  - NRO vs. AF/Navy
  - MIL vs. NASA

Over 150 Models developed and compared
Satellite Schedule Model: Findings

• Must segregate NASA, Commercial programs
  – Different drivers
  – Different durations for same drivers

• Class C/D (a.k.a. experimental) not a driver
  – Technical aspects of program are sufficient (short design life and low weight)

• NR/AUC ratio not a driver in any data subset

• NR+T1 cost is a poor predictor

• BOL power is a poor performer compared to weight
  – In all data subsets
  – In combination with all other variables investigated
Schedule Model: NRO+AF Dataset

Time To First Launch Availability (TT1L)

\[ TT1L = 17.0 + 0.87 \times (\text{dry wt})^{0.406} \times (\text{DesLife} \times \text{Payloads})^{0.136} \]

SEE = 25%
Pearson’s $R^2 = 0.69$
56 Observations

Dry weight in pounds
Design Life in Months
Payloads = number with physically distinct hardware and different users
Schedule-Phasing Interaction

- If Schedule Model Estimates Long Duration (~90+ months)
  - Phasing model is front loaded
  - Typical 80/50

- If Schedule Model Estimates Average Duration (~65 months)
  - Phasing model more “even”
  - Typical 60/50
Summary

- Predictive accuracy can be measured several ways

- Use of independent variables to drive profile shape improves accuracy
  - Single-stage regression gives better results than traditional multi-stage approach
  - Choice of functional form (e.g., Beta/Rayleigh/Weibull) has little effect on accuracy of final model
  - 60/50 Rayleigh is good choice for most development programs

- Start and end-dates must be well defined
  - Don’t phase to PM’s aggressive schedule
  - Use independent schedule estimate

- Cost profiles are not budget profiles
References