A Novel Approach to Early Phase Agile Software Estimating and Sizing

Wilson Rosa
Sara Jardine

2023 ICEAA Professional Development & Training Workshop
May 16-18, 2023

Agenda

Introduction

- Proposed Solution
- Breakthroughs
- Model Usefulness
- Requirements Comparison

Data Analytics

- Data Collection
- Variables
- Data Normalization
- Demographics
- Descriptive Statistics

Results

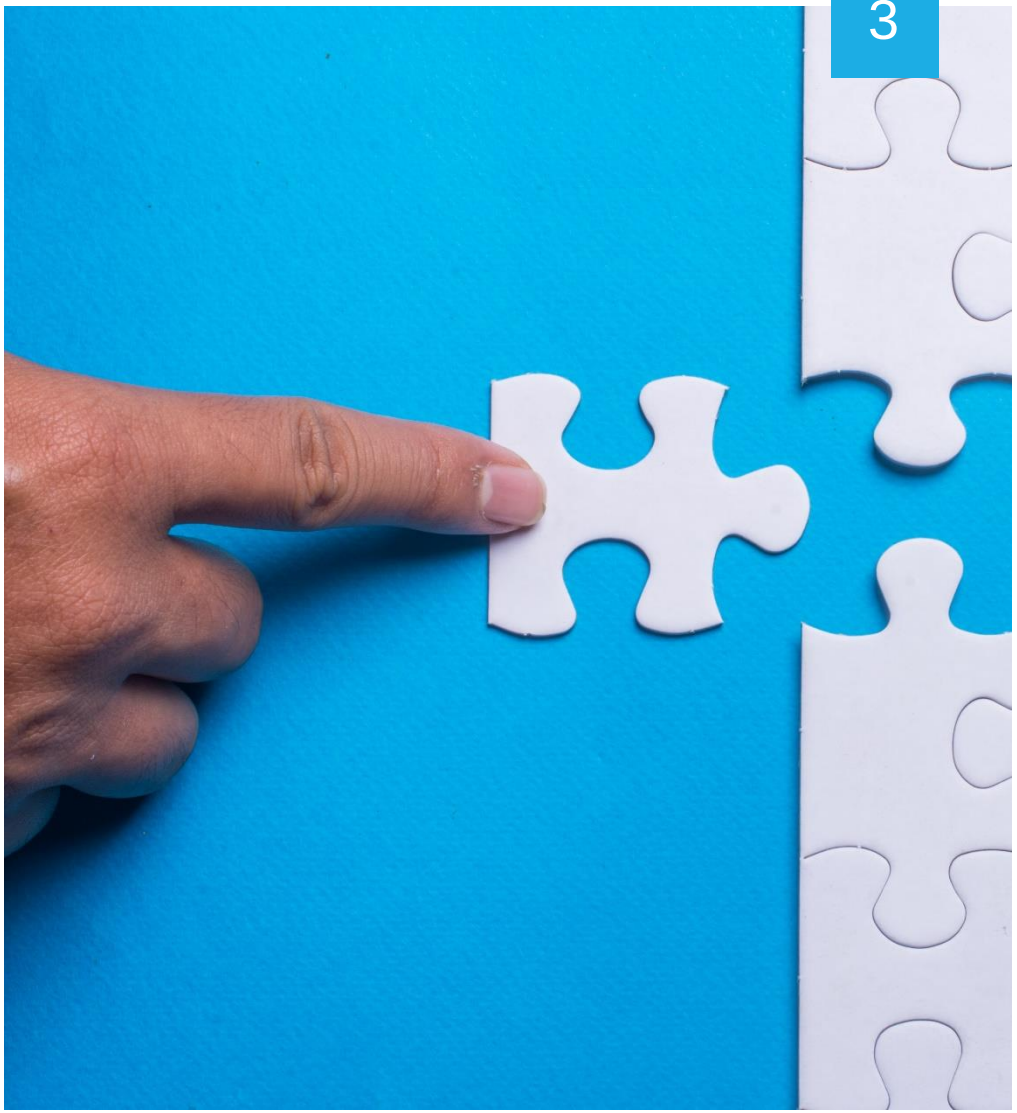
- Benchmarks
- Estimation Models

Conclusion

- Limitations
- Main Takeaways

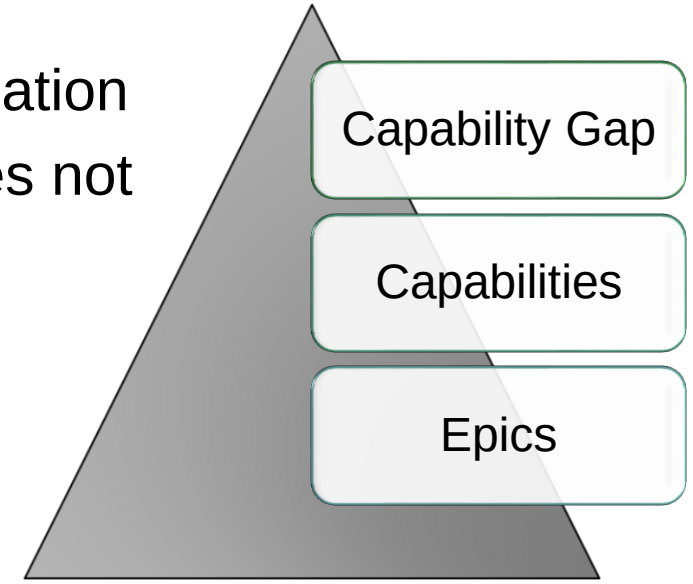
What is the Solution?

- Build effort and schedule estimation models based on high-level sizing measures available at an earlier phase
- Appropriate for:
  - ✓ Initial Budget Estimates
  - ✓ Analysis of Alternatives
  - ✓ Rough Order of Magnitude Cost Estimates



Study Breakthroughs

- Publicize early phase agile cost estimation models using high-level size measures not previously examined elsewhere

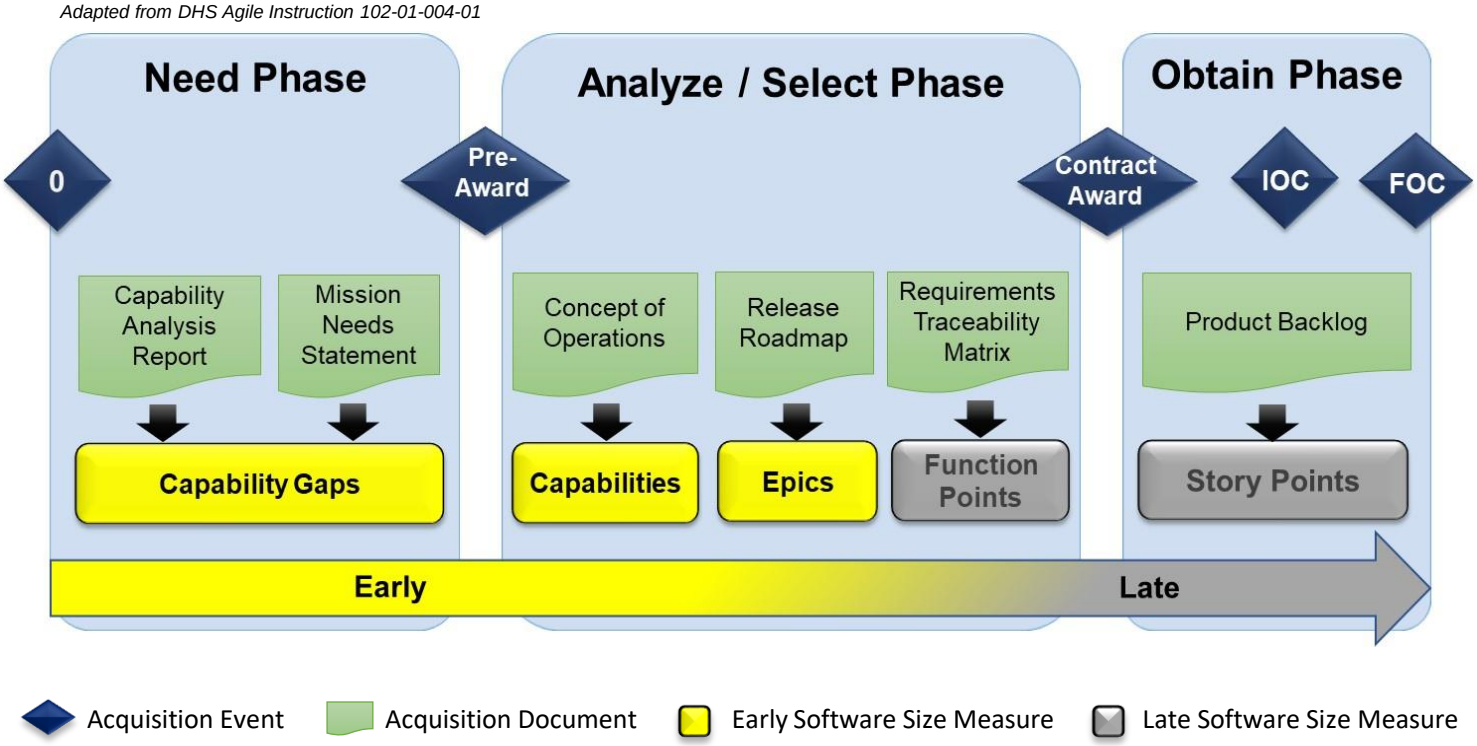


- A **Capability Gap** is a high-level requirement that is needed to perform a mission, but the Department does not currently possess it and there are no plans for it to be provided by existing programs
- One or more unique **Capabilities** are needed to close or resolve each **Capability Gap**

When To Use Our Early Phase Models?

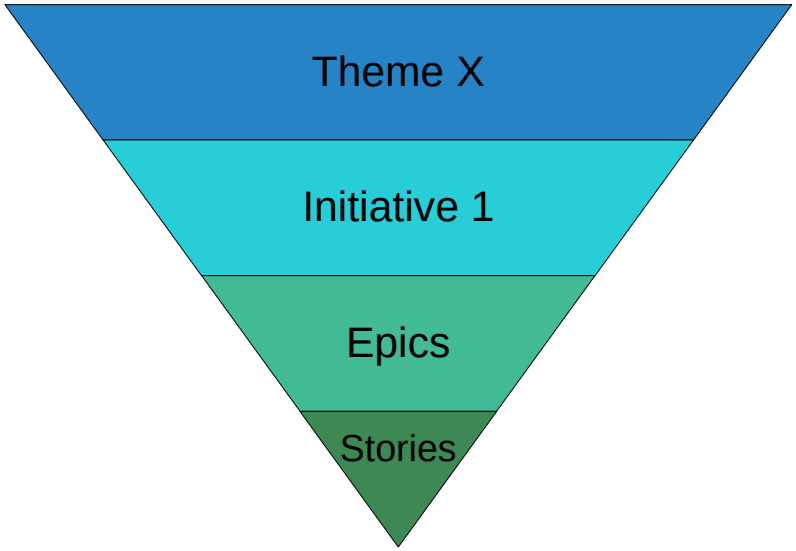
- When choosing the appropriate estimation model, analysts should consider program's lifecycle **phase** and which early size measures are available at that time

Adapted from DHS Agile Instruction 102-01-004-01

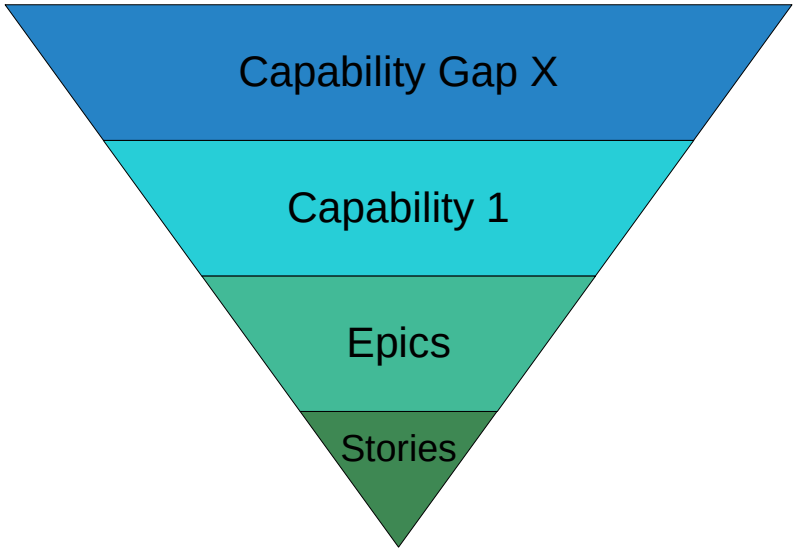


Agile Software Requirements Industry vs DHS

Mainstream Agile



DHS

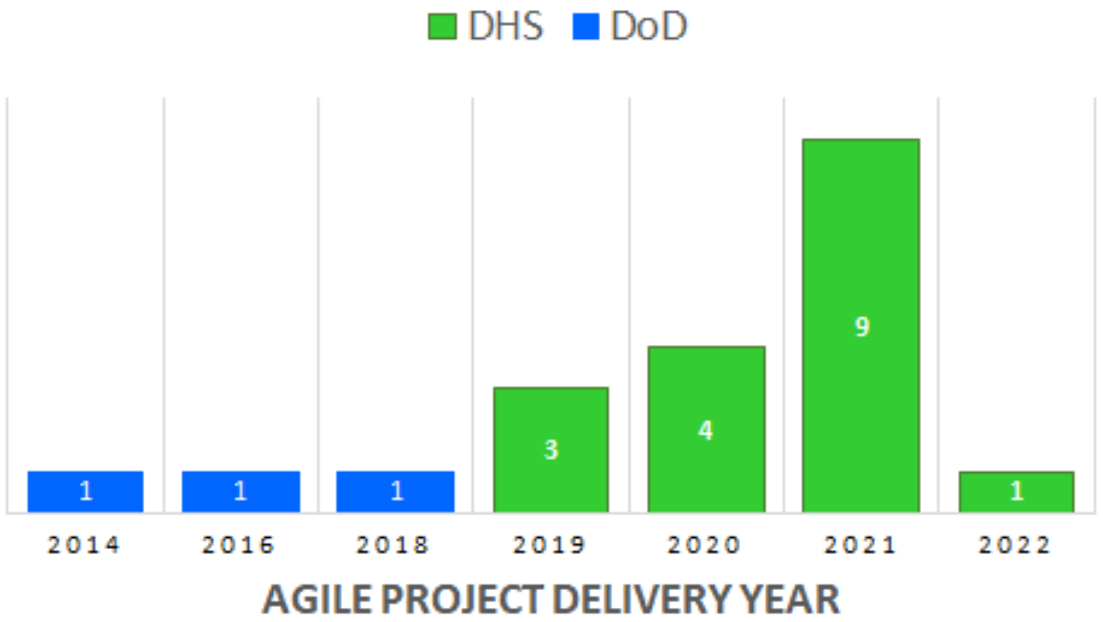


At DHS, a Theme aligns with Capability Gap whereas an Initiative aligns with Capability

Data Analytics

Data Collection

Dataset includes 20 agile projects across 14 different companies



Same Team collected, normalized, and validated the data

Data Sources

Who: Provided by Program Management Offices

Where: Obtained from Authoritative Documents:

Effort and Schedule	Size	Project Context
Monthly Contractor Invoice	Mission Needs Statement	Acquisition Documents
Product Backlog	Concept of Operations	Program Reviews
	Release Roadmap	

Variable Selection

Dependent Variables

Effort

- Total contract hours incurred by the agile development teams
- Reported at the release level

Schedule

- Total months to complete all software development activities
- Reported at the release level

Independent Variables

Capability Gap

- A high-level requirement that is needed to perform a mission but DHS does not posses it
- Also known as Theme

Capabilities

- The means to close or resolve capability gap
- Also known as Initiative

Epics

- Large bodies of work that can be broken down into a number of smaller tasks (called stories)

Data Normalization
How Did We Measure Effort?

- Effort captures total labor hours incurred by the agile development contractor(s)
- Labor includes 11 elements aligned to the DHS IT Work Breakdown Structure (WBS)

ID	DHS IT WBS Element
1.i	System Development
1.i.1	Program Management
1.i.2	Systems Engineering
1.i.4.2	Software Development
1.i.4.3	Data Development & Transition
1.i.4.5	Training Development
1.i.4.6.1	Development Test & Evaluation
1.i.4.6.1	Cybersecurity Test & Evaluation
1.i.4.7	Logistics Support Development
1.i.7	System Level Integration & Test
1.i.8.6.1	Help Desk/Service Desk (Tier 3)
1.i.8.6.4	Software Maintenance

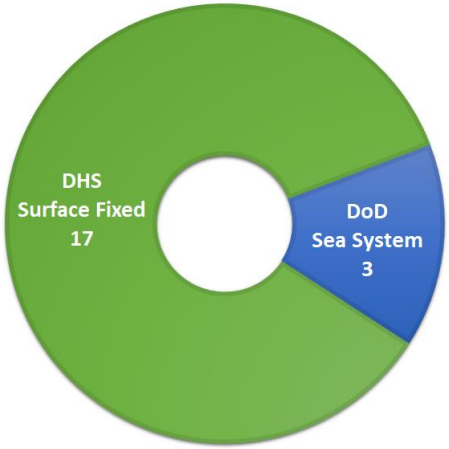
Why use total labor?

Reporting labor at the **total level** is recommended since most DHS agile development contracts are FFP or T&M, and generally do not breakout effort by major cost elements as seen in traditional cost-plus contracts

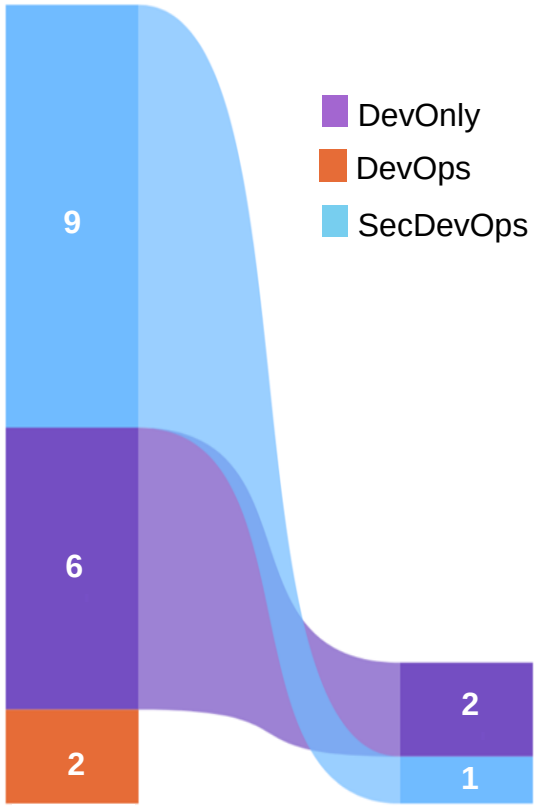
FFP = Firm Fixed Price | T&M = Time and Materials

Dataset Demographics

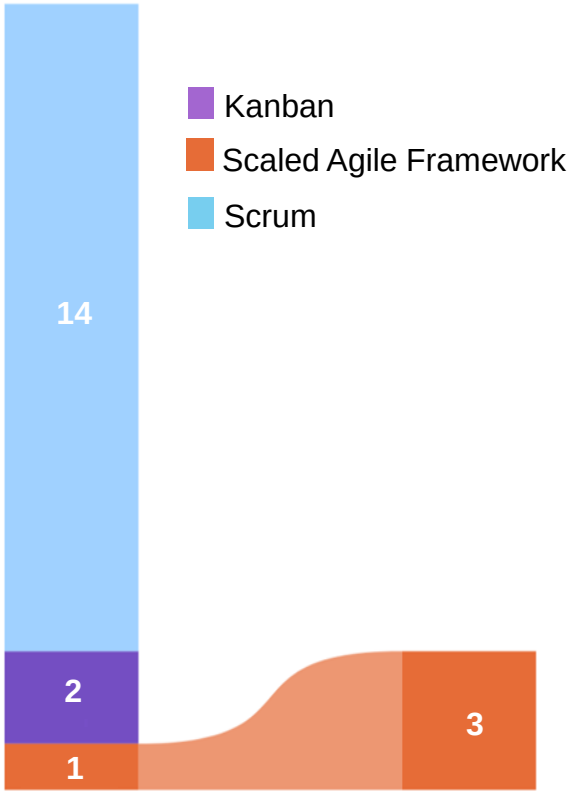
- Sample Size: 20 Projects
- Automated Information Systems
- Majority (16) used AWS as Cloud Provider
- 2 to 4-week Iterations



Dataset Demographics
Agile Process

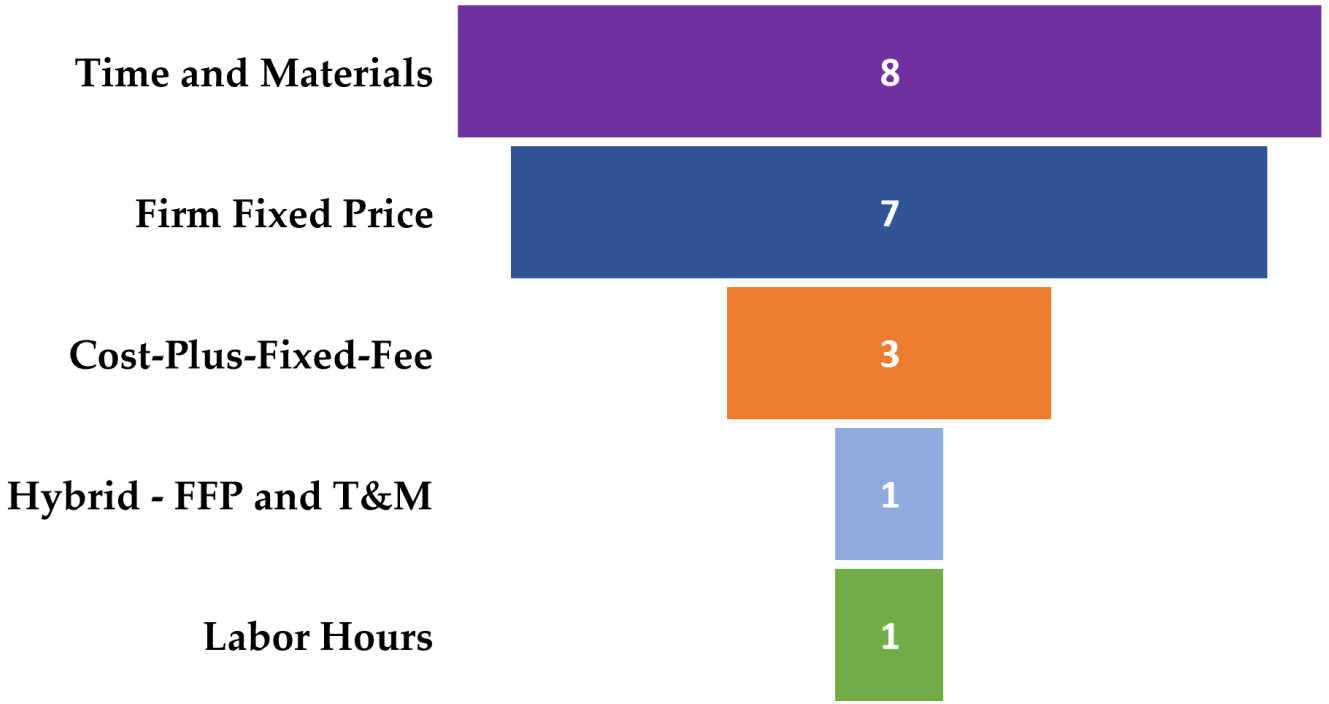


Team Approach



Framework

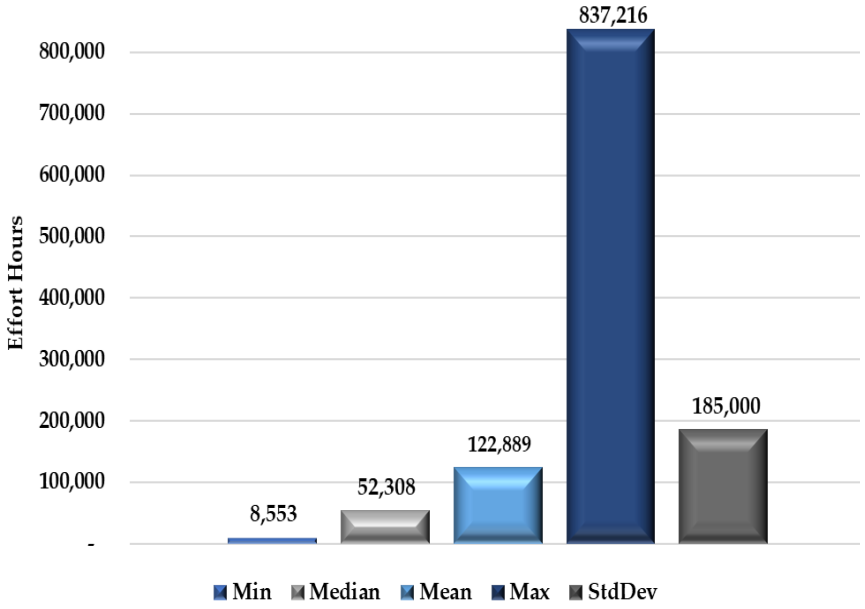
Dataset Demographics
Contract Strategy



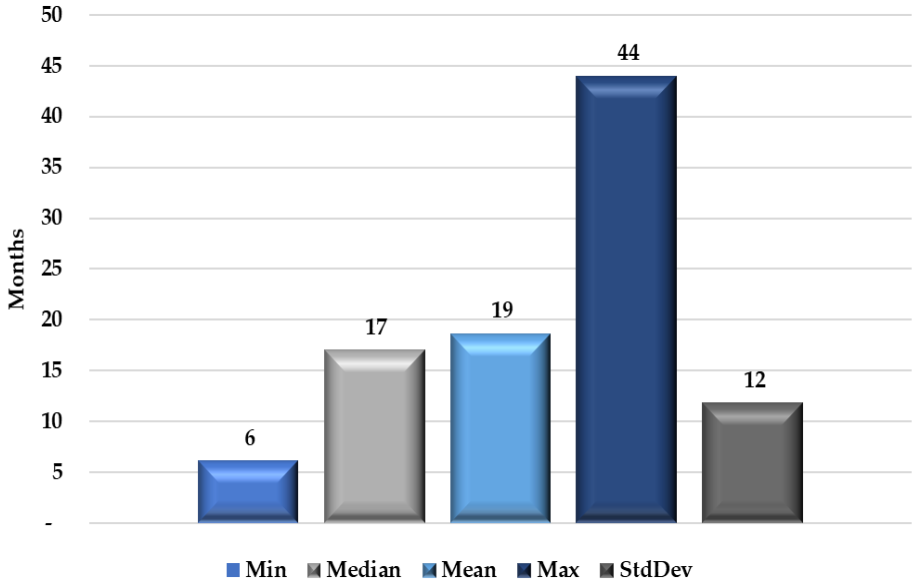
Majority used Firm Fixed price (FFP) or Time and Materials (T&M) contracts

Descriptive Statistics Dependent Variables

Effort



Schedule



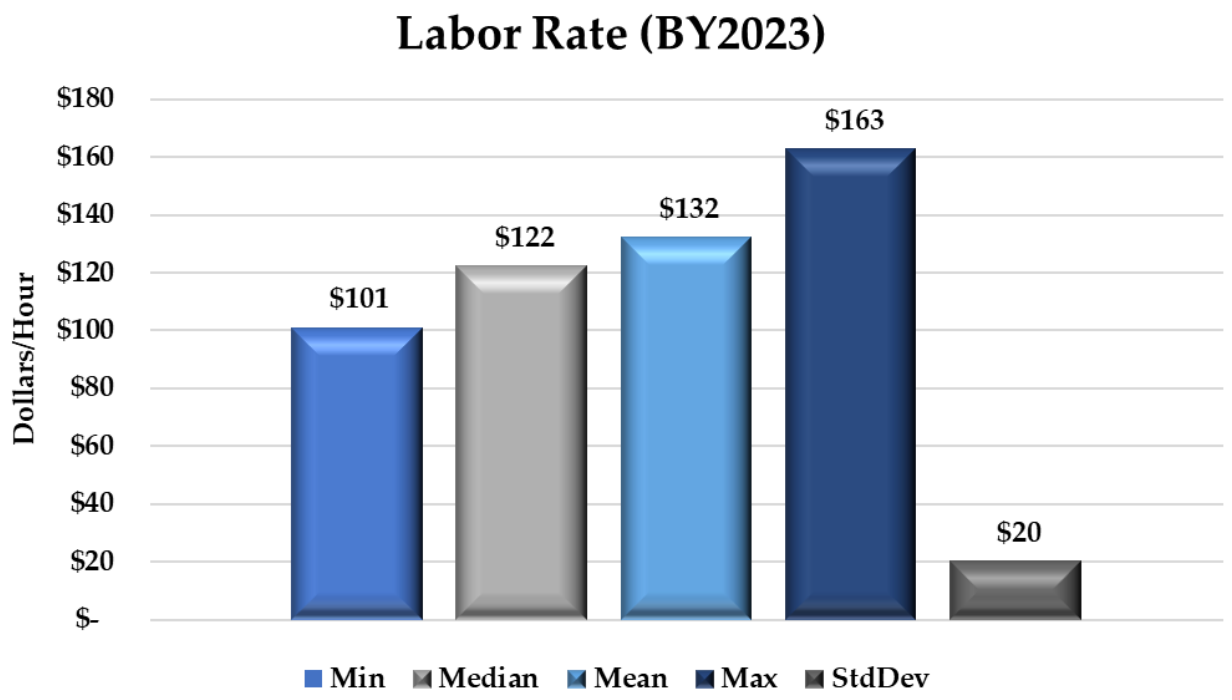
The averages for the dataset are 122,889 total hours and 19 total months

Descriptive Statistics
Size Measures

Size Measure	Min	Median	Max	StdDev
Capability Gap	1	5	26	6
Capability	4	10	50	11
Epic	13	35	406	89

When selecting a regression model, consider the relevant range of each independent variable

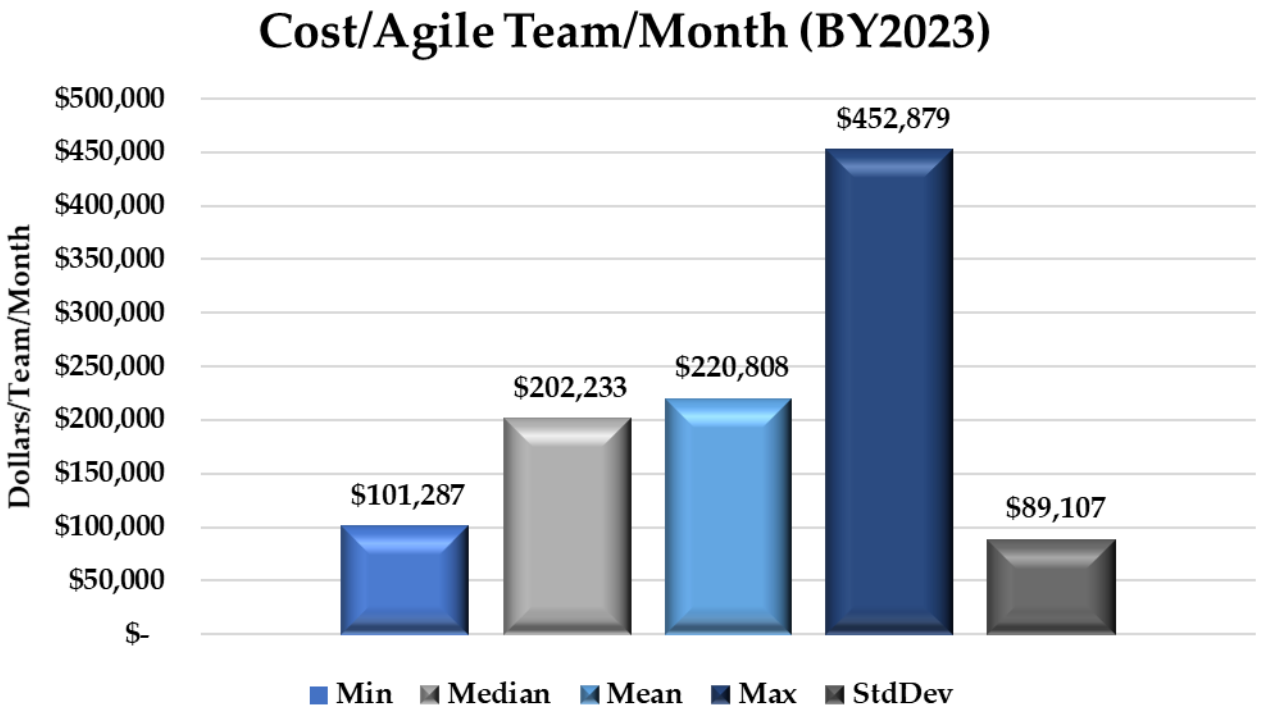
Descriptive Statistics
Labor Rate



Labor Rate (Fully Burdened) = $\frac{Total\ Contract\ Price}{Total\ Contract\ Hours}$

Appropriate for Crosschecking Bidder's Labor Rates

Descriptive Statistics
Cost per Agile Team



Useful for Independent Government Cost Estimates or Cost Realism Analysis

Model Results

Effort Productivity Benchmarks

Benchmark	1st Quartile	Median	3rd Quartile	StdDev
Hours/CAP_GAP	22,210	26,310	30,443	14,890
Hours/CAP	4,490	5,696	9,112	3,925
Hours/Epic	1,180	1,789	2,048	792

CAP = Capability | CAP_GAP = Capability Gap

Practical Application:

- For example, in practice, analysts can predict effort by taking the size (e.g., Epics = 100) from a Release Roadmap, then multiplying by the appropriate effort benchmark (median value from lookup table above):

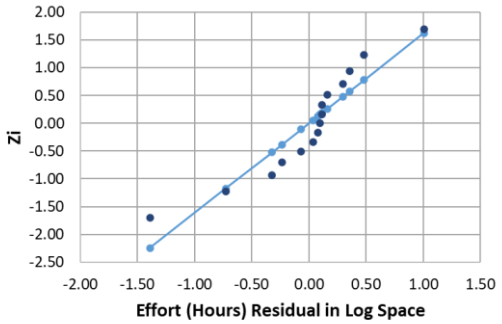
$$Effort = Size\ x(\text{Benchmark}) = 100\ x\ (1{,}789) = 178{,}900\ \text{hours}$$

Effort Estimation Model 1

Model	Equation	N	SE	R^2	R^2_{adj}	R^2_{pred}	MMRE
1	$\mathbf{E} = 26368\ x\ \mathrm{CAP\_GAP}^{0.9518}$	15	0.57	75.0%	73.1%	66.1%	43.8%

Where:

E	=	Total final contract **E**ffort (in Hours)
CAP_GAP	=	# of Capability Gaps from MNS*



- ✓ Appropriate for ROM** cost estimates at the Need Phase (in DHS) or Planning Phase (in DoD)
- ✓ Capability Gaps in DoD can be counted from the Capability Needs Statement

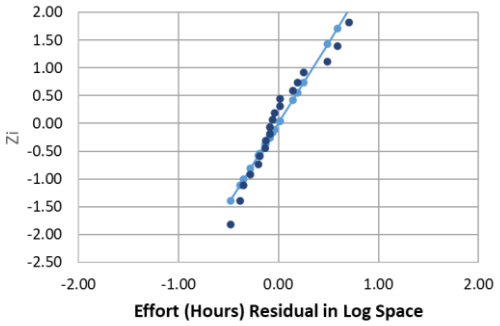
*MNS = Mission Needs Statement **ROM = Rough Order of Magnitude

Effort Estimation Model 2

Model	Equation	N	SE	R^2	R^2_{adj}	R^2_{pred}	MMRE
2	$\mathbf{E} = 1170\ x\ \mathrm{CAP}^{1.712}$	20	0.33	92.5%	92.1%	91.2%	23.5%

Where:

E	=	Total final contract **E**ffort (in Hours)
CAP	=	# of Capabilities from the CONOPS*



- ✓ Useful for building the baseline cost estimate at the Analyze/Select Phase
- ✓ Capabilities in DoD can be counted from the Software Initial Capabilities Document

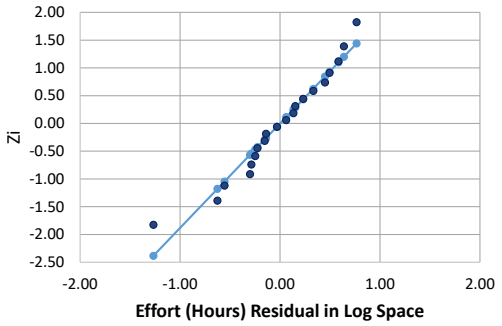
*CONOPS = Concept of Operations

Effort Estimation Model 3

Model	Equation	N	SE	R^2	R^2_{adj}	R^2_{pred}	MMRE
3	$\mathbf{E} = 710.6\ x\ \mathrm{EPIC}^{1.215}$	20	0.51	81.8%	80.8%	78.2%	42.5%

Where:

E	=	Total final contract **E**ffort (in Hours)
EPIC	=	# of Epics from the Release Roadmap



- ✓ Appropriate for IGCE* or assessing contractor's performance during the Obtain Phase
- ✓ Epics in DoD can be measured from the Product Roadmap or Product Backlog

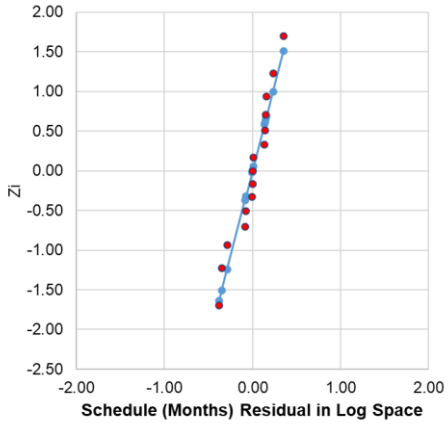
*IGCE = Independent Government Cost Estimate

Schedule Estimation
Model 1

Model	Equation	N	SE	R^2	R^2_{adj}	R^2_{pred}	MMRE
1	$S = 12.13 \; x \; \text{CAP\_GAP}^{0.4272}$	15	0.22	80.0%	78.5%	68.8%	16.4%

Where:

S	=	Total final development **S**chedule (in months)
CAP_GAP	=	# of Capability Gaps from MNS*



✓ Useful for crosschecking initial program schedules at the Need Phase

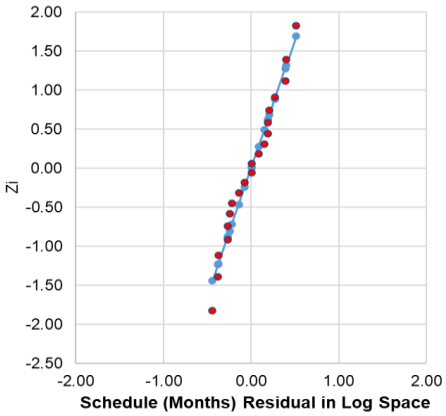
*MNS = Mission Needs Statement

Schedule Estimation
Model 2

Model	Equation	N	SE	R^2	R^2_{adj}	R^2_{pred}	MMRE
2	$S = 2.45 \; x \; \text{CAP}^{0.507} x \; 2.4^{D1}$	20	0.30	82.4%	80.4%	76.0%	24.5%

Where:

S	=	Total final development **S**chedule (in months)
CAP	=	# of Capabilities from the CONOPS*
D1	=	1 if Full Development or 0 if Enhancement



✓ Appropriate for evaluating baseline schedules at the Analyze/Select Phase

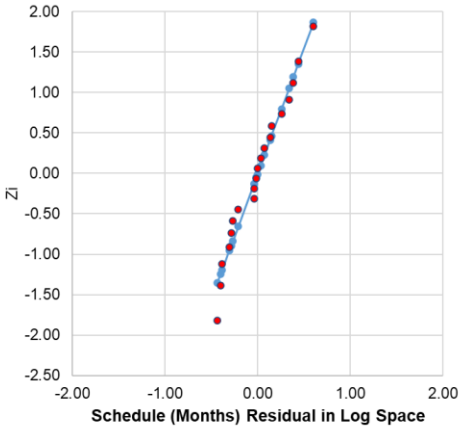
*CONOPS = Concept of Operations D1 = Dummy Variable

Schedule Estimation
Model 3

Model	Equation	N	SE	R^2	R^2_{adj}	R^2_{pred}	MMRE
3	$S = 2.069\ x\ \text{EPIC}^{0.3634} x\ 2.4^{D1}$	20	0.32	80.5%	78.2%	72.8%	24.3%

Where:

S	=	Total final development **S**chedule (in months)
EPIC	=	# of Epics from the Release Roadmap
D1	=	1 if Full Development or 0 if Enhancement



✓ Useful for assessing schedule in the Obtain Phase (in DHS) or Implementation Phase (in DoD)

Conclusion

Model Limitations

Internal Threats

- Dataset timeframe raises potential issues as 3 earlier projects may have used agile processes tailored to fit the agency's need

External Threats

- To produce a quality count of Capability Gaps from a MNS or Capabilities from an ORD or CONOPS, it is important that the program documentation be in a mature or final state

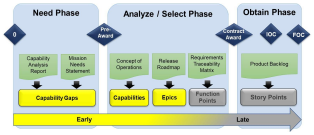
Constructive Threats

- Small sample size (20) poses a threat to statistical conclusion as it does not allow for detecting effects with greater power
- A larger sample is needed for confirmatory hypothesis testing

Main Takeaways

Program offices should consider using **Capability Gap** for ROM cost estimates in program early phases, especially since it is the only measure available at the Need Phase

Capability and **Epic** are more accurate and reliable effort predicters known throughout the lifecycle than **Story Points**, which is known later in the lifecycle



Practitioners can also use **Epics** to estimate software development effort and schedule post contract award, because it can easily be counted from Product Backlog or Release Roadmap