Software Effort Estimation Models for Contract Proposal Evaluation

Authors:
Wilson Rosa
Corinne Wallshein

June 6, 2017
2017 ICEAA Professional Development & Training Workshop

Purpose

- Present a set of effort estimating relationships and benchmarks for predicting both, **traditional and agile** software development projects using empirical data from 196 very recent US DoD programs.
  - Appropriate for crosschecking contract cost proposals as input parameters used in the analysis are typically available at bidding phase or earlier
  - Examines the validity of using of ***Initial Software Requirements*** as a proxy size measure

Outline

- Experimental Design
- Dataset Demographics
- Productivity Benchmarks
- Effort Estimation Models
  - Entire Dataset
  - Agile Software Subset
- Conclusion

Experimental Design

Instrumentation

- Questionnaire:
  - Software Resource Data Report" (SRDR) (DD Form 2630)
- Source:
  - Cost Assessment Data Enterprise (CADE) website:

    http://cade.osd.mil/Files/Policy/Initial_Developer_Report.xlsx

    http://cade.osd.mil/Files/Policy/Final_Developer_Report.xlsx
- Content:
  - Allows for the collection of project context, company information, requirements, product size, effort, schedule, and quality

Dataset used in Study

Empirical data from recent US DoD programs:

176 Paired SRDR Records from the Cost Assessment Data Enterprise (CADE)

Each paired record includes:

SRDR Initial Developer Report (Estimates)
&
SRDR Final Developer Report (Actuals)

http://dcarc.cape.osd.mil/Default.aspx

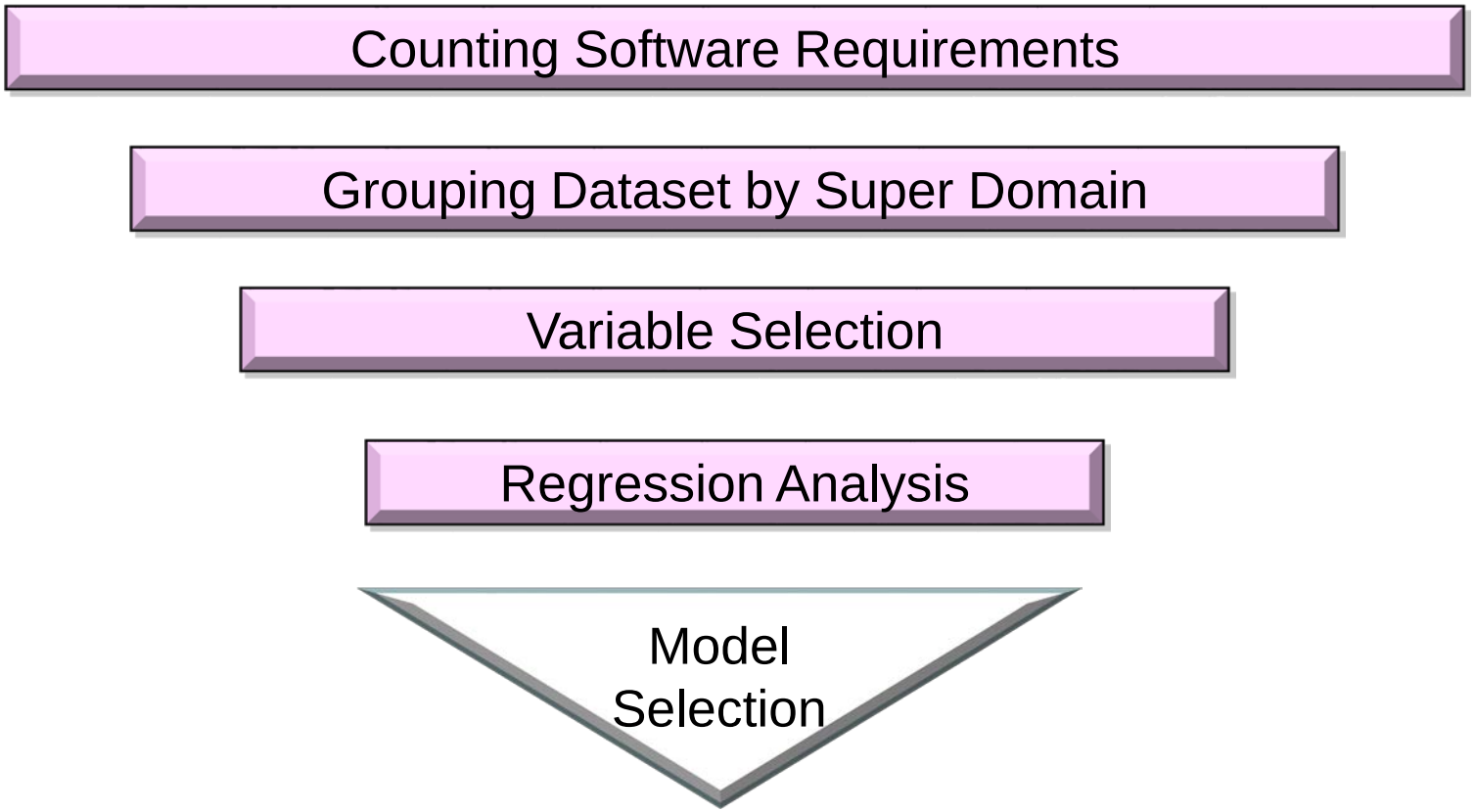
\+ **16** Agile SRDR Records from CADE

\+ **4** additional Agile Records from Proprietary Source

= **196** Records analyzed in this study

Data Normalization and Analysis Workflow

- Dataset normalized to “account for sizing units, application complexity, and content so they are consistent for comparisons” (source: GAO)



Counting Software Requirements

FORMULA	Initial Functional Requirements	+	Initial External Interfaces	= Initial Software Requirements
MEASURE	"shall" statements contained in the baseline Software Requirements Specification (SRS)		" shall" statements contained in the baseline Interface Requirements Specifications (IRS)	
SOURCE	Initial SRDR Report		Initial SRDR Report	

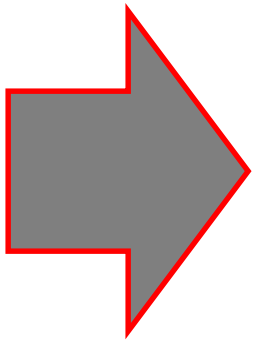
Initial Software Requirements typically available before contract award

Grouping Dataset by Super Domain

Approach

1) Dataset initially mapped into 17 Application Domains*
2) Then into 4 complexity groups called Super Domains

Application Domain
Software Tools
Training
Enterprise Information System
Enterprise Services
Custom AIS Software
Mission Planning
Test, Measurement, and Diagnostic Equipment
Scientific & Simulation
Process Control
System Software
Command & Control, Communications
Real Time Embedded
Vehicle Control/Payload
Signal Processing, Microcode & Firmware



Super Domain
Mission Support (SUPP)
Automated Information System (AIS)
Engineering (ENG)
Real Time (RTE)

*New DOD policy (http://cade.osd.mil/policy/srdr) requires that Application Domains are identified for reported software projects

Grouping Dataset by Super Domain Result

- Entire Dataset (196)

	Support	AIS	Engineering	Real Time	TOTAL
Aircraft	10	2	9	14	35
C4I	5	9	32	35	81
Business	1	18	0	0	19
Ordinance	0	0	0	1	1
Ship	1	0	0	10	11
UAV	5	1	2	5	13
Satellite	1	0	6	4	11
Missile	4	0	3	18	25
	27	30	52	87	196

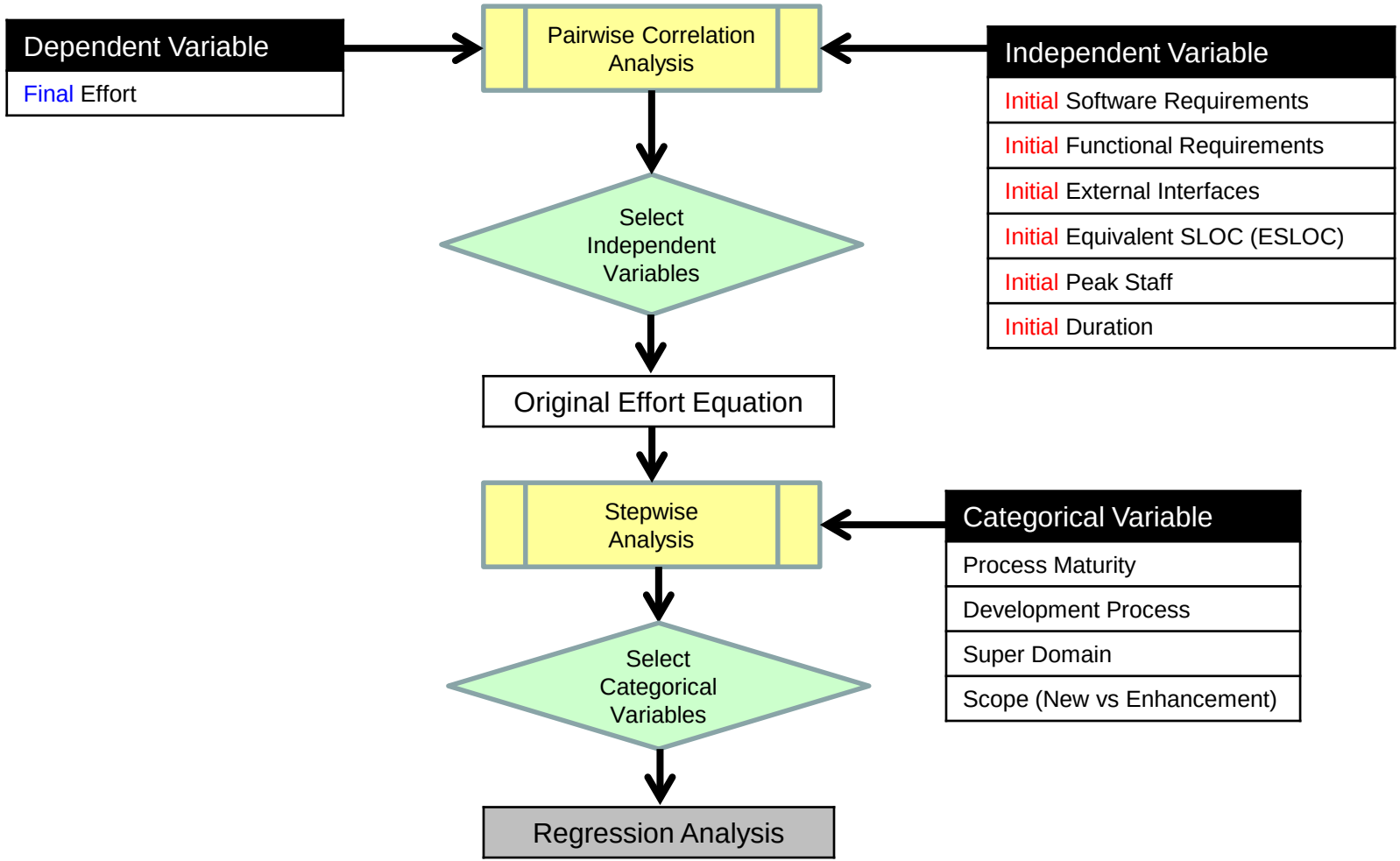
- Agile Subset (20 out of 196)

	Support	AIS	Engineering	Real Time	TOTAL
Aircraft	2	0	4	0	6
Business	1	3	0	0	4
C4I	0	1	3	5	9
Missile	0	0	0	1	1
	3	4	7	6	20

Dataset includes 20 Agile Software Development Projects

Variable Selection

1) Pairwise Correlation to select Independent Variables
2) Stepwise Analysis to select Categorical Variables



Pairwise Correlation Analysis

	Final Hours	Initial Functional Requirements	Initial External Interfaces	Initial Software Requirements	Initial ESLOC	Initial SLOC	Initial Duration
Final Hours	1.00	0.62	0.52	0.68	0.46	0.18	0.11
Initial Functional Requirements (IFQ)	0.62	1.00	0.19	0.98	0.16	0.23	0.07
Initial External Interfaces (IEI)	0.52	0.19	1.00	0.39	0.21	0.13	0.61
Initial Software Requirements (IFQ + IEI)*	**0.68**	**0.98**	0.39	1.00	0.44	0.24	-0.05
Initial ESLOC	0.46	0.16	0.21	0.44	1.00	0.64	0.10
Initial Total SLOC	0.18	0.23	0.13	0.24	**0.64**	1.00	-0.13
Initial Duration	0.11	0.07	**0.61**	-0.05	0.10	-0.13	1.00
Initial Peak Staff	**0.72**	0.38	**0.49**	**0.45**	0.38	0.15	-0.16

Strong Correlation Partial Correlation Weak Correlation

**Initial Software Requirements = Initial Functional Requirements + Initial External Interfaces

Findings:

- ***Initial Software Requirements*** *chosen as Size Measure -- stronger correlation to Final Hours than Initial ESLOC*
- ***Initial Peak Staff*** *added* as input variable -- strong correlation to Final Hours
- Initial Duration NOT considered as input variable -- weak correlation to Final Hours

Stepwise Analysis

Step 1: Initial Model

Initial Variable	Partial Correlations	T-Stat	P-Value	**Effort Estimation Equation**	R^2	MMRE (% Error)
REQ		12.9		Effort = *f* (REQ)	45%	101%

Step 2: Add Stepwise Variable

Stepwise Variables Added to Initial Model	Partial Correlations	T-Stat	P-Value	**Effort Estimation Equation**	R^2	MMRE (% Error)
SD	0.63	11.4	0.0000	Effort = *f* (REQ, SD)	67%	63%
STAFF	0.57	9.9	0.0000	Effort = *f* (REQ, STAFF)	63%	67%
SCOPE	0.10	1.4	0.1508	Effort = *f* (REQ, SCOPE)	63%	70%
Process Maturity	.021	0.2	0.7854	Effort = *f* (REQ, PMAT)	61%	76%

Strong Correlation Moderate Correlation Weak Correlation

Findings:

- *Super Domain and Peak Staff may be added to Initial Equation as these appear to improve model's accuracy*

SD = Super Domain, REQ = Initial Software Requirements, Staff = Initial Peak Staff

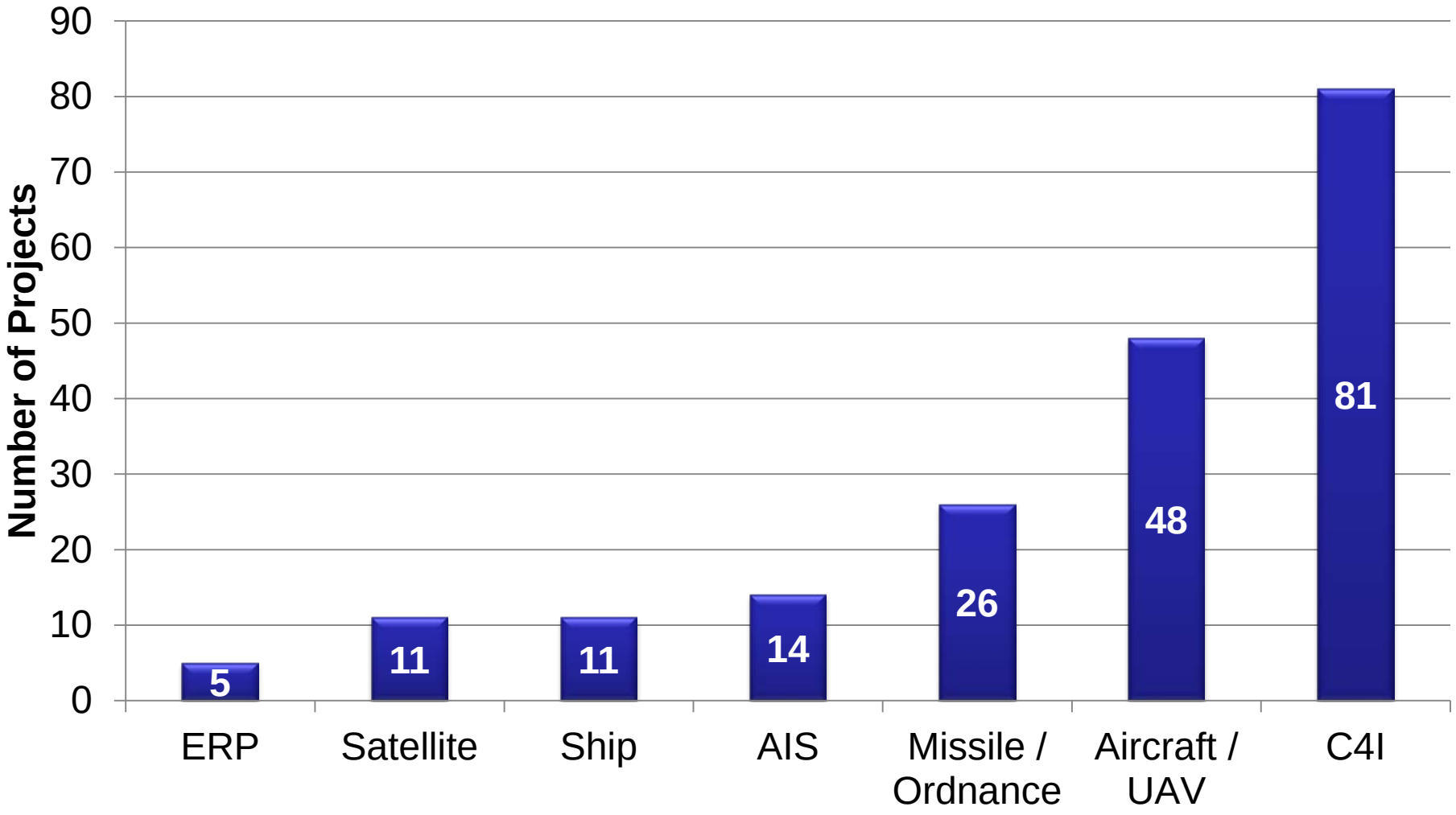
Model Selection

- Model Selection Based on T-Stat, lowest MMRE and CV

Measure	Symbol	Description
Coefficient of Variation	CV	Percentage expression of the standard error compared to the mean of dependent variable. A relative measure allowing direct comparison among models.
P-value	α	Level of statistical significance established through the coefficient alpha ($p \leq \alpha$).
Variance Inflation Factor	VIF	Indicates whether multi-collinearity (correlation among predictors) is present in a multi-regression analysis.
Coefficient of Determination	R^2	The Coefficient of Determination shows how much variation in dependent variable is explained by the regression equation.
Mean Magnitude of Relative Error	MMRE	Low MMRE is an indication of high accuracy. MMRE is defined as the sample mean (M) of the magnitude relative error (MME). MME is the absolute value of the difference between actual and estimated effort divided by the actual effort.

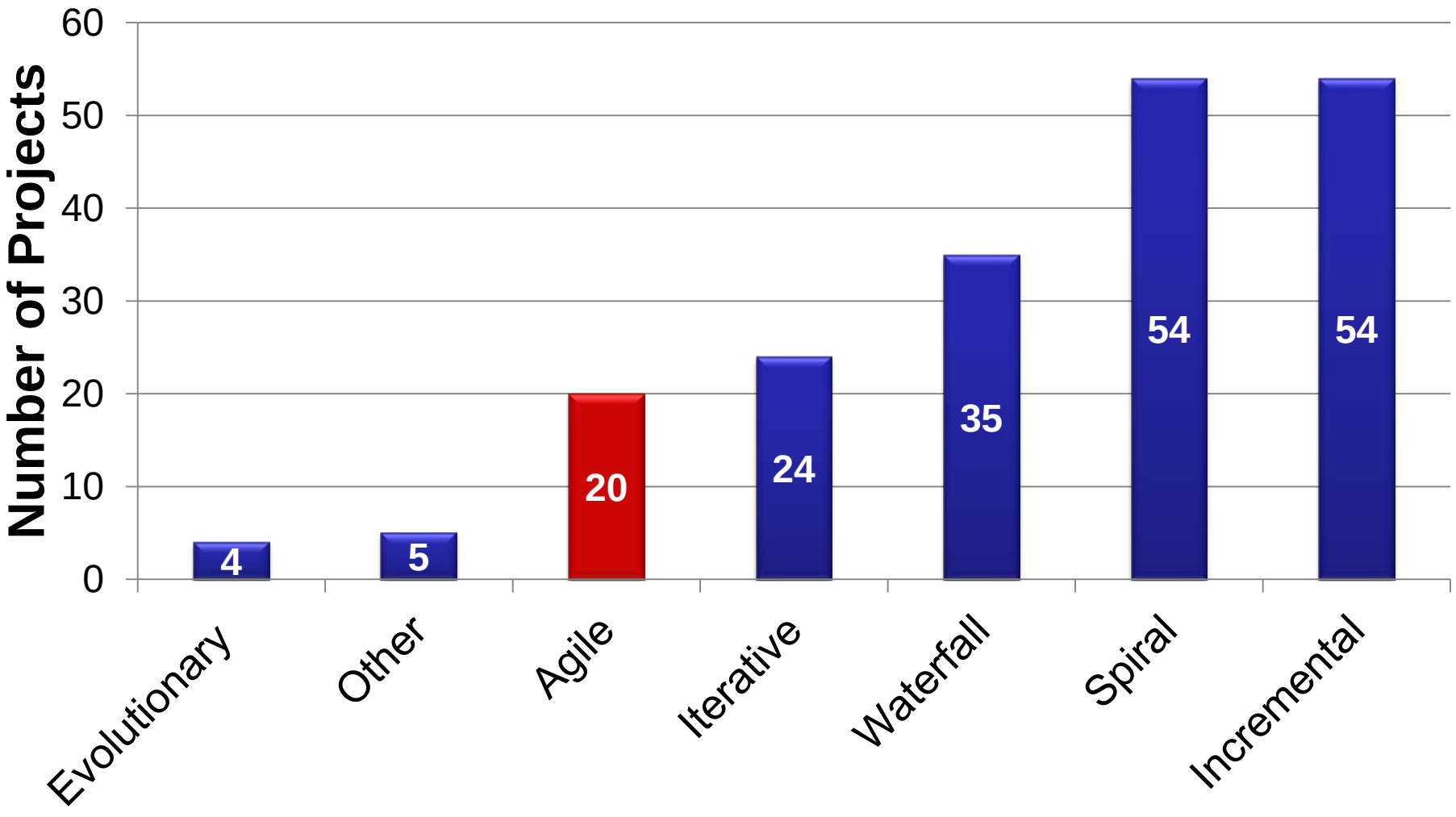
Dataset Demographics

Dataset by Operating Environment

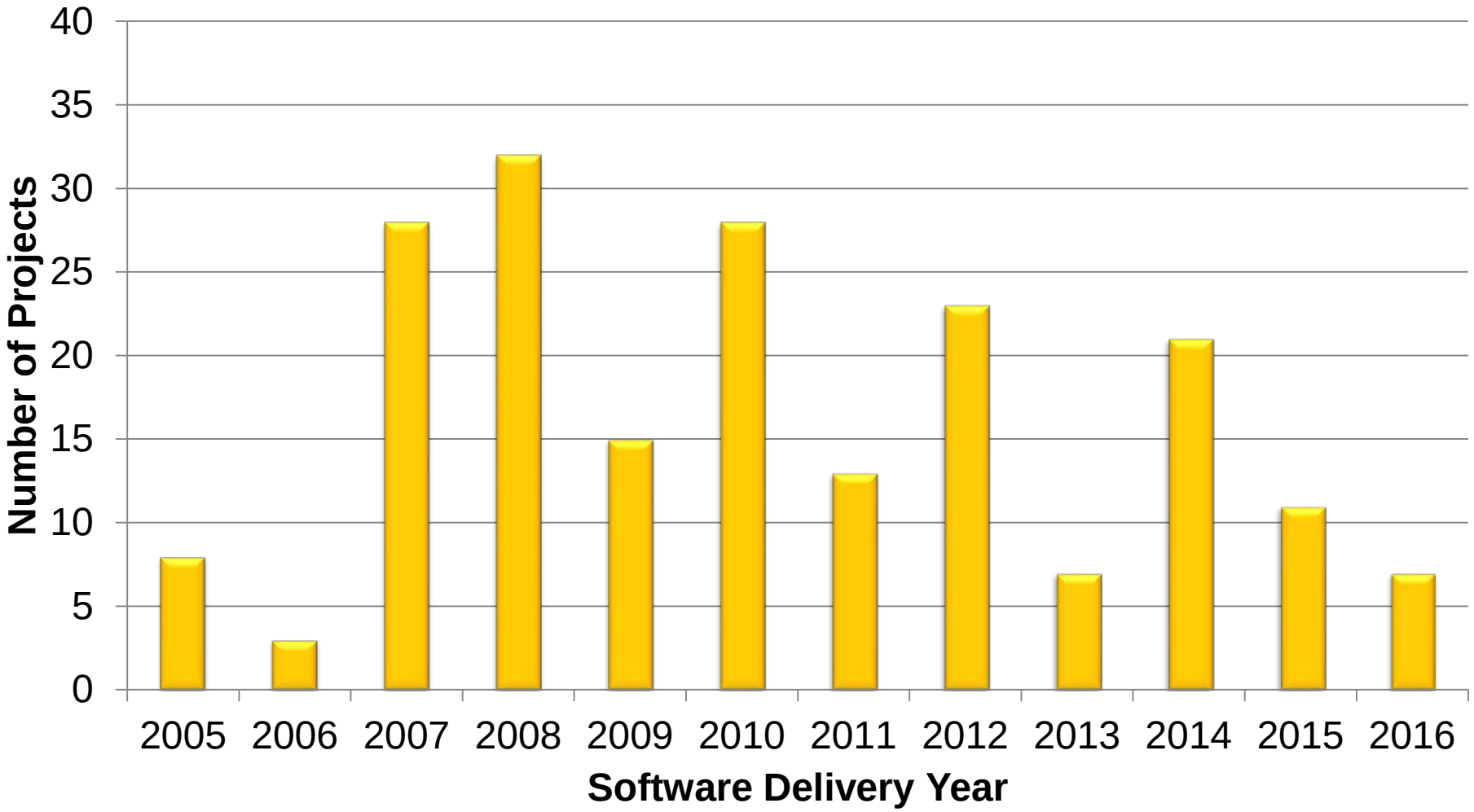


ERP = Enterprise Resource Planning; **AIS** = automated information system; **C4I** = command, control, communication, computer, intelligence

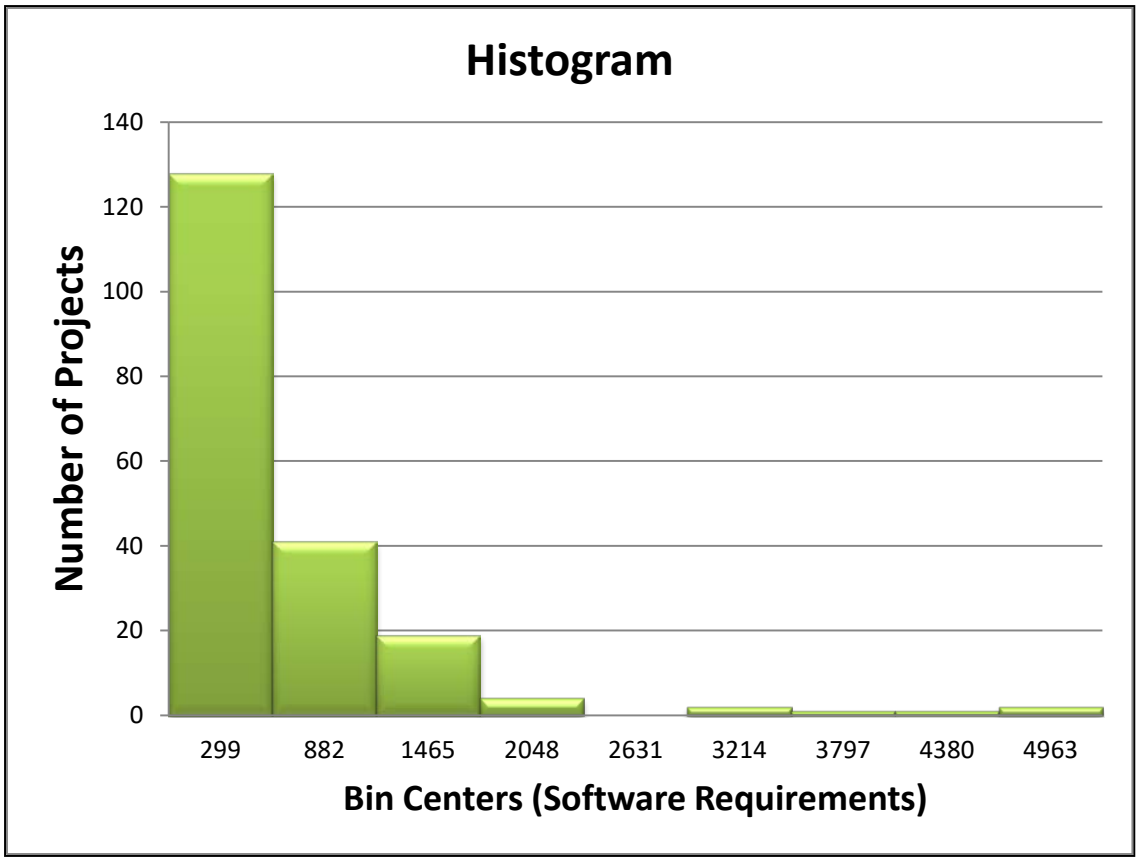
Dataset by Development Process



Dataset by Delivery Year



Dataset by Software Size Range

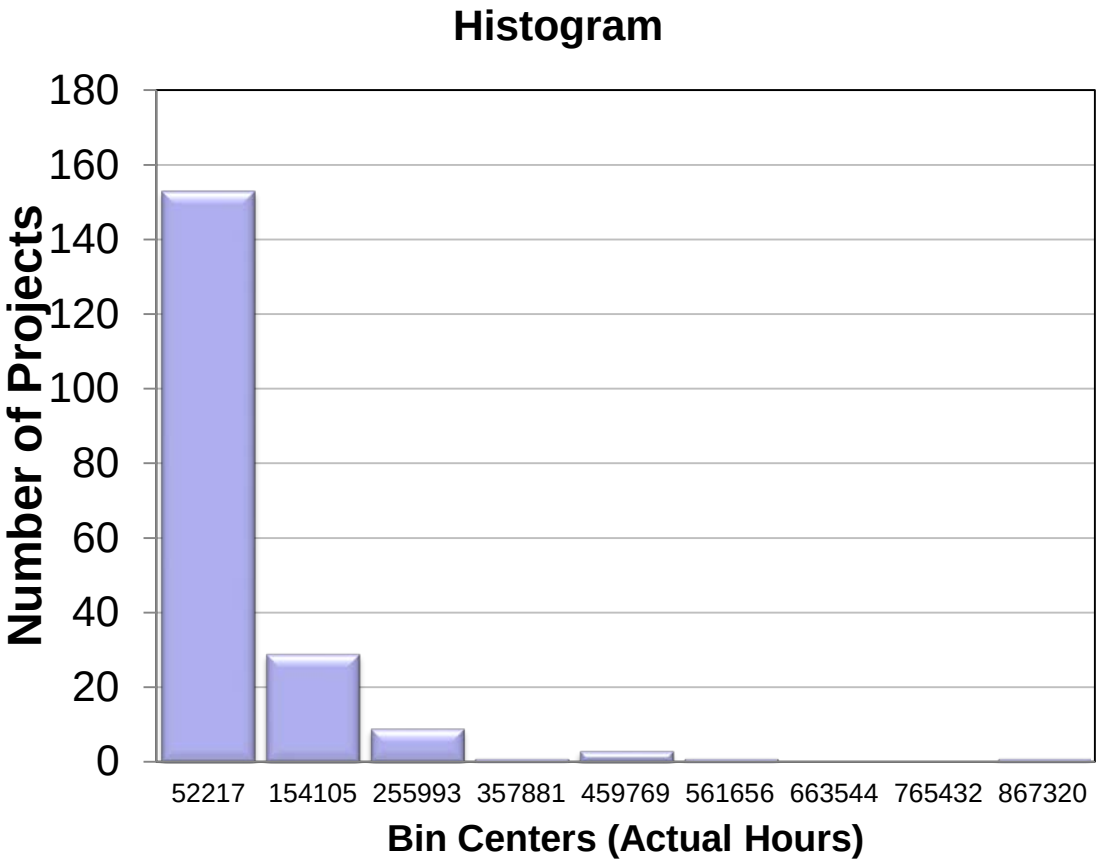


Software Requirements Range	Bin Center	Number of Projects
8.000 - 590.9	299	128
590.9 - 1174	882	41
1174 - 1757	1465	19
1757 - 2340	2048	4
2340 - 2922	2631	0
2922 - 3505	3214	2
3505 - 4088	3797	1
4088 - 4671	4380	1
4671 - 5254	4963	2

Most Projects in the study reported less than 600 Initial Software Requirements

*Software Size refers to the Initial Software Requirements

Dataset by Expended Effort (Actual Hours)



Actual Hours Range	Center	Number of Projects
1273 - 103161	52217	153
103161 - 205049	154105	29
205049 - 306937	255993	9
306937 - 408825	357881	1
408825 - 510712	459769	3
510712 - 612600	561656	1
612600 - 714488	663544	0
714488 - 816376	765432	0
816376 - 918264	867320	1

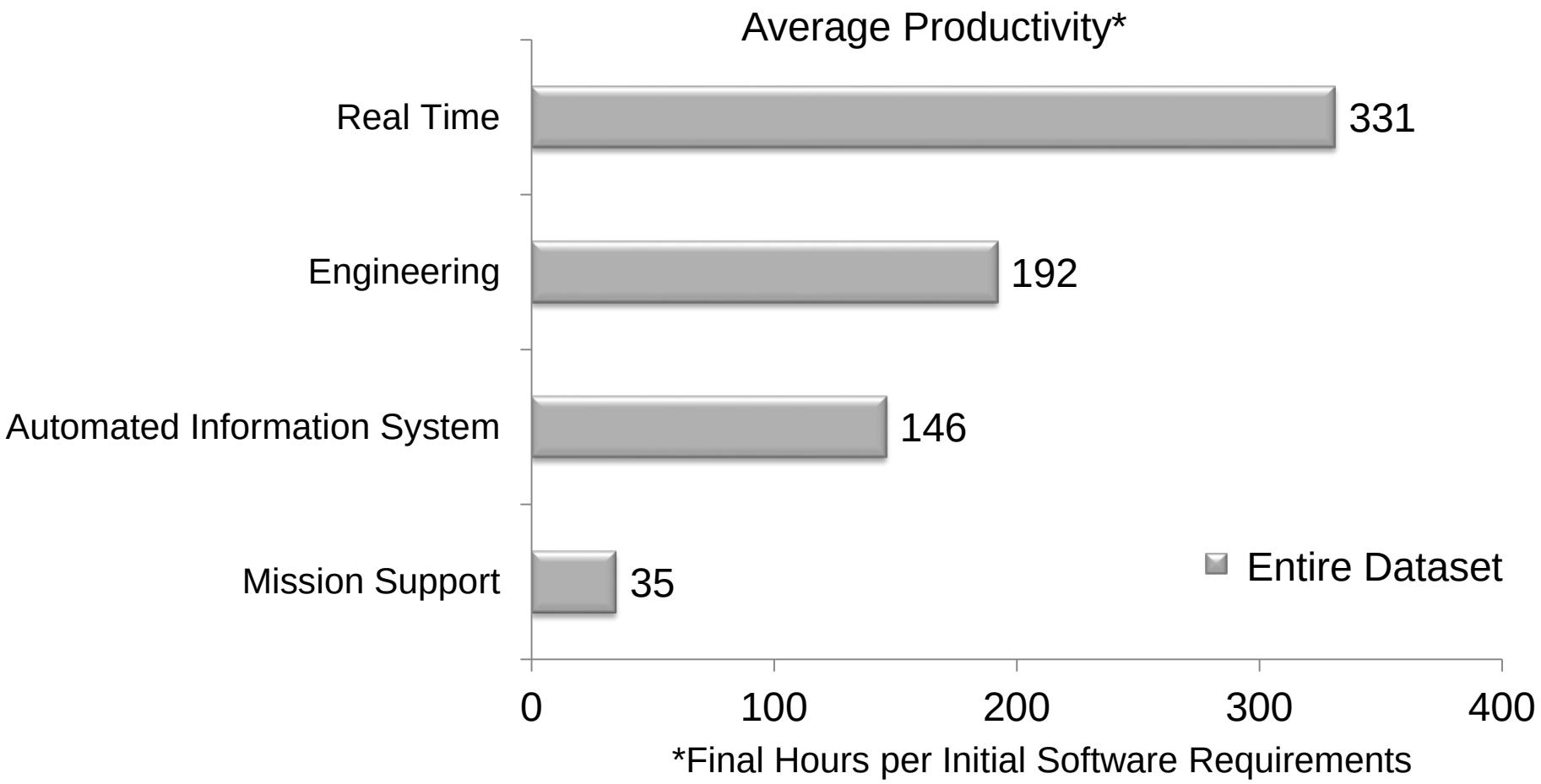
Most Projects in the study expended between 1,200-100,000 Hours

*Software Size refers to the Initial Software Requirements

Productivity Benchmarks

Productivity by Super Domain

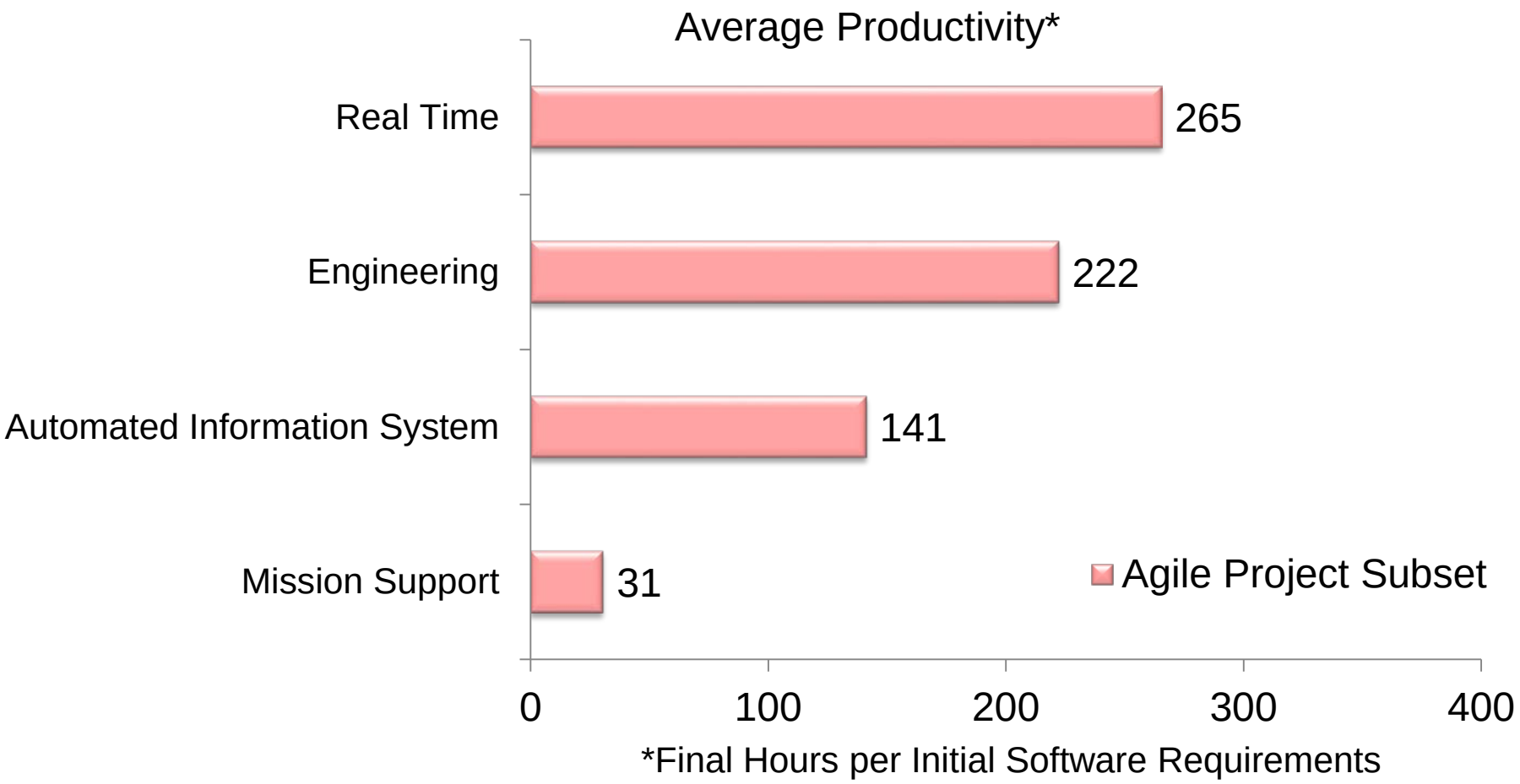
Entire Dataset (n=196)



Super Domain has a significant effect on Software Productivity

Productivity by Super Domain

Agile Project Subset (n=20)



Super Domain also shows significant effect on Agile Project Software Productivity

Productivity Comparison Agile vs Non-Agile

Size Range	Average Productivity* Agile	Non-Agile
1-100	415	466
101-500	159	189
501-5000	77	131
Composite Average	190	229

*Final Hours per Initial Software Requirements

When grouped by Size, Agile Software Projects appear to be more productive

Effort Estimation Models

Entire Dataset

Effort Model Variables

Name	Acronym	Type	Definition
Final Hours	EFFORT	Dependent	Actual software engineering effort (in Hours) at contract completion
Initial Software Requirements	REQ	Independent	Sum of Initial Functional Requirements and Initial External Interface Requirements estimated at contract award. Counting convention based on “shall statements”
Initial Peak Staff	STAFF	Independent	Estimated peak team size at contract award, measured in full-time equivalent staff
Super Domain	SD	Categorical	Software primary application. Four Types: Mission Support,Automated Information System (AIS), Engineering, or Real Time

Effort Estimation Model
Single Variable, Entire Dataset

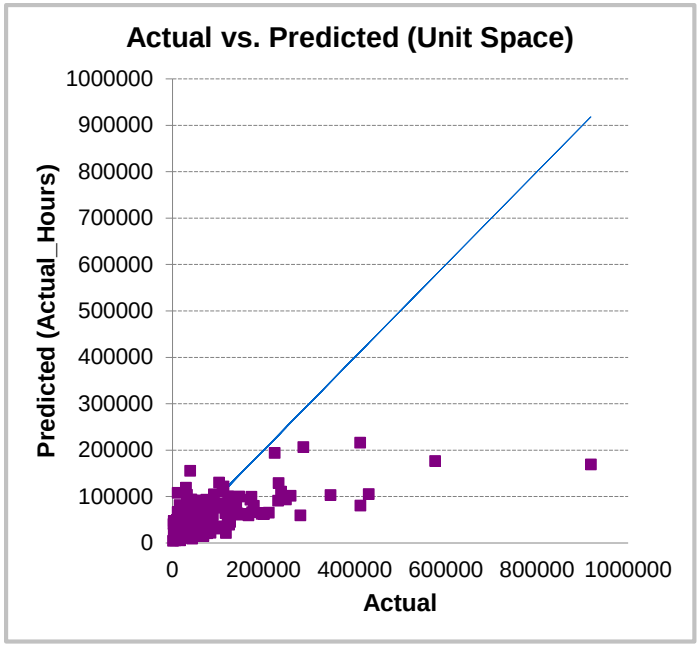
Model	**Equation Form**	N	R^2 %	CV %	Mean	MMRE %	REQ Min	REQ Max
1a	Effort = 1379 **x** $REQ^{0.59}$	196	45	58	74425	101	8	5254

Effort = Final Hours (or Actual Hours) at contract completion

REQ = Initial Software Requirements at contract start

Coefficient Statistics:

Variable	T-stat	VIF
Intercept	27.2	
REQ	12.8	
STAFF		
SD		



Effort Model not accurate when simply using REQ as input

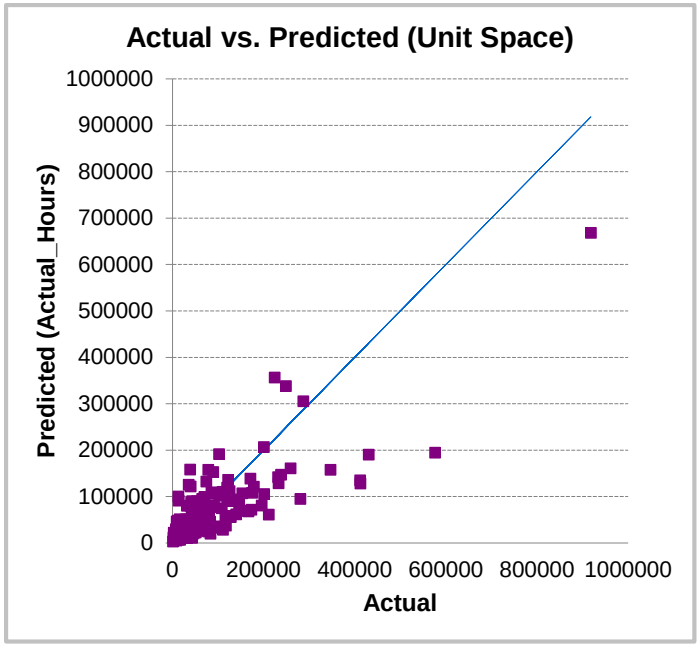
Effort Estimation Model
Two Variable, Entire Dataset

Model	**Equation Form**	N	R^2 %	CV %	Mean	MMRE %	REQ Min	REQ Max
1b	Effort = 1376 **x** $REQ^{0.3662}$ **x** $STAFF^{0.5225}$	196	63	47	74425	71	8	5254

Effort = Final Hours (or Actual Hours) at contract completion
REQ = Initial Software Requirements at contract start
STAFF = Initial (or Estimated) Peak Staff at contract start

Coefficient Statistics:

Variable	T-stat	VIF
Intercept	33.3	
REQ	8.3	1.36
STAFF	9.8	1.36
SD		



Model's accuracy dramatically improves when Peak Staff is incrementally added

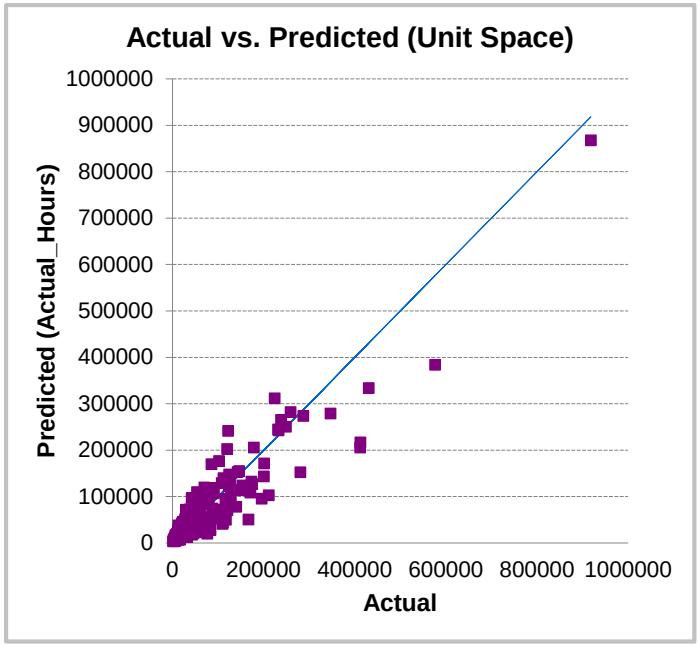
Effort Estimation Model
Three Variable, Entire Dataset

Model	**Equation Form**	N	R^2 %	CV %	Mean	MMRE %	REQ Min	REQ Max
1c	Effort = 244.3 **x** $REQ^{0.4803}$ **x** $STAFF^{0.4889}$ **x** $SD^{1.152}$	196	83	32	74425	41	8	5254

Effort = Final Hours (or Actual Hours) at contract completion
REQ = Initial Software Requirements at contract start
SD = Super Domain (1 if Mission Support, 2 if AIS, 3 if Engineering, 4 if Real Time)
STAFF = Initial (or Estimated) Peak Staff at contract start

Coefficient Statistics:

Variable	T-stat	VIF
Intercept	29.3	
REQ	15.6	1.45
STAFF	13.5	1.37
SD	15.0	1.07



Model's accuracy far better when Super Domain is treated along with REQ and STAFF

Effort Estimation Models

Agile Software Subset

Effort Estimation Model
Single Variable, Agile Software Subset

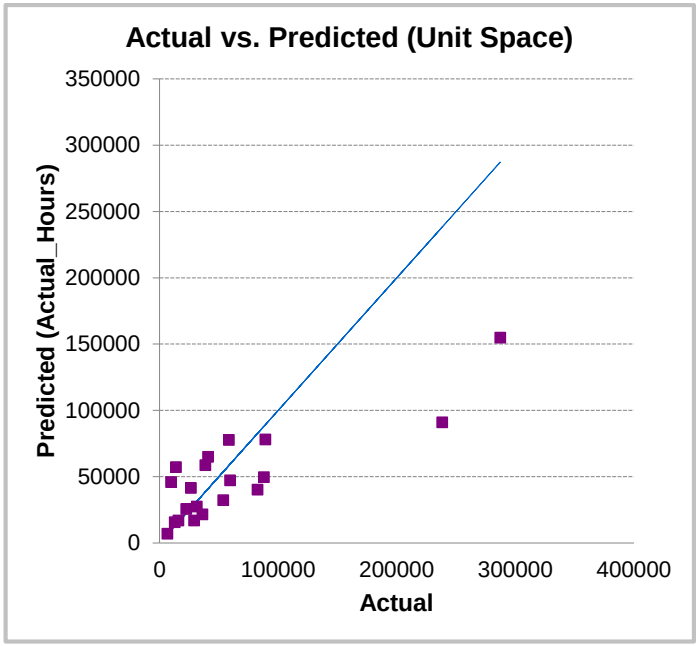
Model	**Equation Form**	N	R^2 %	CV %	Mean	MMRE %	REQ Min	REQ Max
2a	Effort = 2202 **x** $REQ^{0.5009}$	20	53	48	62140	64	10	4867

Effort = Final Hours (or Actual Hours) at contract completion

REQ = Initial Software Requirements at contract start

Coefficient Statistics:

Variable	T-stat	VIF
Intercept	12.3	
REQ	4.7	
STAFF		
SD		



Agile Estimation Model not accurate when simply using REQ as input

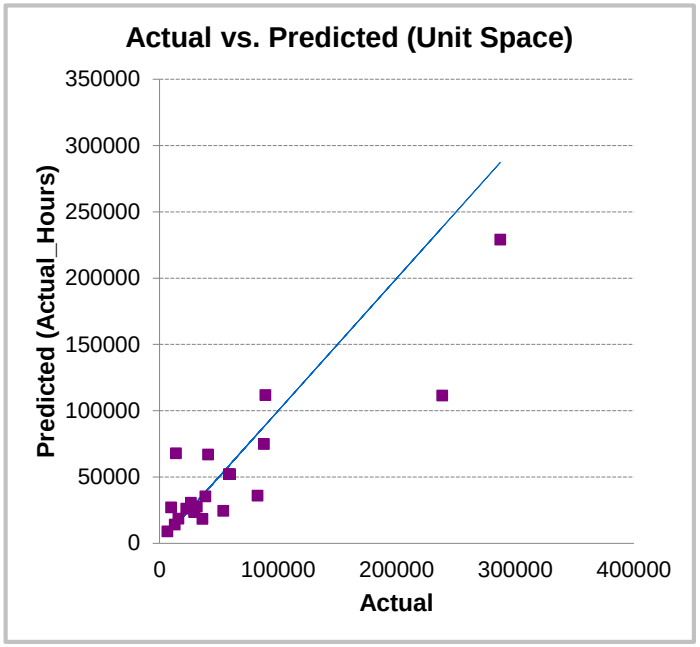
Effort Estimation Model
Two Variable, Agile Software Subset

Model	Equation Form	N	R^2 %	CV %	Mean	MMRE %	REQ Min	REQ Max
2b	Effort = 1045 **x** $REQ^{0.4071}$ **x** $STAFF^{0.4404}$	20	60	36	62140	52	10	4867

Effort = Final Hours (or Actual Hours) at contract completion
REQ = Initial Software Requirements at contract start
STAFF = Initial (or Estimated) Peak Staff at contract start

Coefficient Statistics:

Variable	T-stat	VIF
Intercept	10.2	
REQ	3.7	1.22
STAFF	2.0	1.22
SD		



Agile Estimation Model dramatically improves when Peak Staff is treated with REQ

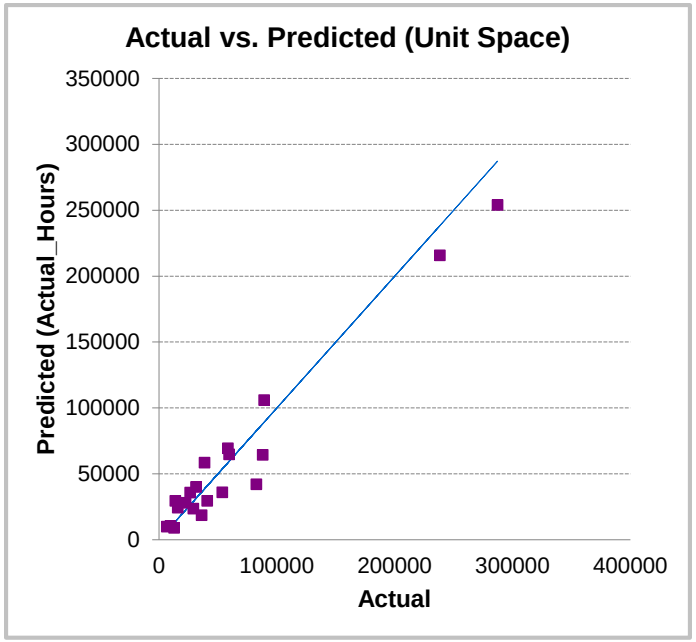
Effort Estimation Model
Three Variable, Agile Software Subset

Model	Equation Form	N	R^2 %	CV %	Mean	MMRE %	REQ Min	REQ Max
2c	Effort = 200.1 **x** $REQ^{0.5126}$ **x** $STAFF^{0.4782}$ **x** $SD^{1.001}$	20	81	22	62140	32	10	4867

Effort = Final Hours (or Actual Hours) at contract completion
REQ = Initial Software Requirements at contract start
SD = Super Domain (1 if Mission Support, 2 if AIS, 3 if Engineering, 4 if Real Time)
STAFF = Initial (or Estimated) Peak Staff at contract start

Coefficient Statistics:

Variable	T-stat	VIF
Intercept	9.1	
REQ	6.7	1.45
STAFF	3.2	1.37
SD	4.6	1.07



Agile Estimation Model is far more accurate when all 3 variables are added

Conclusion

Primary Findings

- ✓ Initial Software Requirements* is a valid size proxy for Agile and non-Agile Software Effort Estimation Models
- ✓ Estimation Models' accuracy improves when Peak Staff and Super Domain, are treated along with Initial Software Requirements*

Model	**Equation Form**	N	R^2 %	CV%	MMRE%
1a	Effort = 1378 **x** $REQ^{0.59}$	196	45	58	101
1b	Effort = 1372 **x** $REQ^{0.3667}$ **x** $STAFF^{0.5218}$	196	63	47	71
1c	Effort = 243.7 **x** $REQ^{0.4809}$ **x** $STAFF^{0.488}$ **x** $SD^{1.151}$	196	83	32	41

- ✓ Agile Estimation Model's accuracy also improves when Peak Staff and Super Domain are gradually added

Model	**Equation Form**	N	R^2 %	CV%	MMRE%
2a	Effort = 2202 **x** $REQ^{0.5009}$	20	53	48	64
2b	Effort = 1045 **x** $REQ^{0.4071}$ **x** $STAFF^{0.4404}$	20	60	36	52
2c	Effort = 200.1 **x** $REQ^{0.5126}$ **x** $STAFF^{0.4782}$ **x** $SD^{1.001}$	20	81	22	32

*Initial Software Requirements = Initial Functional Requirements + Initial External Interfaces

Model Limitations and Usefulness

- ❖ Since data was analyzed at the CSCI level, effort models may not be appropriate for projects reported at the Roll-Up Level.
- ❖ Do not use Effort Estimation Models if your input parameters are outside of the model's dataset range.
- ✓ **P**roposed Effort Models may be used to either crosscheck or validate contract proposals as input parameters used in the study are typically available during proposal evaluation phase
- ✓ **A**pplicable for both, Defense and Business Systems
- ✓ **A**pplicable for Agile and non-Agile Software Projects