

“Trust, but verify is a form of advice given which recommends that while a source of information might be considered reliable, one should perform additional research to verify that such information is accurate, or trustworthy. The original Russian proverb is a short rhyme which states, Доверяй, но проверяй (doveryai, no proveryai).”
-Wikipedia

Trust, but Verify: An Improved Estimating Technique Using the Integrated Master Schedule (IMS)

Eric Lofgren, Technomics
ICEAA 2014

13 June 8:00am MDT, Colorado J
Earned Value Management Track

Introduction

- This presentation will show that by enforcing the baseline of an Integrated Master Schedule (IMS) through subsequent submissions, one may predict a more accurate schedule end date earlier in the contract
- This argument will be supported by analysis of actual Major Defense Acquisition Program (MDAP) contract IMSs collected from contract initiation to close-out

Data & Methods:

Utilized the Earned Value Management Central Repository (EVM-CR) for data access to MDAP contracts. <http://dcarc.cape.osd.mil/EVM/EVMOverview.aspx>

Analyzed all contract schedules which had MS Project IMSs spanning from contract start to end

- Data not spanning entire timespan not instructive in this analysis
- Not realistic to analyze PDFs; no access to a Primavera license
- Currently 12 contracts with 133 schedule observations

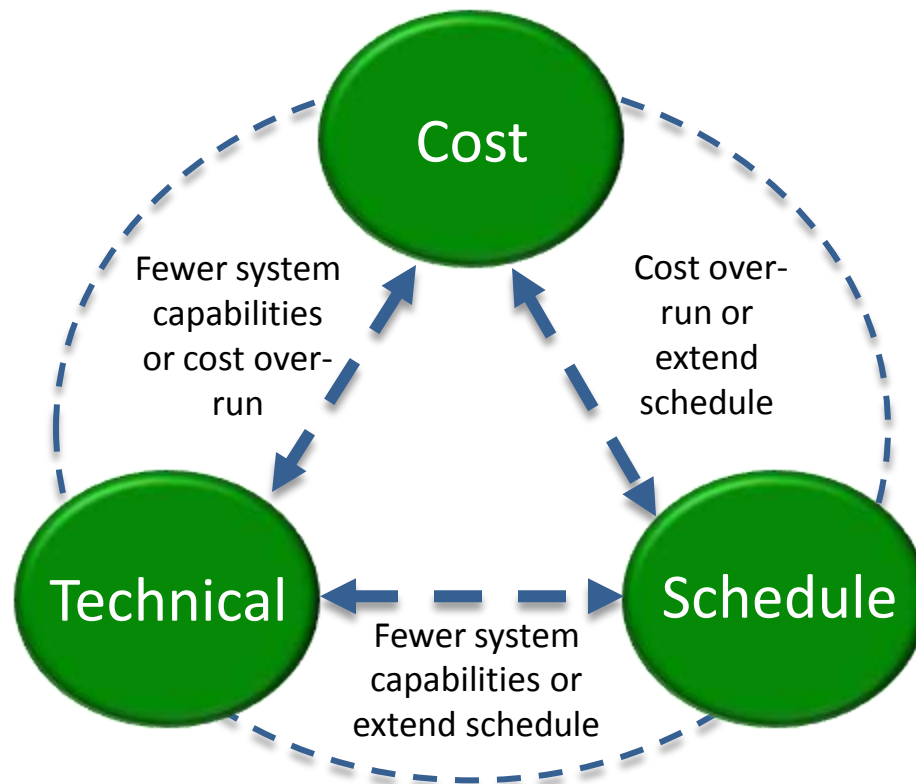
Uniform data extraction methodology across contracts and schedules

- Only non-summary activities
- 19 standard MS Project fields
- Converted into MS Excel flat-files
- Standard MS Excel template for analysis and gathering metrics

Used Defense Acquisition Management Information Retrieval (DAMIR) for additional contract insight. <http://www.acq.osd.mil/damir>

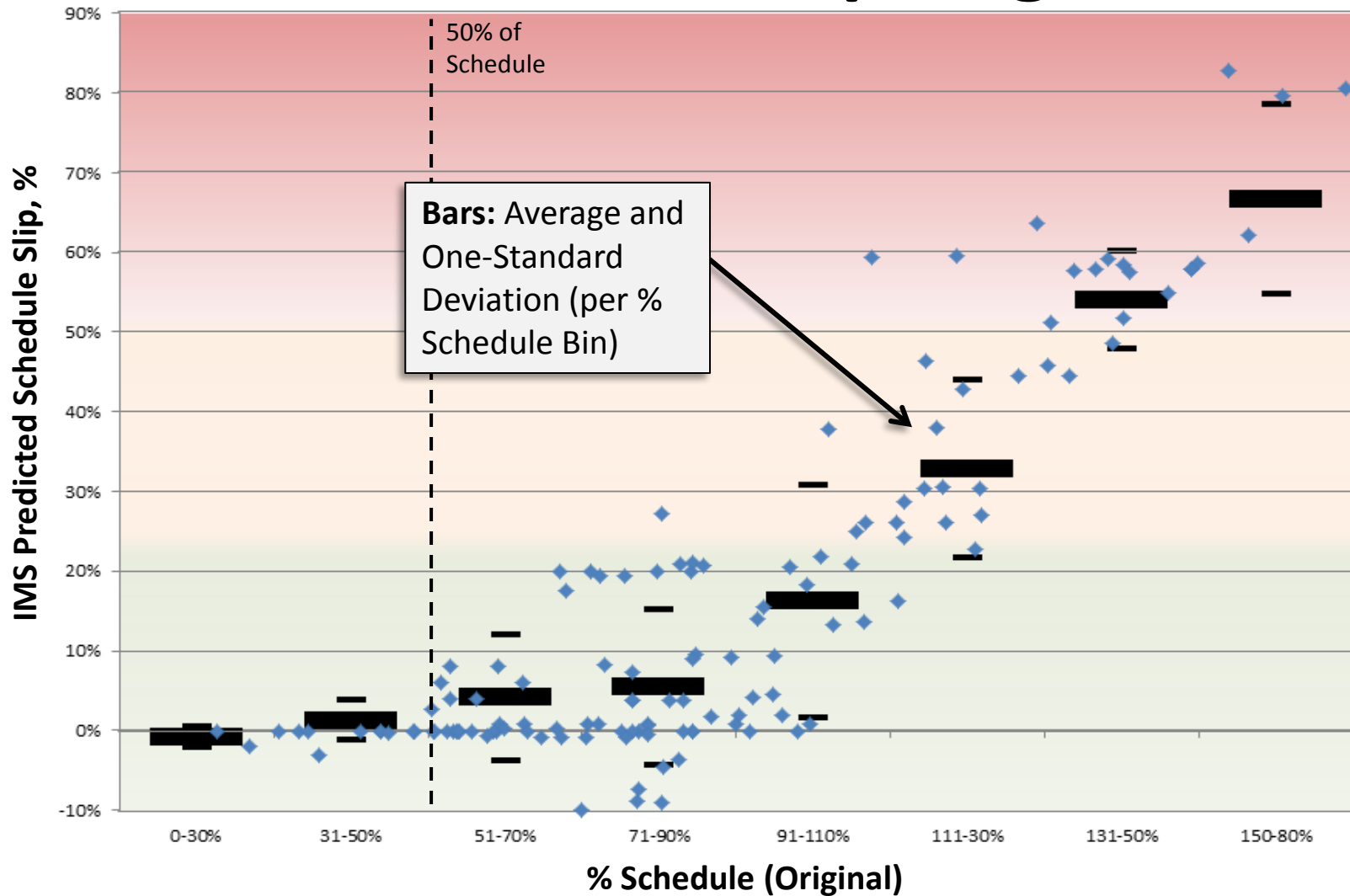
The IMS in Context

- Contract performance risk comprises three categories: cost, schedule, and technical
 - In general, realized risk in any one area can only be offset at the expense of performance in one or both of the others
- For a decision-maker, the value of an IMS is its ability to evaluate risk associated with schedule and provide an early warning for schedule slip.
 - What-if scenarios are also crucial for assessing impacts of technical issues and cost overruns



Note: there is often a positive correlation amongst the three: sometimes poor performance can only be mitigated

Problem: Schedule Slips Signaled Late



For many MDAP contracts, the IMS has a poor history of accurately reflecting schedule risk. Schedule slips are not registered until late in the project.

Status Quo Metrics Not Good Enough

- Because only the schedule itself can answer the question “when will the deliverable arrive?” a plethora schedule metrics have been devised. For example:
 - Baseline Execution Index (BEI): efficiency with which tasks have been accomplished when measured against the baseline tasks¹
 - Critical Path Length Index (CPLI): a measure of efficiency required to complete a milestone on-time²
- Besides evaluating quality, what good is made of metrics like the BEI and CPLI?
 - They can’t measure an end date other than that reflected in the current schedule
 - They give some indication of schedule performance and risk, but the forecast invariably paints a brighter picture
- Many find it difficult to ascertain the realism of any given schedule
 - Schedule **quality** is necessary but not sufficient condition for schedule **realism**
 - Follow the GAO Schedule Assessment Guide best practices
 - DCMA 14-Point Schedule Metrics for IMS
 - ❖ Test on logic; leads; lags; relationship types; constraints float; duration; resources; etc.
 - Reference the Joint Cost and Schedule Risk and Uncertainty Handbook

¹ See DCMA, “Earned Value Management System (EVMS) Program Analysis Pamphlet (PAP)”, pp. 16-18.

² See DCMA, “Earned Value Management System (EVMS) Program Analysis Pamphlet (PAP)”, pp. 18-20.

The Case for a New Metric



P-51 Mustang



F-15 Eagle



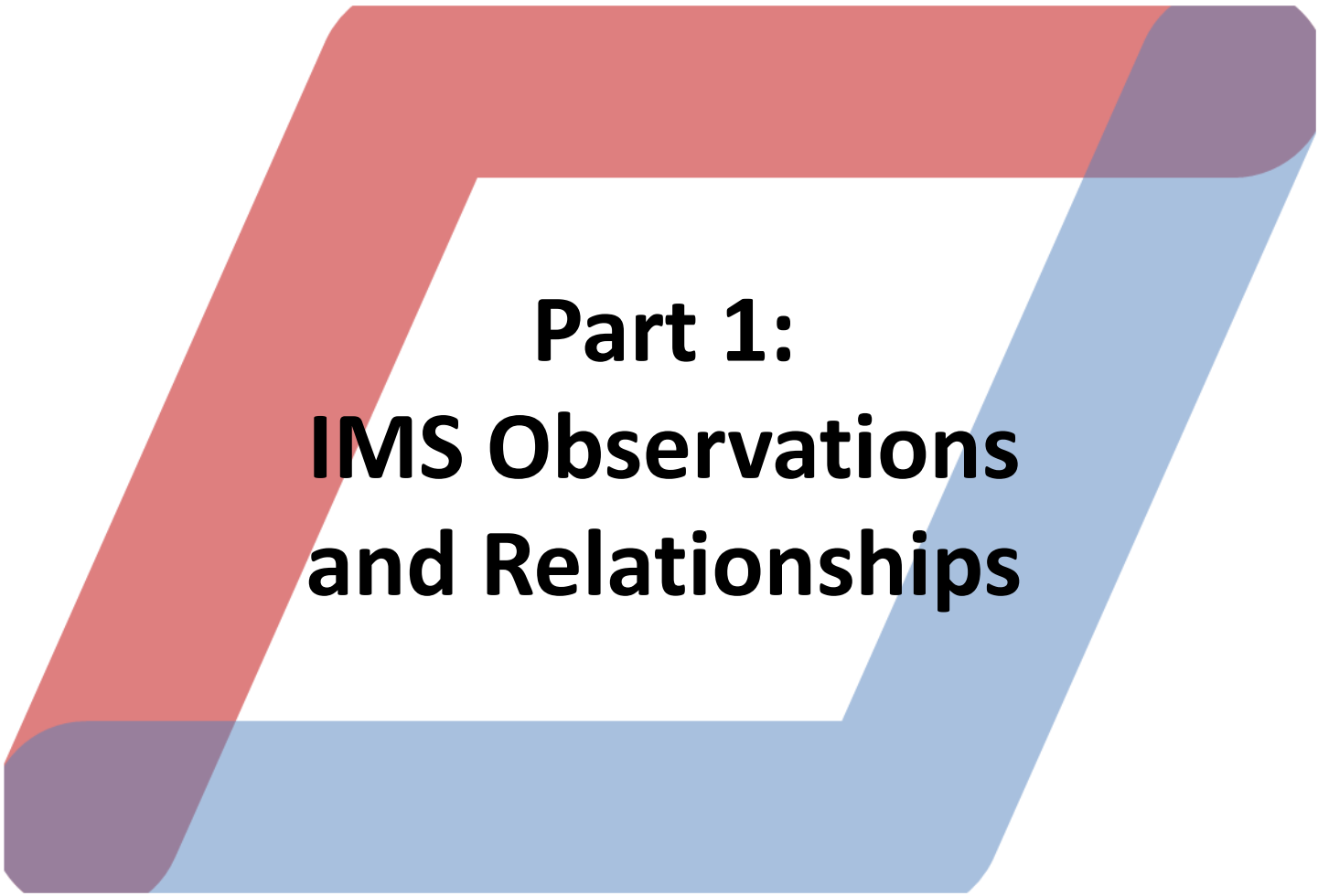
F-35 Lightning II

Above: Evolution of the fighter. How do current challenges in system development compare to the past and what is their effect on schedule estimation?

1. While schedules are a “living document,” planners tend to know the major activities involved in the execution of a project
 - There’s nothing new under the sun - *Ecclesiastes*
 - Contractors generally have well-defined processes
2. Whenever assessing an IMS, it is important to understand how it has evolved
 - What baseline changes have occurred?
 - How was performance to last submission’s plan?

An activity’s baseline from the initial IMS is relevant through subsequent submissions

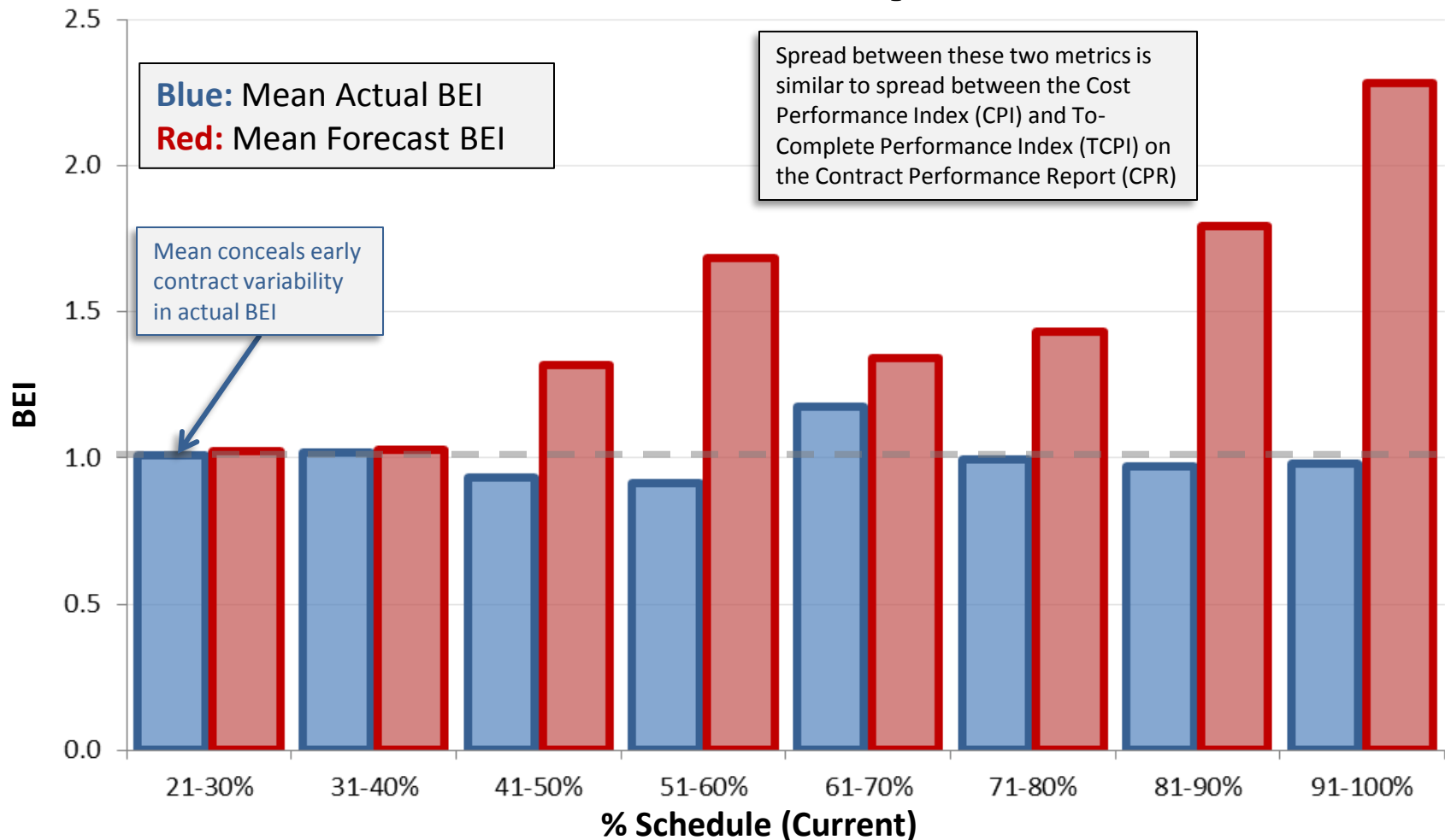
The Approach: By tracing near-term activities through subsequent IMSs and comparing them to their original baseline, as opposed to the “current” baseline, one may extrapolate a more realistic contract finish date far earlier in the project



Part 1:

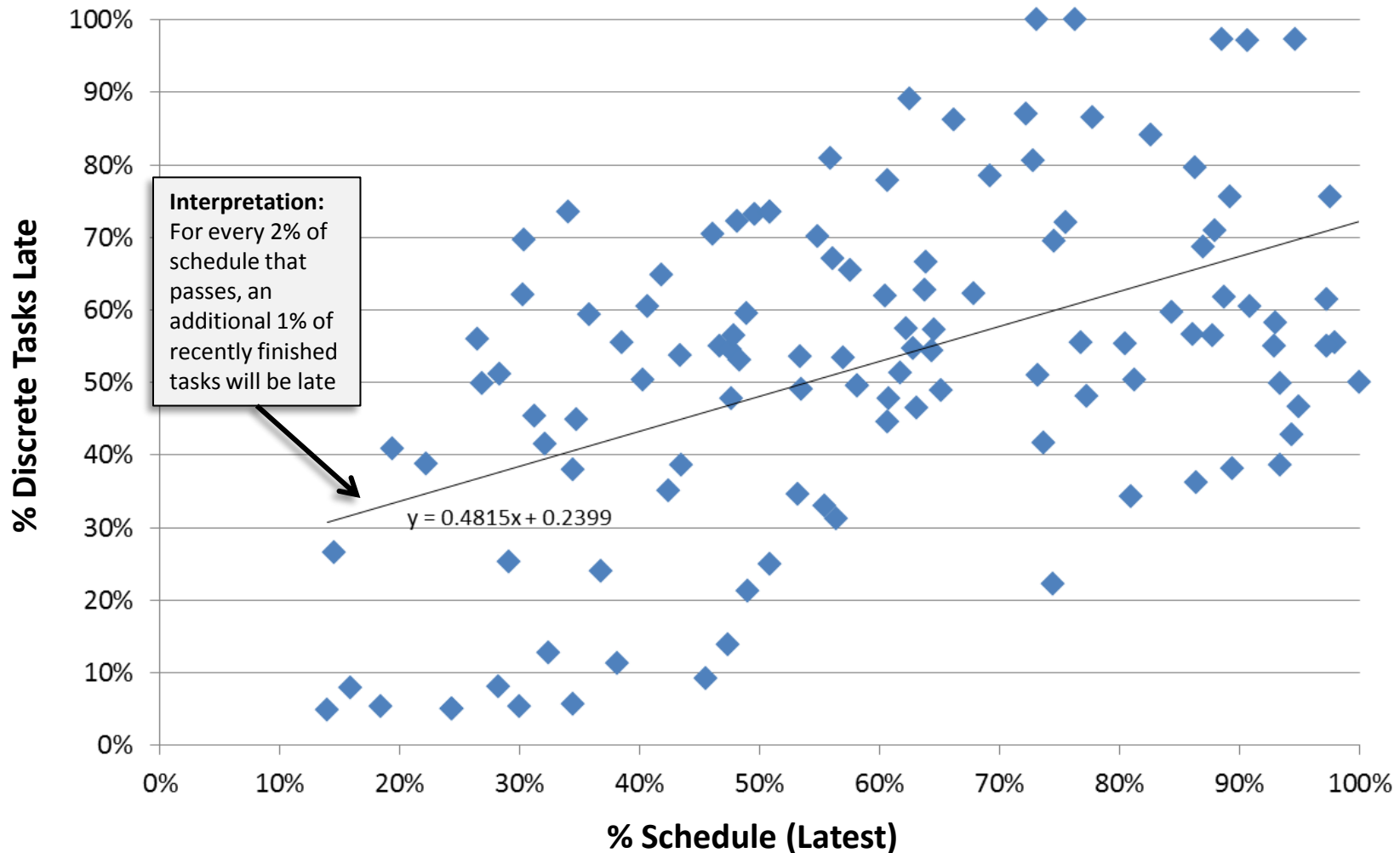
IMS Observations and Relationships

Forecasts are Optimistic



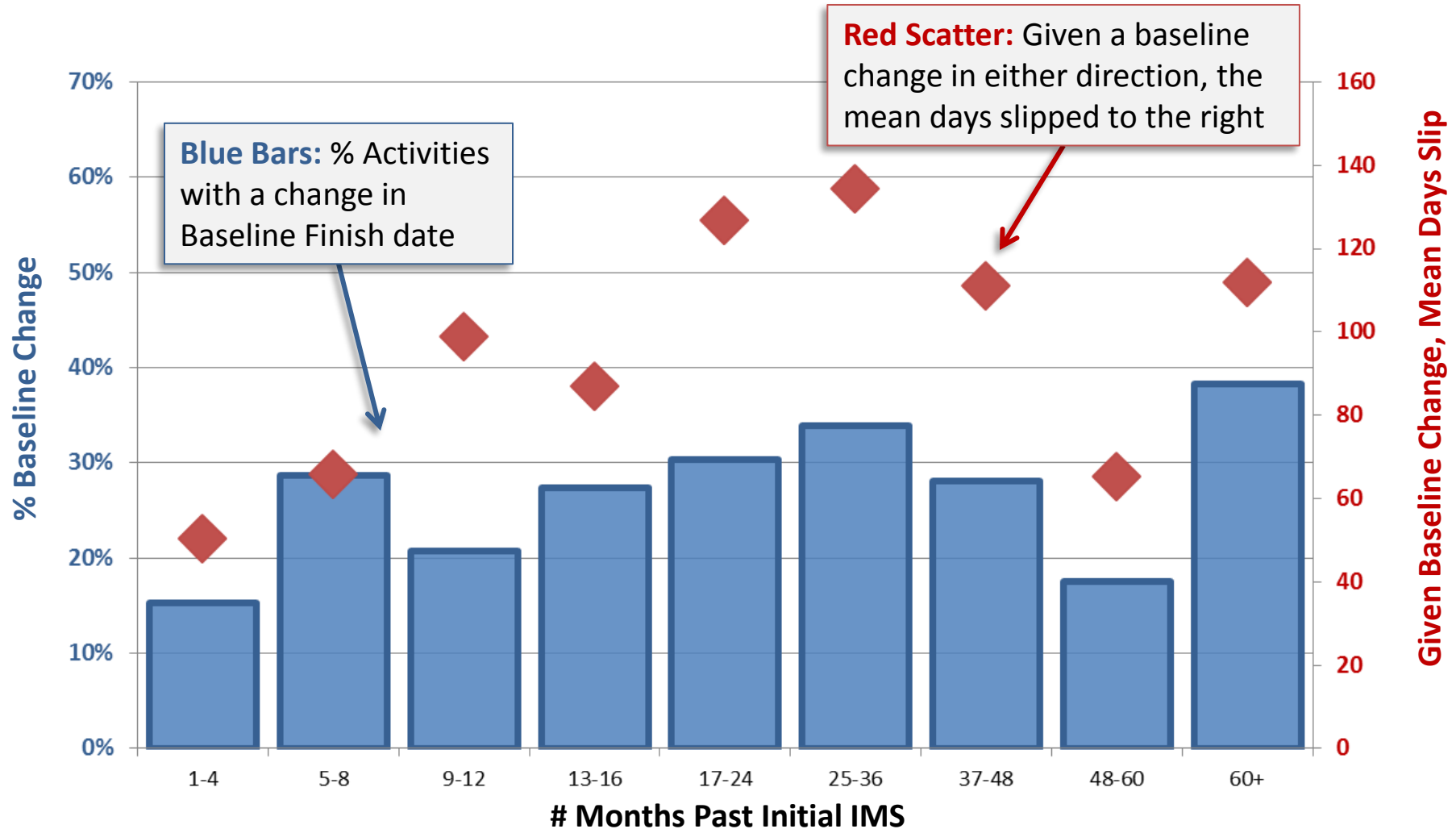
While mean actual BEIs do not deviate far from 1.0, the forecast BEI on future discrete tasks becomes increasingly optimistic. Relative discrepancies between actuals and forecasts may reflect current near-term schedule slips.

Task Performance Rarely Improves



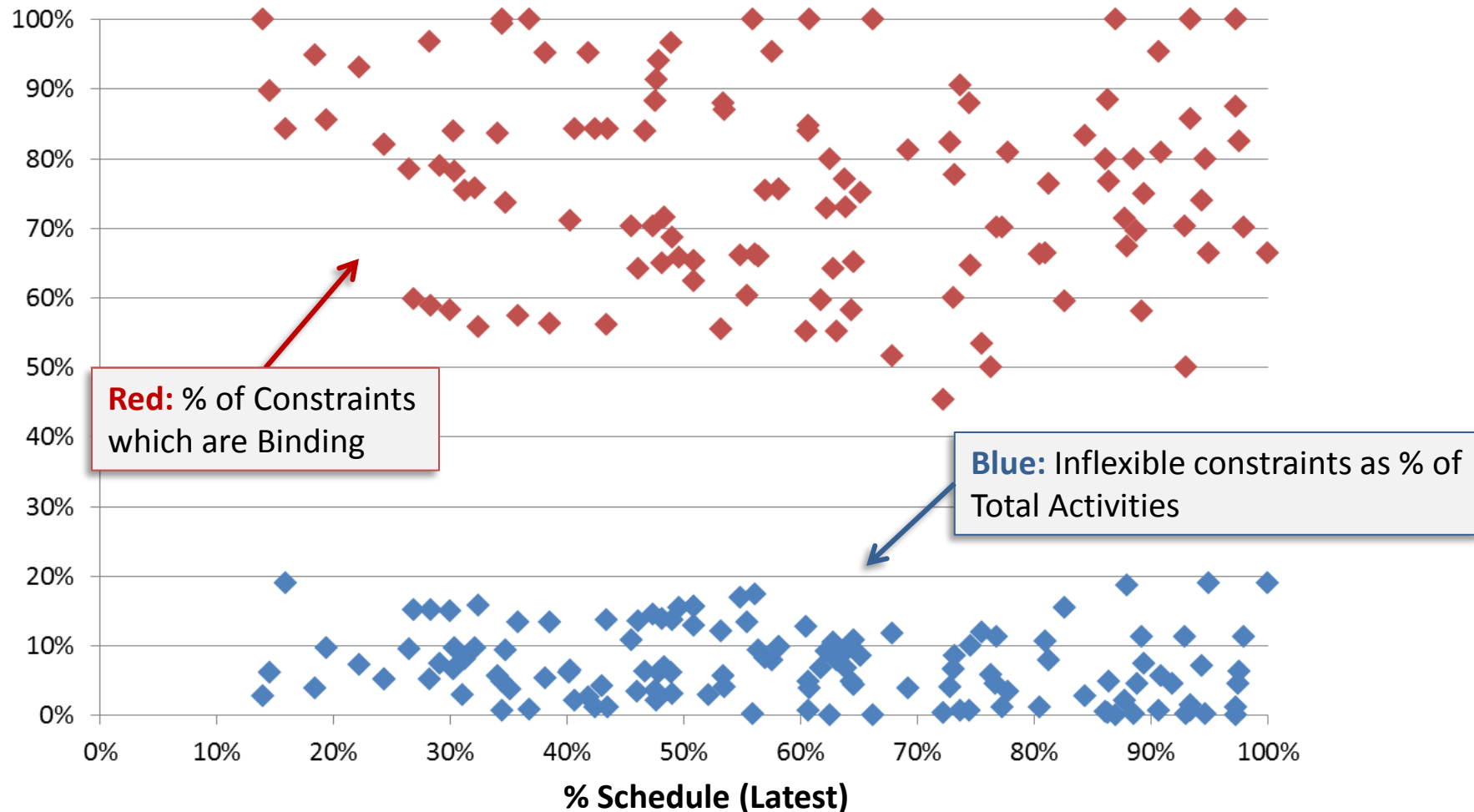
Of discrete tasks which finished in the past 3 quarters, the percent which are late to current baseline increases over time.

Relatively Few but Large Baseline Changes



Of activities which persist through IMS submissions, roughly a quarter of them have changes to their Baseline Finish Date. **On average, those activities with baseline changes eventually slip over 2 months.**

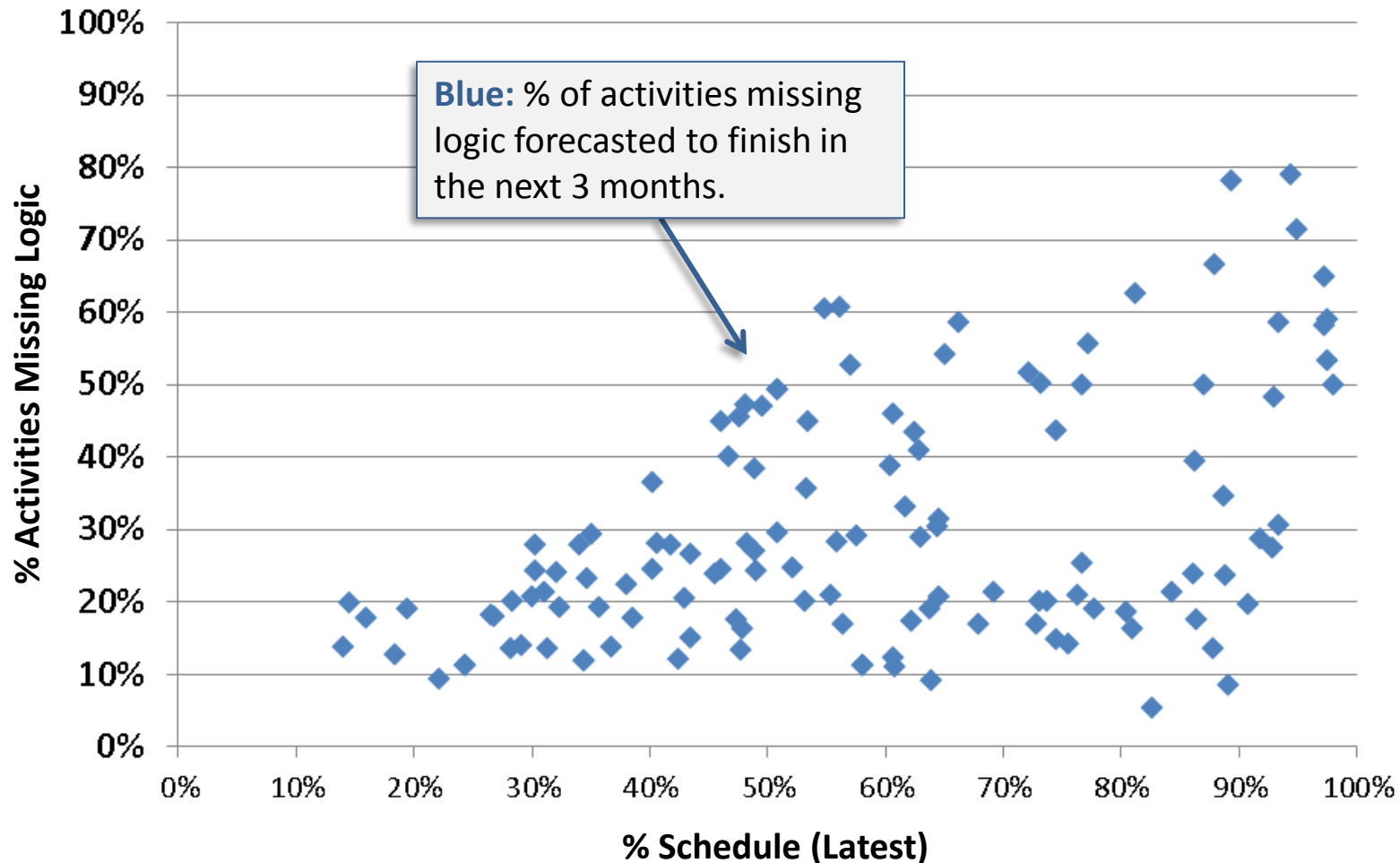
Numerous Binding Constraints



On average, schedules constrain 7.5% of total activities, but often range up to 20%. Of these constraints, generally 50% or more are binding.

➤ Binding: forecast start or finish date equals the constraint date

Schedules Become Less Logical Over Time



Schedule logic provides each activity with at least one predecessor and successor. It appears that instead of building logic into the schedule over time, the tasks become less inter-related.

IMS Observation Summary

- We have seen that, in general, schedule quality and performance tend to decrease over the course of a contract
 - Less Logic; More Late Tasks; Large Baseline Changes; Increasingly Unrealistic Forecasts
- **Question:** What affects schedule realism?
 - Project changes: re-plans, work-arounds, evolution, etc.
 - Participant incentives: schedulers, control account managers (CAMs), program managers (PMs)
 - Do schedules with high quality ratings at any given point evolve logically over time?
- Importantly, it is often late in the schedule when slips are realized in the IMS
- **Question:** Can current IMS data be better used to measure schedule risk and extrapolate a realistic end date?



Part 2: An Improved Schedule Estimator

A Simple Idea

- Set the first IMS available as a baseline
 - Preferably immediately after the Integrated Baseline Review (IBR)
- Use activity names and/or unique IDs to trace activity end dates through subsequent schedules and compare them to their original baseline end date and slack days (total float). Disregard all:
 - New work packages
 - Activities which change identifiers
 - New activity sequencing
- The activity which has slipped the most to its original baseline is used to extrapolate a project end date
- When the schedule has evolved to the extent that the current work packages no longer reflect the baseline, the predicted end date stabilizes and the estimate is determined
 - Original baseline irrelevance: when more than 95% of the activities from the original baseline are either finished or no longer appear in the current schedule
 - Analysis may start anew if schedule re-plan occurs before “original baseline irrelevance”

***This analysis is a good early predictor of schedule end for contracts **not including** Indefinite Delivery, Indefinite Quantity (IDIQ) The 4 IDIQ contracts collected are not shown in the following slides, future work needed for analysis at a lower (task-order) level.**

An Example

IMS #1

Activity Name	Baseline Finish	Total Float
Rough grade property	03-20-13	3
Frame first floor walls	03-25-13	3
Install roof trusses	03-27-13	3
Install roof decking	03-29-13	3
Rough-in framing inspection	04-01-13	5
Form and pour driveway	04-02-13	3
Finish excavate and pour garage	04-03-13	3
Framing complete	04-03-13	3

IMS #2

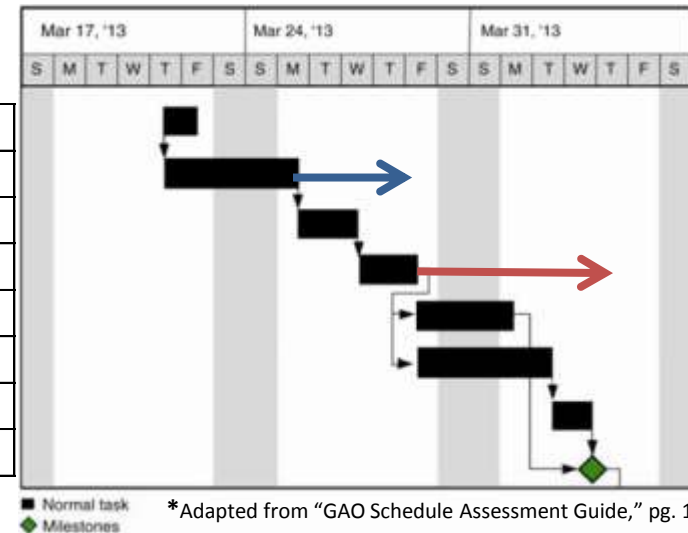
Finish	Slip
03-20-13	0
03-29-13	1
03-29-13	0
03-31-13	0
04-03-13	0
04-04-13	0
04-05-13	0
04-05-13	0

Max Slip: 1

IMS #3

Finish	Slip
03-20-13	0
03-29-13	1
03-29-13	0
04-05-13	4
04-03-13	0
04-05-13	0
04-06-13	0
04-06-13	0

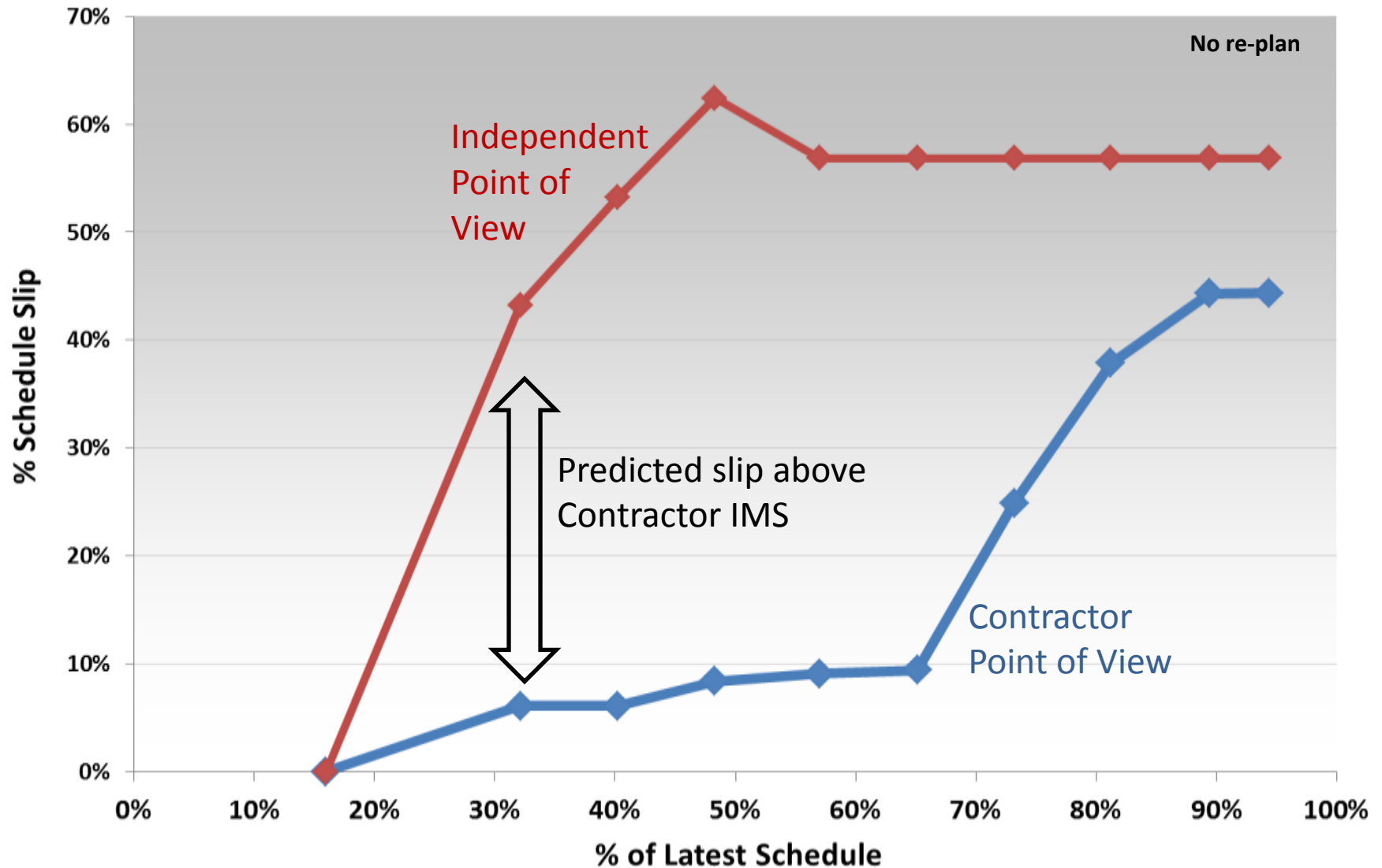
Max Slip: 4



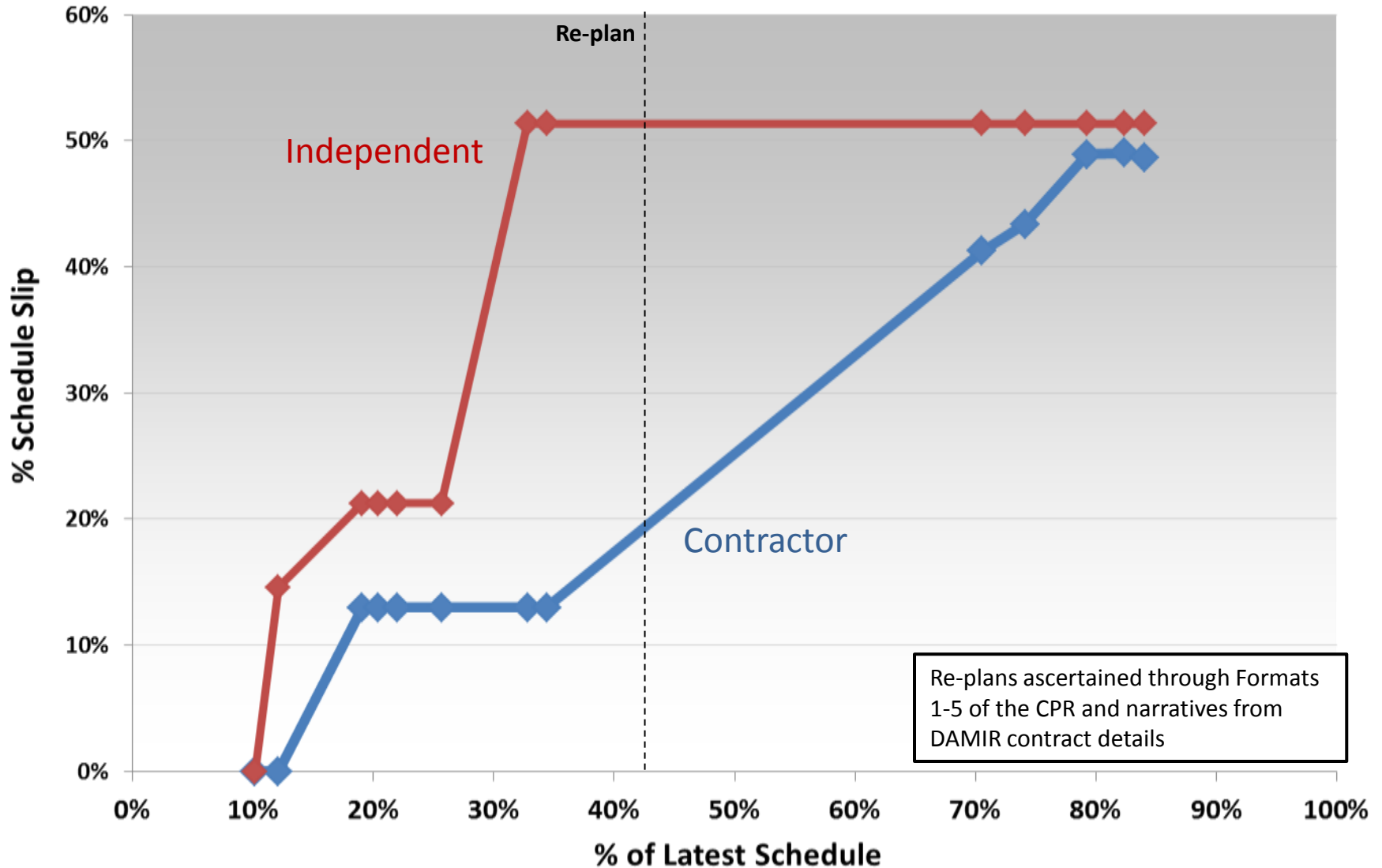
1. IMS #1 (first submission) sets the baseline for future activities
 - For each subsequent IMS, evaluate all updated finish dates and compare them to the baseline from IMS #1
 - ⇒ Slip Effect = $\max(\{f(x_i): i = 1, \dots, n\})$
 where $f(x_i) = [\text{Current Finish}_i - \text{Baseline Finish}_i - \text{Baseline Total Float}_i]$
 and $\max(\{f(x_i): i = 1, \dots, n\}) \geq 0$
2. In IMS #2, "Frame first floor walls" slipped 4 days to baseline. Factoring in 3 days of total float, the total effect on schedule is now 1 day of predicted slip.
3. In IMS #3, "Install roof decking" slipped 7 days to baseline. Factoring in 3 days of total float, the total effect on schedule is now 4 days of predicted slip.

As detail gets built into the schedule over time and a majority of baseline activities have their finish dates realized, the metric finds a stable value

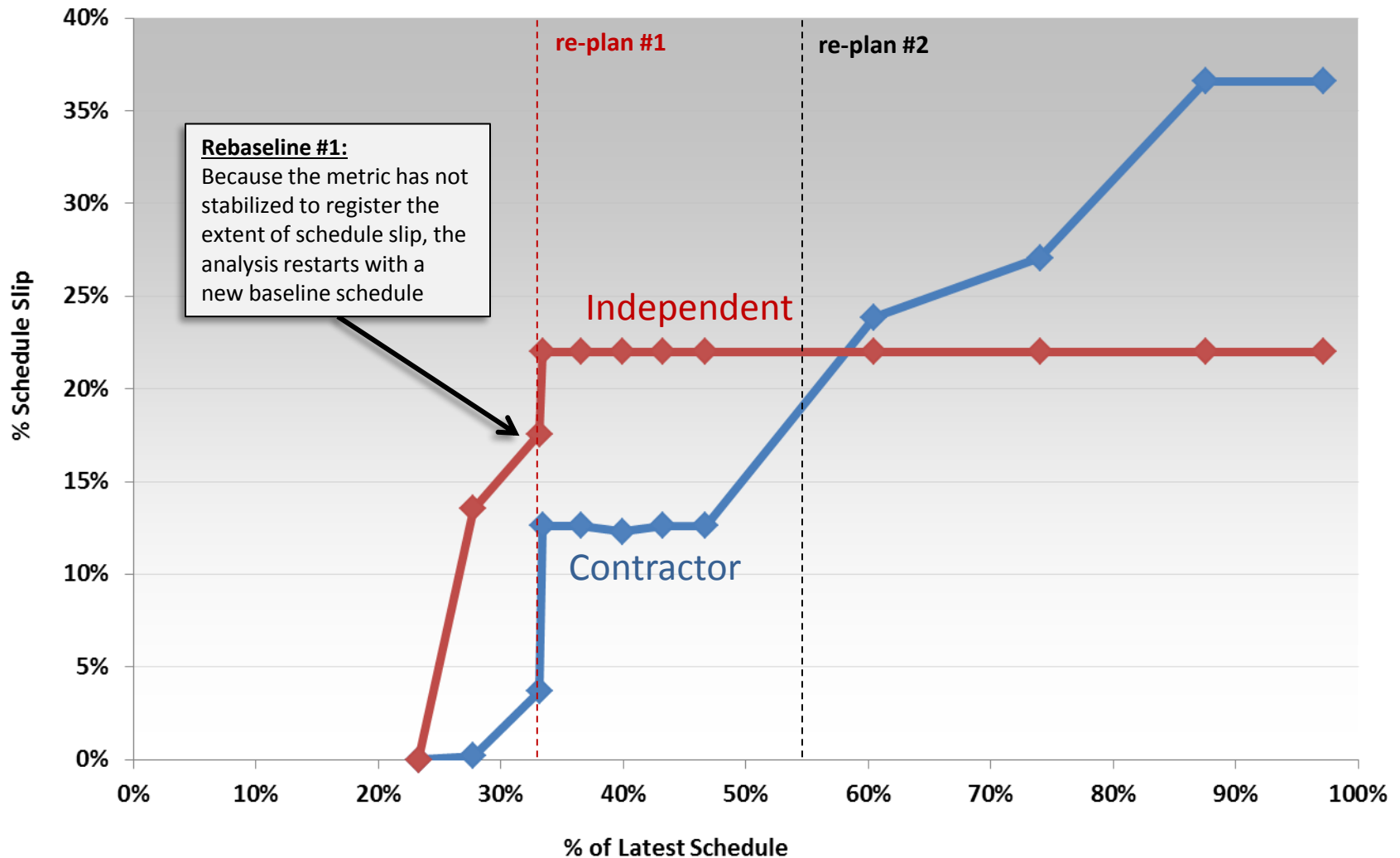
Results: Contract #1



Results: Contract #2

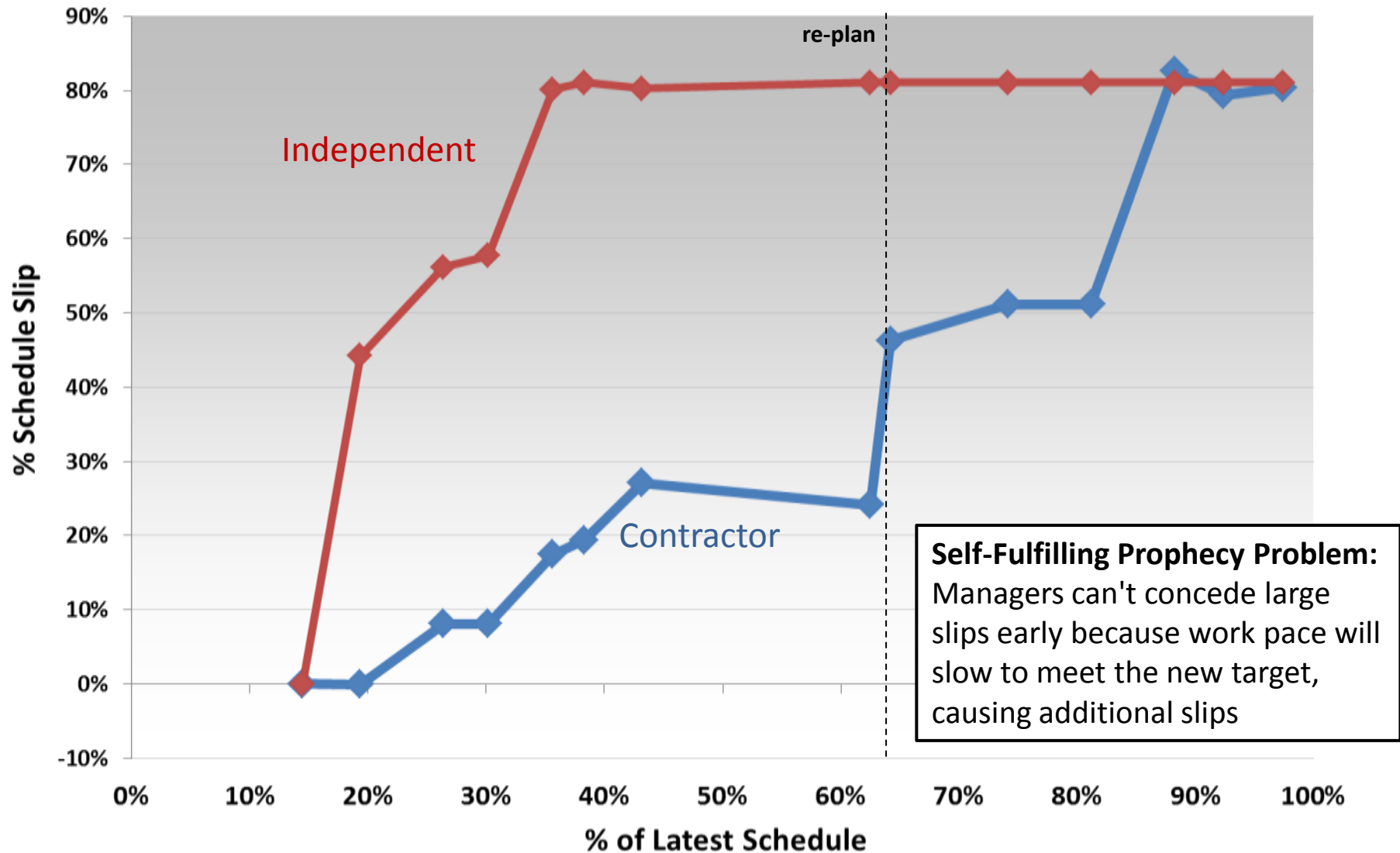


Results: Contract #3

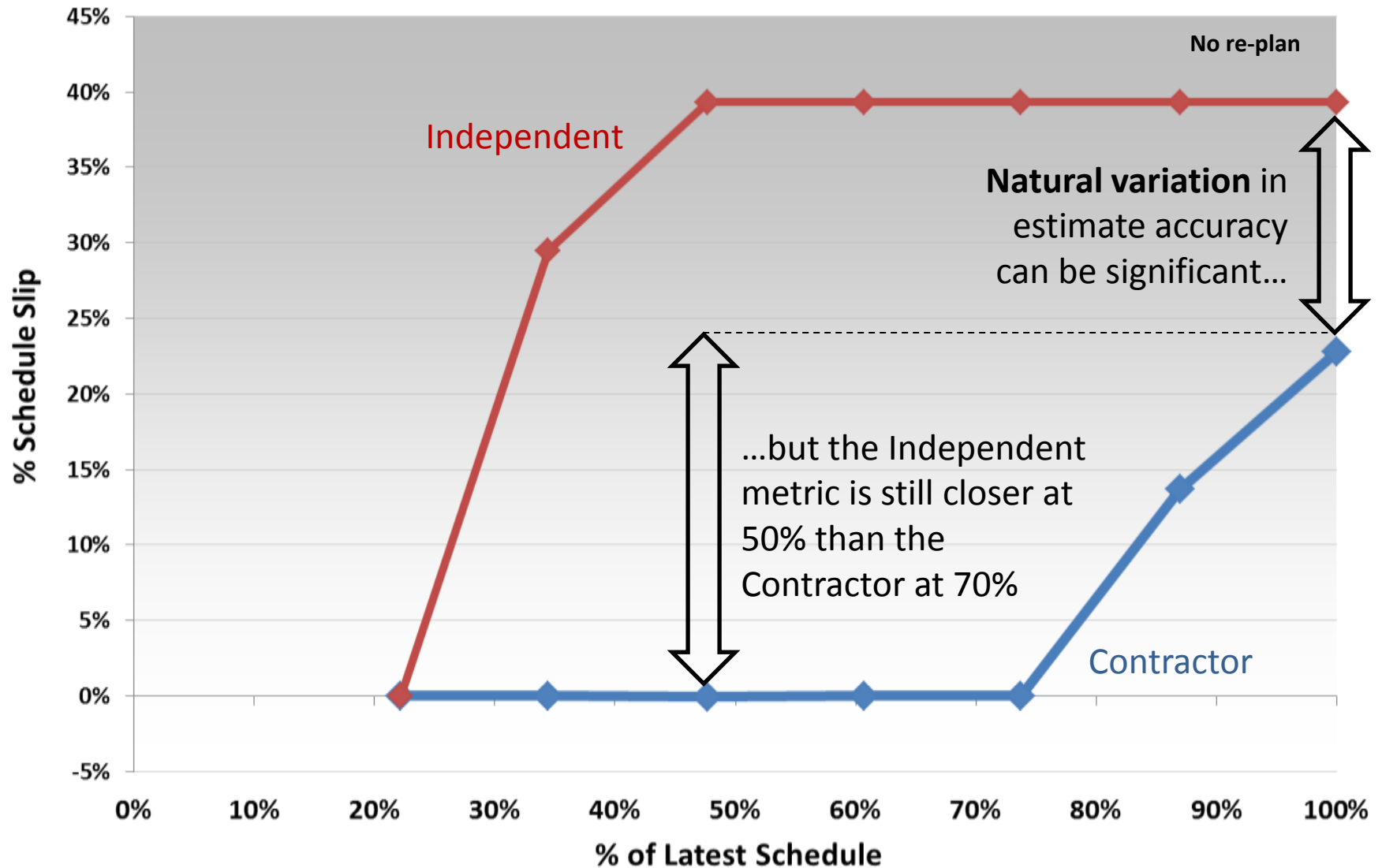


If the contract undergoes a major re-plan early on while many near-term tasks are still changing to baseline, then it is reasonable to use the re-planned schedule as a new baseline. However, the metric will have already accounted for future re-plans once it “flat-lines.”

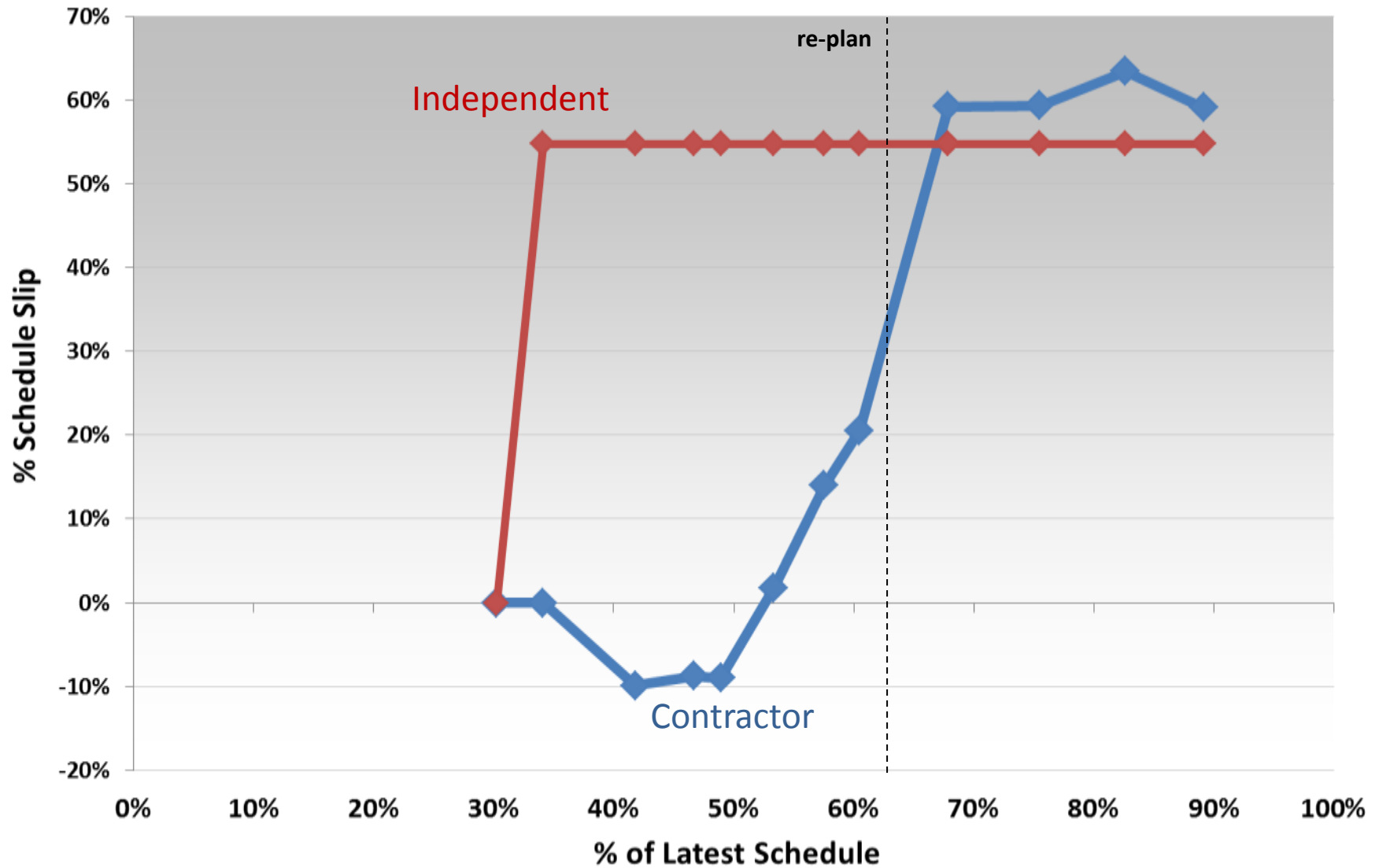
Results: Contract #4



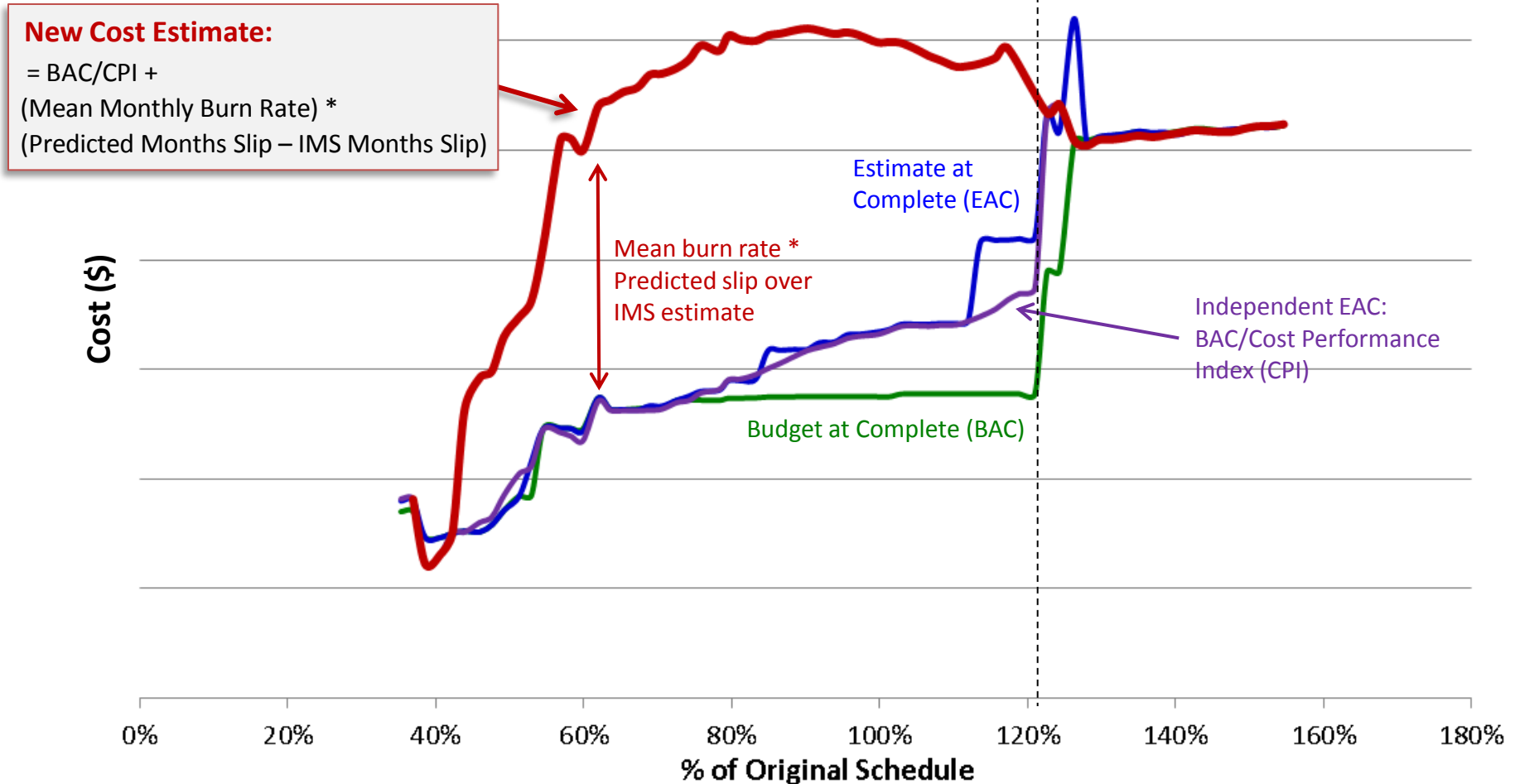
Results: Contract #5



Results: Contract #6

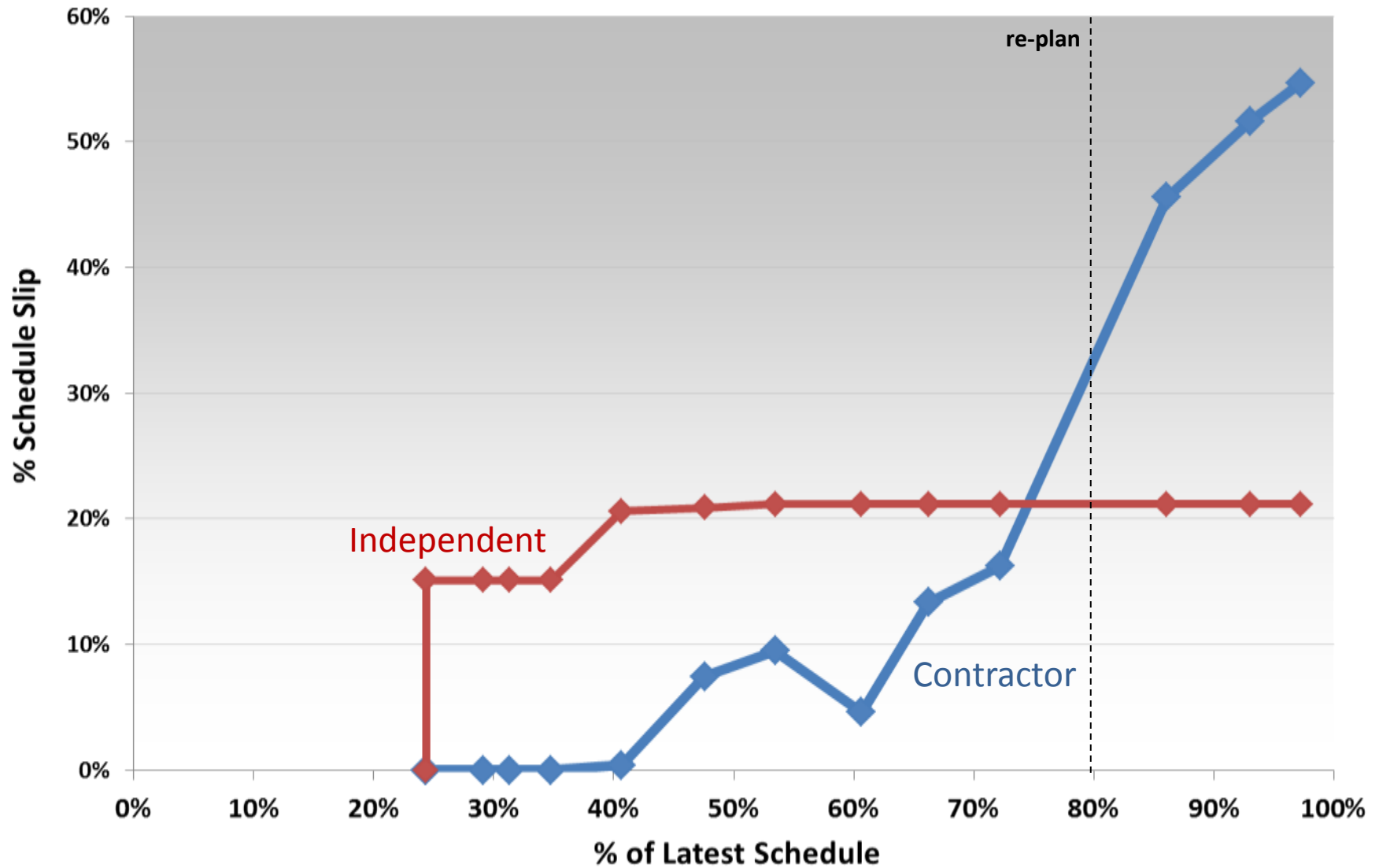


Potential Use in Cost Estimates

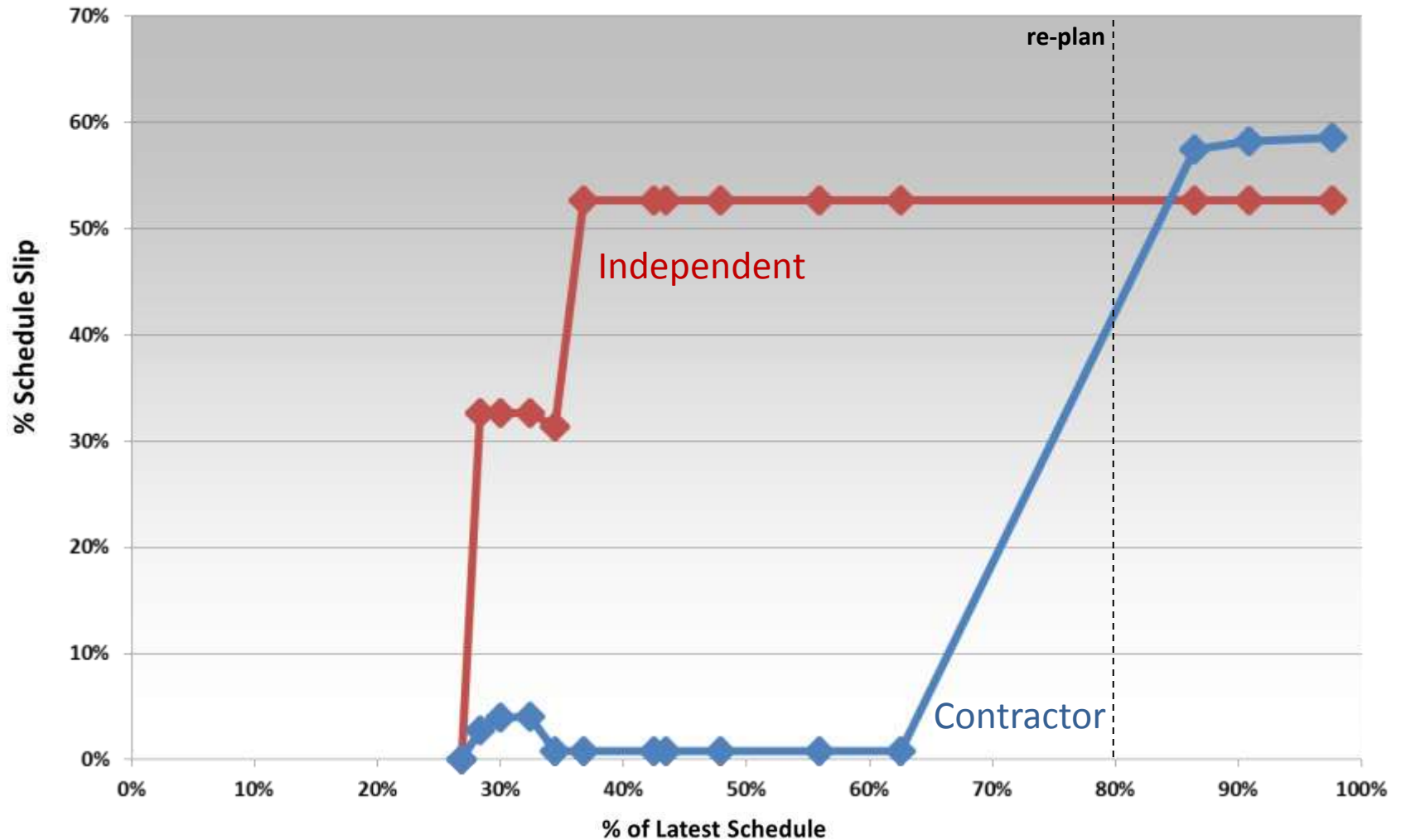


Contract cost Estimates at Complete (EACs) also suffer from lack of realism. By extending the cost burn rate for predicted schedule slip over and above that reflected in the IMS, one may be able to account for additional schedule risk.

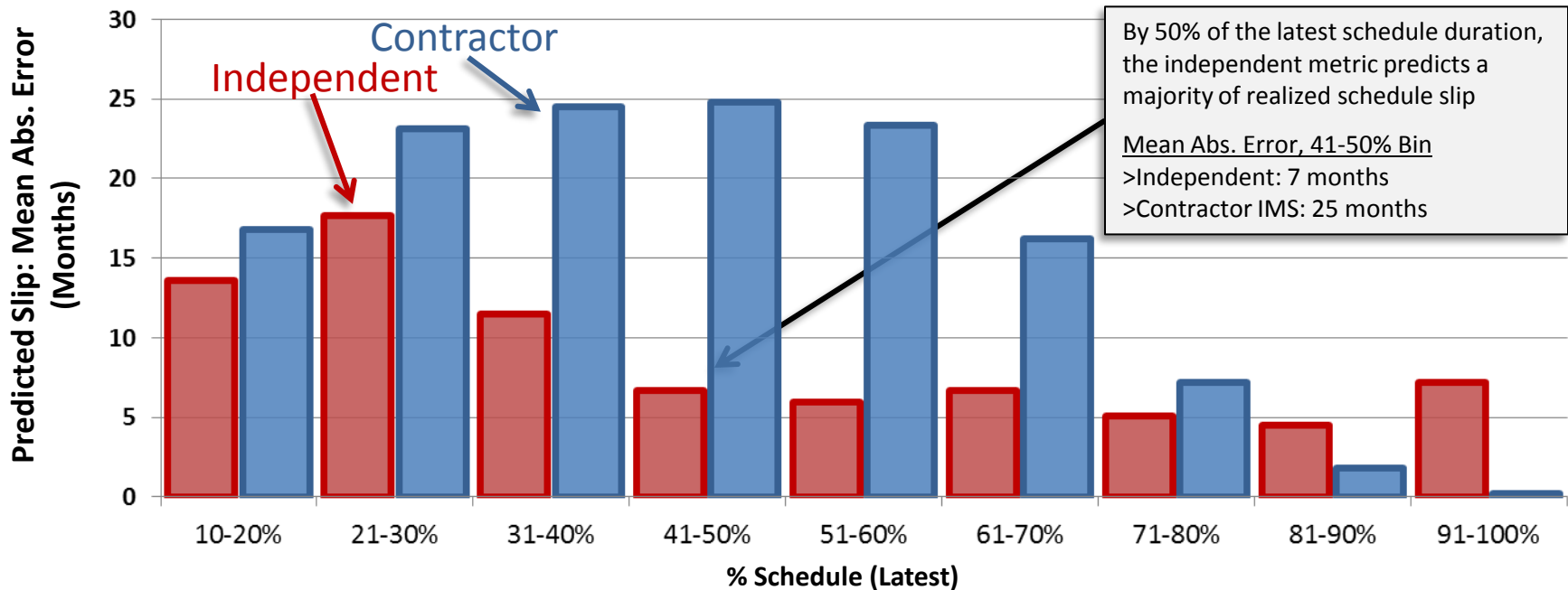
Results: Contract #7



Results: Contract #8



We Now Have an Early Indicator



How can this metric be used to improve outcomes on MDAP contracts?

1. Gives the decision-maker an early indicator to the magnitude of a potential schedule slip
 - Plan early for cost-schedule-technical tradeoffs
 - Fewer sunk costs leads to greater flexibility in project termination
 - May prove useful in providing cost estimates
2. Allows for traceability between submissions
 - High schedule quality as traditionally calculated may be misleading
 - Is the schedule continually re-invented? How volatile are activities?

Conclusions

➡ There is a better way of predicting schedule slip

- While schedules develop over time, planners do a good job at the outset of phasing major milestones and near-term work packages
 - Things will change, but not that much relative to other projects
- As a project progresses, cumulative changes to a schedule deteriorate its quality and weaken the signal of schedule slip
 - Schedule performance is best registered early
 - Performance metrics are leading rather than concurrent indicators
 - Better schedule maintenance is needed
- Performance to baseline early in the project is a good indicator of realized schedule slip
 - Contractors quickly reveal their pace of work and “settle” into performance
 - Schedule often cannot be made up through work-arounds, forced constraints, or optimistic forecasts

Moving Forward...

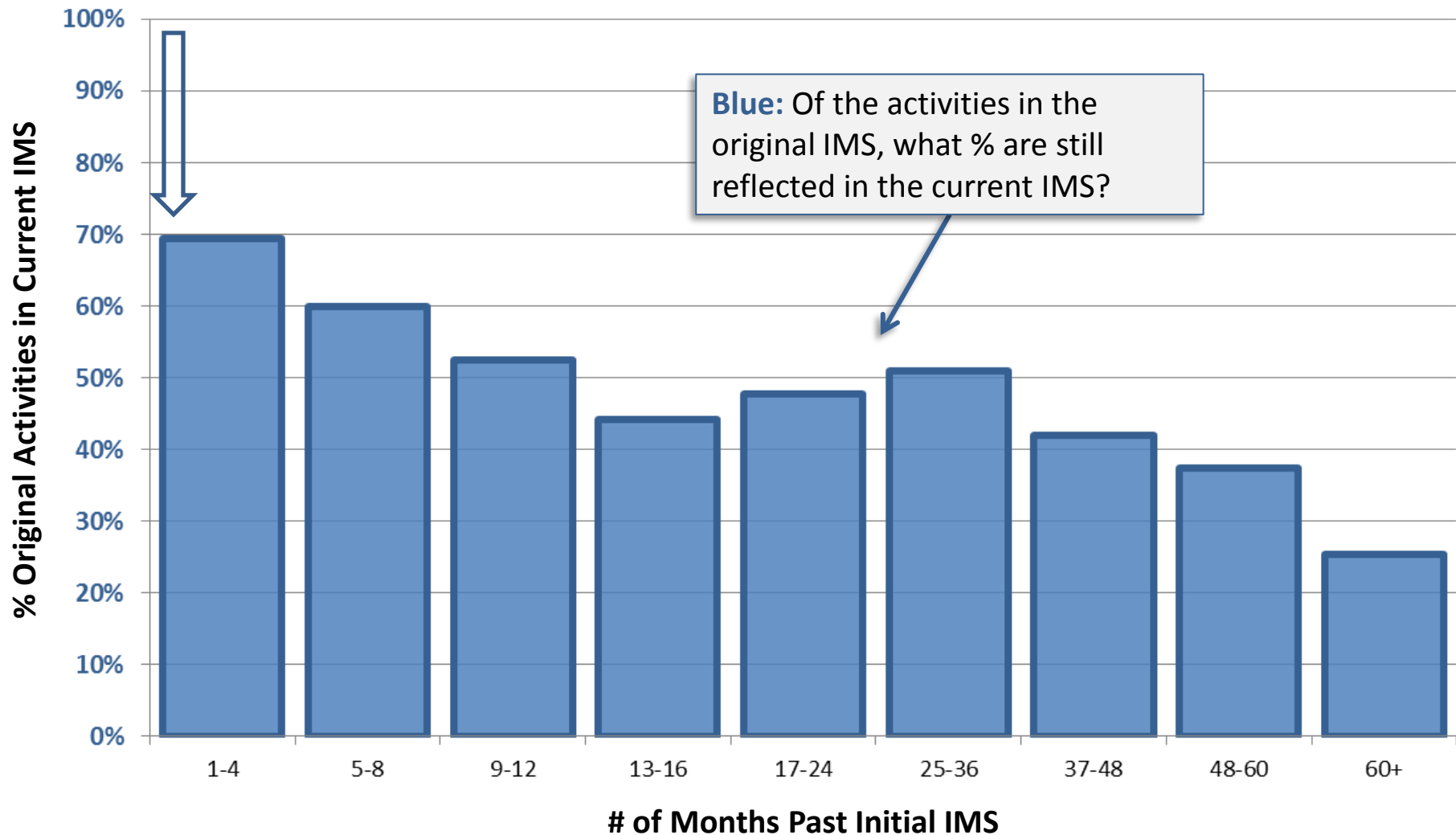
1. Expand data set to test metric robustness
2. Assess metric value for IDIQ contracts
 - Need activities broken out by task-order
 - More detailed approach
3. Consider applications to augment cost estimation
 - Extending average burn rate
4. Provide data-driven generalizations of schedule quality/realism over course of a contract
 - Why do activities lose logic links over time?
 - How much do binding constraints affect schedule realism?
 - What is the effect of activity churn between submissions?

Bibliography

- “GAO Schedule Assessment Guide.” U.S. Government Accountability Office (GAO). 30 May 2012. <http://www.gao.gov/assets/600/591240.pdf>. Accessed 10 April 2014.
- “Joint Cost and Schedule Risk and Uncertainty Handbook (CSRUH).” Naval Center for Cost Analysis (NCCA). 16 July 2013. <https://www.ncca.navy.mil/tools/csruh/index.cfm%20>. Accessed 10 April 2014.
- “Earned Value Management System (EVMS) Program Analysis Pamphlet (PAP).” Defense Contract Management Agency (DCMA). July 2012. <http://www.dcmamilitary.com/policy/200-1/PAM-200-1.pdf>. Accessed 10 April 2014.
- “IMS DID.” Office of the Undersecretary for Defense Acquisition, Technology, & Logistics Performance Assessments and Root Cause Analyses (OUSD AT&L PARCA). 30 March 2005. <http://dcarc.cape.osd.mil/EVM/Documents.aspx>. Accessed 10 April 2014.
- “IPMR Implementation Guide.” Office of the Undersecretary for Defense Acquisition, Technology, & Logistics Performance Assessments and Root Cause Analyses (OUSD AT&L PARCA). July 21 2012. <http://dcarc.cape.osd.mil/EVM/Documents.aspx>. Accessed 10 April 2014.
- “Ending the EAC Tail Chase: An Unbiased EAC Predictor using Progress Metrics, Eric R. Druker, et al., SCEA/ISPA, 2007.

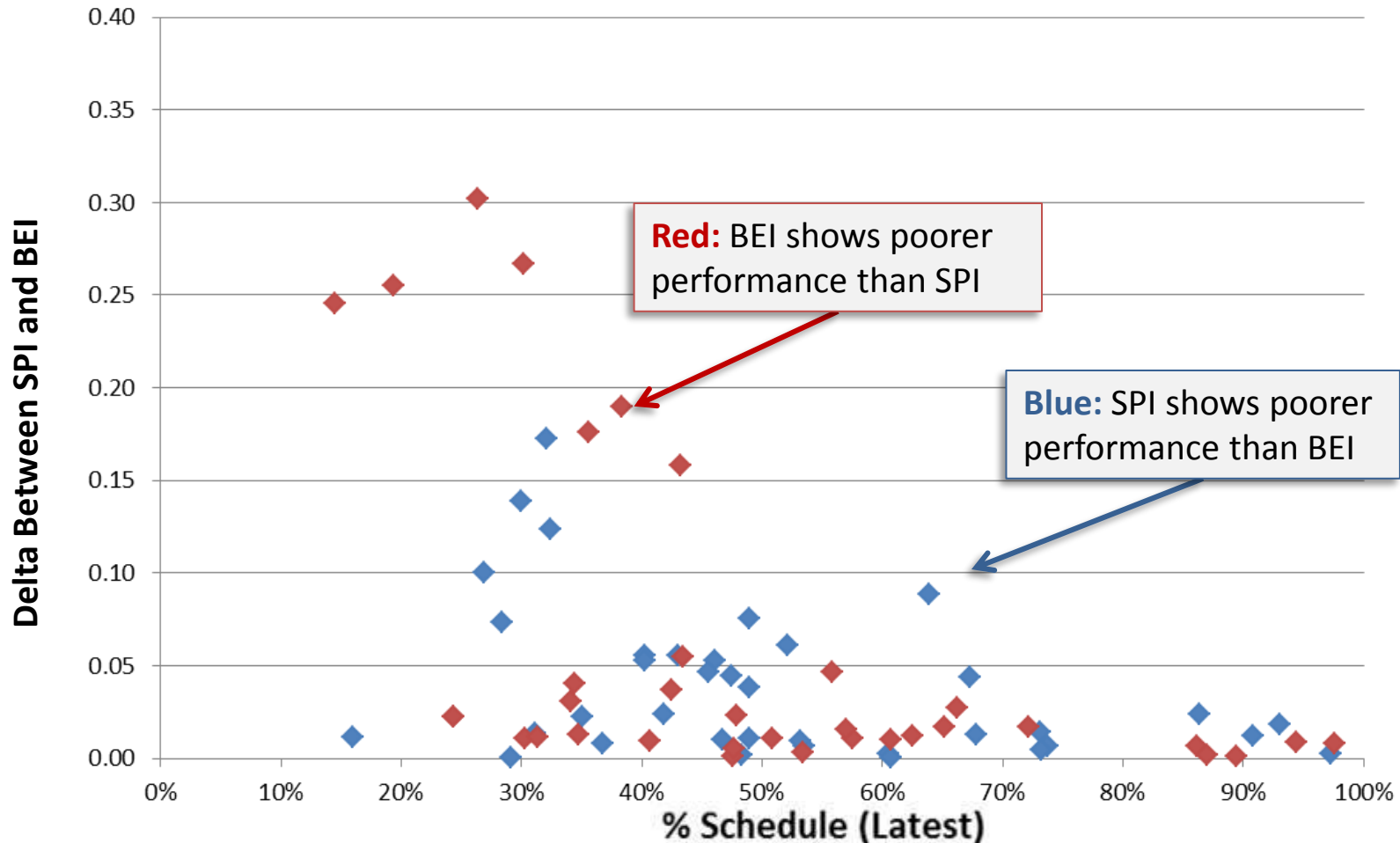
Back-Up

Significant Churn in Schedule Activities



Roughly 30% of the activities in the original IMS are quickly untraceable or removed from the schedule. By contract end, the IMS often retains a history of less than 50% of activities from the original IMS.

BEI and SPI Tell Different Stories



The Schedule Performance Index (SPI) relates dollars earned to dollars planned, rather than tasks completed to tasks planned for the BEI. Though often compared, they vary widely early in a contract.