2° International Conference on
IT Data collection, Analysis and Benchmarking
Tokyo (Japan) - October 22, 2014

Analysis of Factors that Affect Productivity of Enterprise Software Projects

Tsuneo Furuyama
Tokai University
furuyama@tokai-u.jp

Motivation

A large-scale database of software projects

Can a cost model be derived from the database?

- Background
- Analysis data
- Data analysis methods
- Analysis results & discussions
  - Regression analysis for quantitative variables
  - One-dimensional analysis for qualitative variables
  - Two-dimensional analysis for qualitative variables
  - Comparison to the COCOMO II
- Summary

Backgrouond

Cost model = Theory /Approach × Experience database /Expert's knowledge and experience

- (Multiple) regression analysis
- Analogy
- Expert judgment
- Neural networks
- Bayesian networks

…

Multiple Regression Analysis

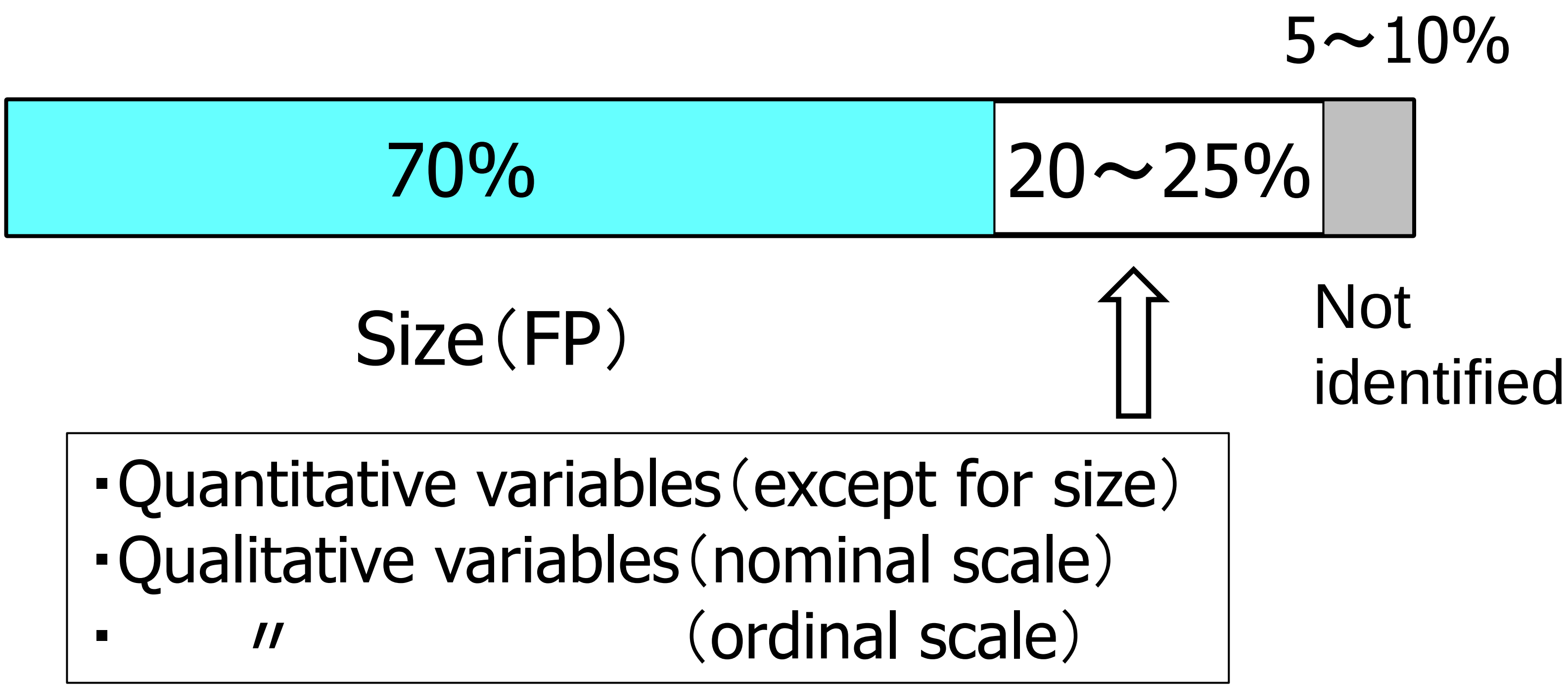
Multiple regression analysis is a fundamental approach or baseline of various approach for constructing cost model.

Models applying multiple regression analysis with an experience database have been the majority even in recent years.[1]

Ordinary least squares regression in combination with a logarithmic transformation performed the best in comparison with various types of cost models.[2]

[1] Jorgensen, M, and Shepperd, M.: A Systematic Review of Software Development Cost Estimation Studies, IEEE Trans. SE, Vol.33, No.1, pp.33-53 (2007).
[2] Dejaeger, K., Verbeke,W., Martens, D., and Baesens, B.: Data Mining Techniques for Software Effort Estimation: A Comparative Study, IEEE Tr. SE, Vol. 38, No.2, pp. 375-397 (2012).

Factors affecting effort



COCOMO & COCOMO II

- Formula of COCOMO:

$$Effort = A \times (LOC)^{B} \times \prod_{i=1}^{15} Cd_i$$

- Formula of COCOMO II:

$$Effort = A \times (size)^{B+\sum_{j=1}^{5} Sf_j} \times \prod_{i=1}^{17} Cd_i$$

Cd: Cost driver
Sf: Scale factor
Size: SLOC or FP

Overview of PROMISE repository

Name of database		coc81	coc81-dem	coc2000	nasa93	nasa93-dem	Maxwell	usp05	China	Kitchen-ham
Number of projects		63	63	125	101	93	62	919	499	145
Attributes		19	27	50	24	27	27	17	19	10
	Assigned to value in ordinal scale	*15	22	*22	15	22	15	2	0	0
Number of missing values		0	0	0	0	0	0	83	0	13

Almost complete dataset !

Note: Databases including more than or equal to 50 projects are listed.
* The value of each variable is selected among a few pre-defined numeric constants.

SEC database

- Enterprise software project data collected by the Software Reliability Enhancement Center (SEC) of the Information-Technology Promotion Agency (IPA) in Japan.
- Strengths: 1) More than 200 variables are defined, even excluding detailed variables. (*)
  2) The size of this database is more than 3000.
- Weakness: a large amount of missing values

(*) ISBSG database has at least 86 variables, but most are details of efforts, defects, and project profiles subject to nominal scales.

How to effectively use SEC database?

(1) Multiple regression analysis after list wise deletion

- If high priority is given to the number of projects when making a complete subset, some important variables may be lost before analysis, or

- if high priority is given to the number of variables, the number of selected projects is not sufficient for analysis.

(2) Multiple regression analysis with a step-by-step progressive selection of variables

- The subset data in every step are different so that the combinations of variables unselected in the prior steps may be more appropriate than the current combinations.

Can missing values be interpolated?

- Multiple imputation method for interpolating missing values to complete the dataset, have been proposed.
- For missing data in the category of missing completely at random (MCAR) or missing at random (MAR), this method is effective.
- However, for missing values in the category of missing not at random (MNAR) or too many missing values, interpolation is not effective.

➔How to effectively utilize information included in project database to feedback useful findings to development teams without interpolating missing values.

Analysis for SEC database

First step:

- Multiple regression analysis for quantitative variables
  - -- FP, number of test cases, and number of faults
- Analysis of variance for qualitative variables alone
  - -- 39 variables subject to ordinal scale
  - -- Cross effects of these variables were attempted.

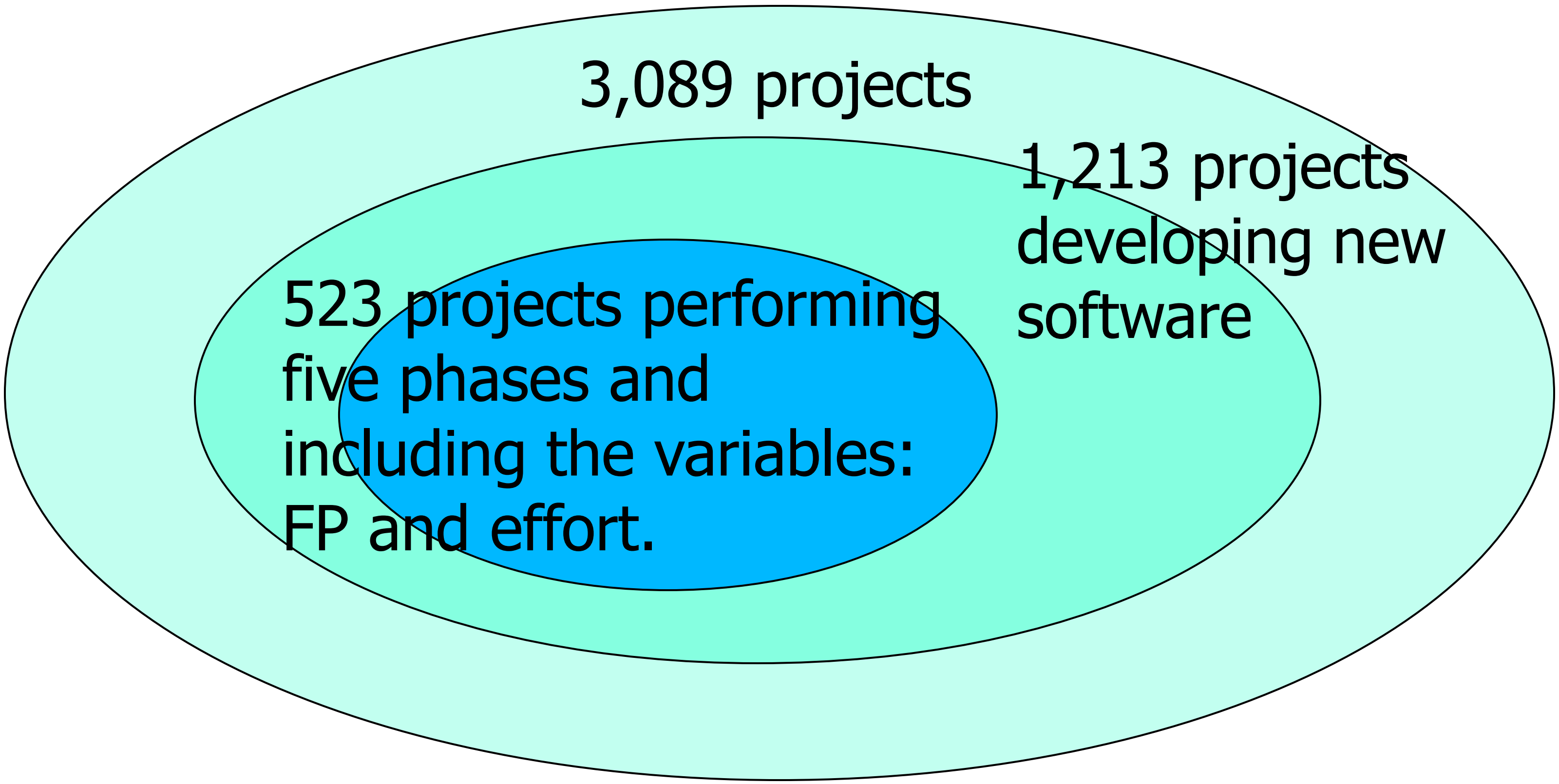
Future challenges:

- Multiple regression analysis for multiple variables
- Path analysis

Analysis of Factors that Affect Productivity of Enterprise Software Projects

- Background
- Analysis data
- Data analysis methods
- Analysis results & discussions
  - -- Regression analysis for quantitative variables
  - -- One-dimensional analysis for qualitative variables
  - -- Two-dimensional analysis for qualitative variables
  - -- Comparison to the COCOMO II
- Summary

Number of projects for analysis



Five phases: fundamental design, detail design, manufacture, system test, and total test by vendor

Criterion and predictor variables

(1) Quantitative predictor variables

- Criterion variable: effort
- Predictor variables:
    - -- Size based on FP
    - -- Number of test cases
    - -- Number of faults detected during development

Fundamental log values of quantitative variables

	Effort/FP	Effort	FP	Number of test cases	Number of faults
Mean	0.988	3.858	2.870	3.133	2.061
Medium	1	3.819	2.862	3.168	2.043
Standard deviation	0.368	0.667	0.489	0.754	0.663
Number of projects	523	523	523	324	310

Note: Unit of effort is person-hours.

(2) Quantitative predictor variables

- Criterion variable: productivity= $\frac{\mathrm{effort}}{\mathrm{FP}}$

※ Note that the smaller the ratio, the higher the productivity.

- Predictor variables: 39 variables subject to ordinal scale in 6 categories

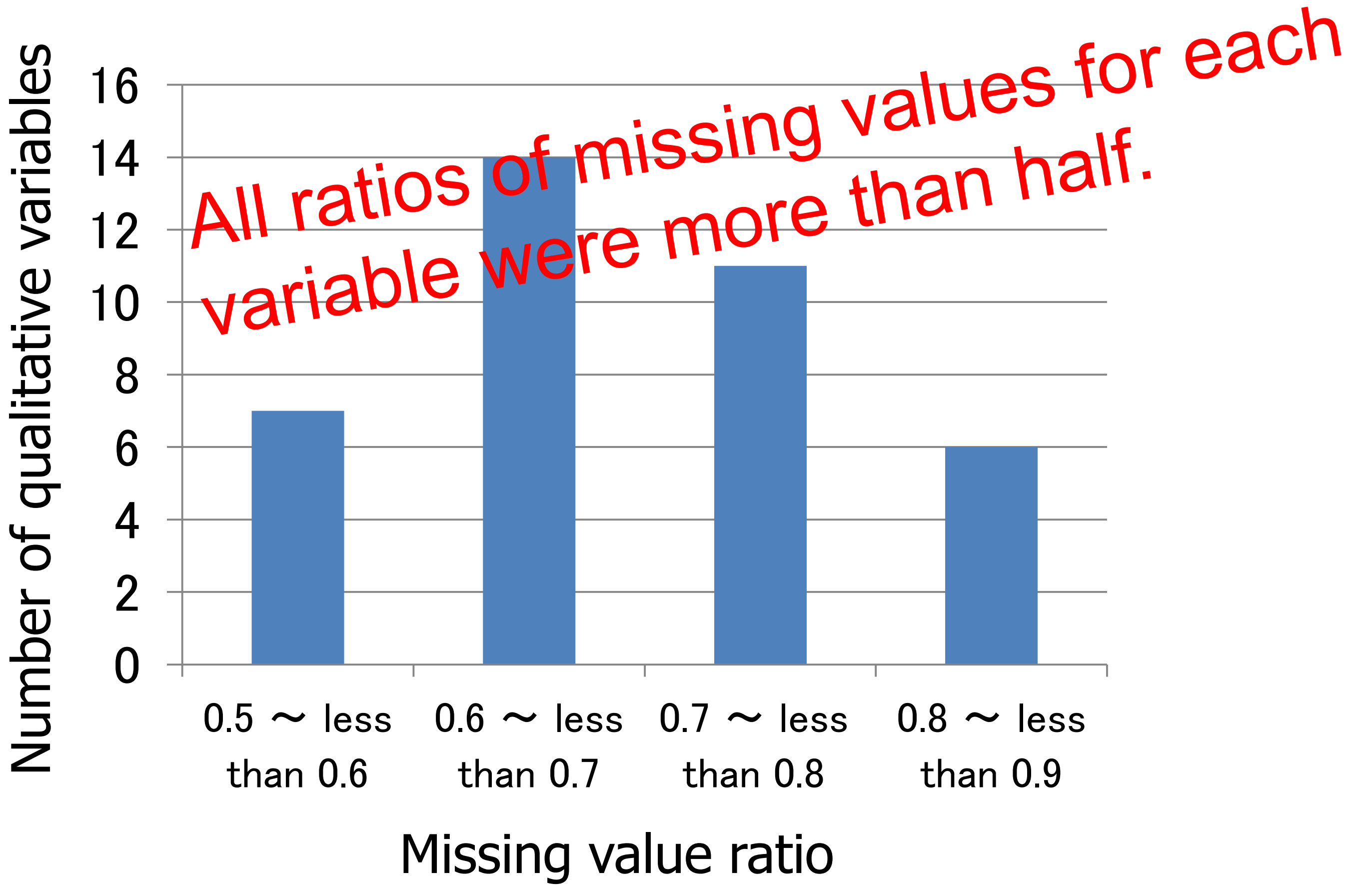
1) Equal to or more than 50 data values

2) All levels have more than or equal to 10 data values and 15% of the total amount of data values of the variable.

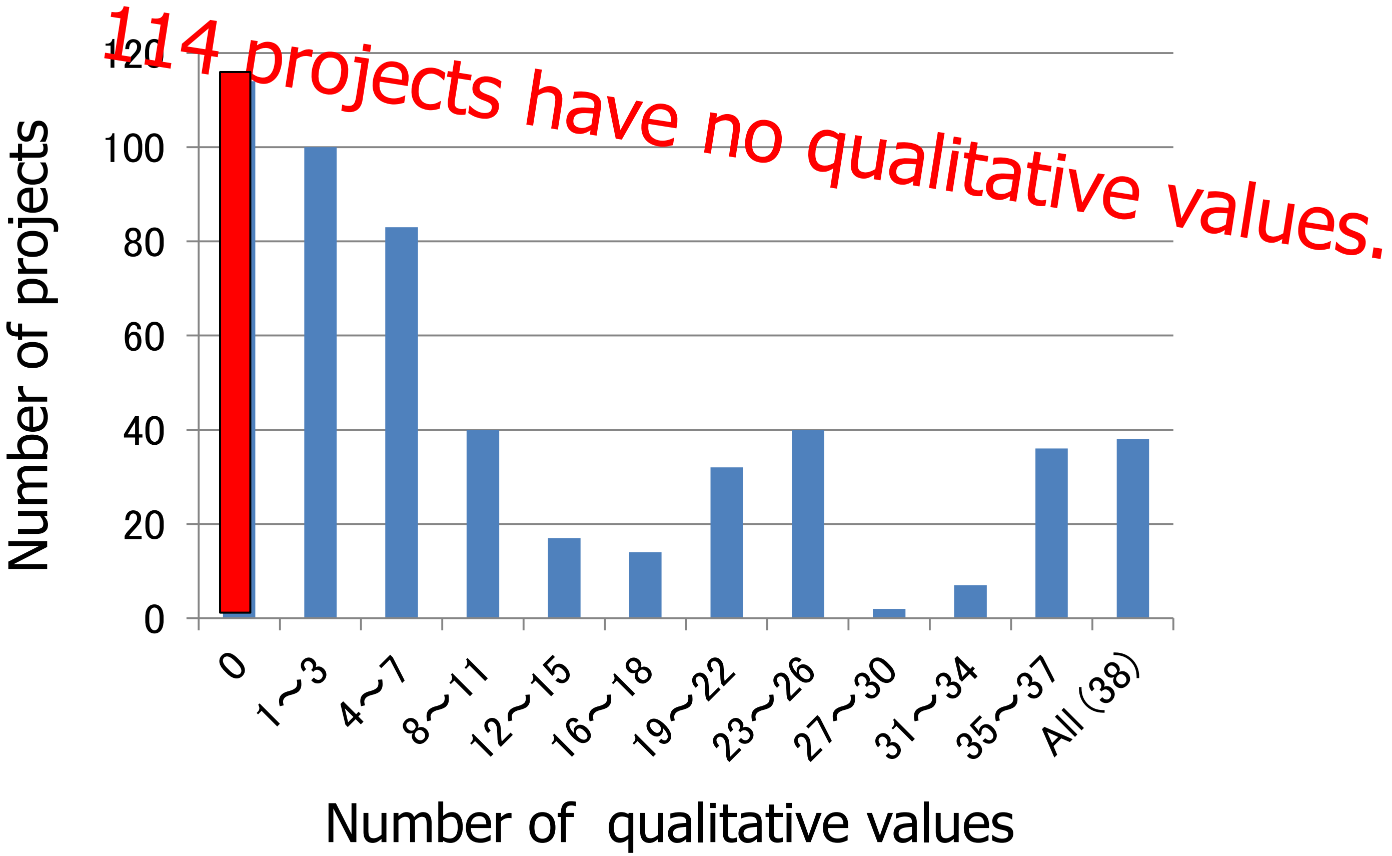
Candidate qualitative predictor variables

Category	Variables
Overall project	Introduction of new technology/Clarity of role assignments and each staff member's responsibilities /Clarity of objectives and priorities/Quantitative quality/Quantitative quality criteria for delivery/Quality assurance system in fundamental design (FD) phase
	Working space/Noise conditions
Evaluation of plan	Evaluation of plan (cost/duration/quality)
Tool usage	Similar project/Project management tool/Configuration control tool/Design support tool/Document-generation tool/Debug and test tool/Code generator/ Development framework
User's skill levels and commitment	Commitment of user to defining requirement specifications/Commitment of user to acceptance test/User's experience of in developing systems/User's business experience/User's understanding level for design content/Clarity of role assignments and each organization's responsibilities between user and development staff
Requirement levels	Clarity level of requirement specifications/Requirement level (reliability/usability/ performance and efficiency/portability/maintainability/security)/Legal regulation
Development staff's skill levels	Project manager's skill level
	Staff's skill levels (business experience/analysis and design experience/ experience with languages and tools/experience with development platform)
	Test team's skill levels and size

Missing value ratios for qualitative variables



Number of qualitative values each project has



- Background
- Analysis data
- Data analysis methods
- Analysis results & discussions
  - -- Regression analysis for quantitative variables
  - -- One-dimensional analysis for qualitative variables
  - -- Two-dimensional analysis for qualitative variables
  - -- Comparison to the COCOMO II
- Summary

Data pre-processing

(1) Transformation of quantitative variables

All quantitative variables, including that for productivity, were logarithmically transformed before analysis.

(2) Merging of levels of qualitative data

The levels of variables with more than two levels were merged into two levels by combining the upper two levels into one level and the lower two levels into another, or combining the top level or the lowest level into one level and the other three levels into the other level.

The boundary was determined so that the numbers of two levels were as balanced as possible.

Data analysis methods

(1) Quantitative variables – Multiple regression analysis

(2) Qualitative variables

- One-dimensional analysis of variance
  - -- The Welch-adjusted analysis of variance
  - -- Significant level: 5%
  - -- Both means of two levels are more than or less than the mean of all 523 projects: 0.988.

    ➔Variable was regarded as "biased" and reserved.
- Two-dimensional analysis of variance
  - -- At least three of all combinations of the 2-by-2 levels of the cross table, that is, six pairs, must be significant.

- Background
- Analysis data
- Data analysis methods
- Analysis results & discussions
  - -- Regression analysis for quantitative variables
  - -- One-dimensional analysis for qualitative variables
  - -- Two-dimensional analysis for qualitative variables
  - -- Comparison to the COCOMO II
- Summary

Regression analysis of quantitative variables

Number of predictor variables	Regression coefficient			Constant	Multiple corre-lation coeffi-cient	Adjusted coefficient of determina-tion	Number of projects
	Size	Number of test cases	Number of faults				
1	1.147	-	-	0.566	0.841	0.706	523
	-	0.556	-	2.237	0.653	0.425	324
	-	-	0.686	2.575	0.706	0.497	310
2	0.893	0.230	-	0.637	0.855	0.730	324
	0.850	-	0.276	0.906	0.844	0.711	310
	-	0.278	0.483	2.108	0.772	0.594	288
3	0.746	0.193	0.182	0.797	0.868	0.750	288

Meaning of results of quantitative variables

$$\log E = 0.746 \log S + 0.193 \log T + 0.182 \log B + 0.797$$

$\sim \frac{1}{8}$ $\sim \frac{1}{5}$ $\sim \frac{1}{5}$

$$\frac{E}{S} = 6.26 \times S^{0.121} \left(\frac{T}{S}\right)^{0.193} \left(\frac{B}{S}\right)^{0.182}$$

Adjusted coefficient of determination: 70.6% ➔75.0%

※ E: effort, T: number of test cases, B: number of faults,
E/S: productivity, T/S: test case density, B/S: fault density

- Background
- Analysis data
- Data analysis methods
- Analysis results & discussions
  - -- Regression analysis for quantitative variables
  - -- One-dimensional analysis for qualitative variables
  - -- Two-dimensional analysis for qualitative variables
  - -- Comparison to the COCOMO II
- Summary

Two questions

a) What is the productivity ratio of typical projects(*) of two groups?

b) Is the selected predictor variable appropriate for affecting productivity as determined from the literature and my experience?

(*) A typical project is defined as a project whose size, effort, and productivity are the inverse logarithmic transformation of their means in the logarithmic scale.

Analysis of variance for overall project

Variable	Level	Number of projects	Productivity*		Typical project	
			Mean	Vari-ance	Produc-tivity **	Produc-tivity ratio
Clarity of role assign-ments and each staff member's responsibilities	Very clear	84	0.820	0.147	6.61	1.71
	Fairly clear + Little clear + Unclear	130	1.053	0.151	11.31	
Clarity of objectives and priorities	Very clear	70	0.754	0.128	5.67	1.88
	Fairly clear + Little clear + Unclear	121	1.029	0.152	10.68	
Working space	Levels: a + b (broad)	89	0.798	0.118	6.28	1.56
	Levels: c + d (narrow)	66	0.991	0.208	9.80	
Quality Assurance system in FD phase	Done by project members	125	0.985	0.136	9.65	1.84
	Done by quality assurance staff	59	1.249	0.121	17.73	

*Logarithmic scale, **Unit is person-hours/FP

Findings for overall project

- Role assignments and each staff member's responsibilities is very clearly defined.
- Objectives and priority is very clearly defined.
- Working space is broad enough.

➔These circumstances make developers work effectively without physical stress or mental confusion.

➔Higher productivity than otherwise

- Project members ensure the quality of the design specifications is higher than that in which the staff of the quality assurance team does.

➔ A little different result from what is written in textbooks or reported in research papers.

➔ Analysis from the viewpoint of quality (or reliability) instead of productivity may lead to a different conclusion.

Analysis of variance for tool usage

Variable	Level	Number of projects	Productivity		Typical project	
			Mean	Variance	Produc-tivity	Productiv-ity ratio
Similar project	Usage	54	1.009	0.165	10.21	1.47
	No usage	57	0.843	0.165	6.96	
Project management tool	Usage	111	1.004	0.181	10.09	1.63
	No usage	64	0.791	0.110	6.19	
Document-generation tool	Usage	60	0.653	0.109	4.50	2.21
	No usage	93	0.998	0.133	9.95	
Debug and test tool	Usage	52	1.003	0.208	10.07	1.58
	No usage	99	0.806	0.126	6.39	
Development framework	Usage	91	0.923	0.158	8.37	1.40
	No usage	75	1.070	0.156	11.75	

Findings for tool usage

- Development framework usage
- Document-generation tool usage

➔Higher productivity than otherwise (reasonable results)

- Similar project usage
- Project managing tool usage
- Debug and test tool usage

➔Lower productivity than otherwise (unexpected results)

➔Further study is needed since the usage of these tools or similar project may contribute to improving reliability.

Analysis of variance for other categories

Category	Variable	Level	Number of projects	Productivity		Typical project	
				Mean	Variance	Productivity	Productivity ratio
User's skill levels and commitment	Commitment to defining requirement specifications	Sufficient commitment + Fair commitment	132	0.917	0.162	8.27	1.34
		Insufficient commitment + No commitment	81	1.043	0.155	11.05	
Require-ment levels	Requirement level for reliability	Extremely high + High	81	1.016	0.194	10.38	1.85
		Medium + Low	87	0.750	0.101	5.62	
	Requirement level for security	Extremely high + High	64	1.128	0.158	13.43	2.85
		Medium + Low	89	0.672	0.074	4.70	
Develop-ment staff's skill levels	Project manager's skill level*	Levels 5, 6 and 7 (high level)	58	1.088	0.195	12.25	1.81
		Levels 3 &4 (low level)	108	0.831	0.140	6.77	

*Categorized in accordance with IT skill standard defined by METI.

Finding for other categories

- User's insufficient or no commitment to defining the requirement specifications
  - ➔Lower productivity than otherwise (reasonable result)
- High requirement levels for security or reliability
  - ➔Lower productivity than otherwise (reasonable results)
- Projects managed by a person with a high skill level was 1.81 times lower than that of projects managed by a person with a low skill level, is inappropriate.
  - ➔ Further investigation is needed.

Difference in project features conducted by high and low skill PMs

	Project size in FP		Test case density		Fault density	
PM skill level	high	low	high	low	high	low
Mean*	3.114	2.882	0.303	0.047	-0.884	-0.932
Variance*	0.221	0.199	0.166	0.388	0.543	0.396
Number of projects	58	108	40	65	40	65
P value	0.3%		1.2%		18.2%	
Ratio of typical projects	1.71		1.80		1.12	

* Logarithmic scale

Peason's Chi-squared Test between PM skill levels and three predictor variables

		User's commitment to defining requirement specifications		Requirement level for reliability		Requirement level for security	
		Sufficient commitment + Fair commitment	Insufficient commitment + No commitment	Extremely high + High	Medium + Low	Extremely high + High	Medium + Low
PM skill level	High	40 >>	4	31 >>	12	26 >>	17
	Low	78	27	47	58	37	67
P value		3.9%		0.4%		1.0%	

PM with a high skill level

- A PM with a high skill level tends to manage a software project developing large-scale software with high requirement levels for reliability or security.
- One of their duties is to run much more test cases per FP to maintain software quality than a PM with low skill level.

Productivity ratios for variables alone

- Requirement level for security: 2.85
- Document-generation tool usage: 2.21
- Other 11 variables: less than 1.9

➔Most variables alone affect productivity of less than 2.0.

Analysis of Factors that Affect Productivity of Enterprise Software Projects

- Background
- Analysis data
- Data analysis methods
- Analysis results & discussions
  - -- Regression analysis for quantitative variables
  - -- One-dimensional analysis for qualitative variables
  - -- Two-dimensional analysis for qualitative variables
  - -- Comparison to the COCOMO II
- Summary

Results of two-dimensional analysis of variance

Combinations that cause synergy				Number of projects	Productivity		Productivity ratio*
Variable	Level	Variable	Level		Mean	Variance	
Requirement level for security	Extremely high + High	Working space	Levels: c + d (Narrow)	34	1.264	0.146	3.48
				110	0.722	0.086	
		Development framework	No usage	30	1.291	0.091	3.36
				77	0.765	0.101	
		Clarity of role assignments and each staff member's responsibilities	Fairly clear + Little clear + Unclear	41	1.210	0.130	3.06
				109	0.724	0.092	

*Ratio of typical projects

Results of two-dimensional analysis of variance (continued)

Combinations that cause synergy				Number of projects	Productivity		Productivity ratio
Variable	Level	Variable	Level		Mean	Variance	
Require-ment level for reliability	Extremely high + High	Clarity of objectives and priorities	Fairly clear + Little clear + Unclear	45	1.197	0.123	3.07
				108	0.711	0.098	
		Clarity of role assignments and each staff member's responsibilities	Fairly clear + Little clear + Unclear	42	1.179	0.165	2.71
				111	0.746	0.115	
		Similar project	Usage	17	1.091	0.176	2.22
				70	0.744	0.114	

Results of two-dimensional analysis of variance (continued)

Combinations that cause synergy				Number of projects	Productivity		Productivity ratio
Variable	Level	Variable	Level		Mean	Variance	
Working space	Levels: c + d (Narrow)	Development framework	No usage	25	1.299	0.125	3.09
				80	0.809	0.132	
		Clarity of objectives and priorities	Fairly clear + Little clear + Unclear	47	1.138	0.183	2.35
				108	0.768	0.107	
Project management tool	Usage	Document-generation tool	No usage	50	1.124	0.117	2.48
				100	0.729	0.106	
		Clarity of objectives and priorities	Fairly clear + Little clear + Unclear	46	1.048	0.138	1.98
				98	0.75	0.122	

Productivity ratio

- Of the 10 combinations, 5 productivity ratios were greater than 3, and 4 were between 2 and 3.

➔Most combinations of these predictor variables affect productivity more than the predictor variables do alone.

- The productivity of a project developing software whose requirement level for security was high in a narrow working space was 3.48 times lower than otherwise.
- The productivity of a project developing software whose requirement level for security was high without a development framework was 3.36 times lower than otherwise.

Axioms

- All synergy effects greatly lower the productivity.
- The following combinations of synergy effects are the most important.

"When developing software required for high security or for high reliability, role assignments and each staff member's responsibilities, objectives and priorities should be very clear, and working space should be broad enough. Such a project has a high possibility of extremely lower productivity. To prevent lower productivity, the usage of a development framework is also important when developing software required for high security".

- Background
- Analysis data
- Data analysis methods
- Analysis results & discussions
  - -- Regression analysis for quantitative variables
  - -- One-dimensional analysis for qualitative variables
  - -- Two-dimensional analysis for qualitative variables
  - -- Comparison to the COCOMO II
- Summary

Comparison of results derived from SEC database and cost drivers of COCOMO II

Overall project: four factors were identified in the SEC database, but they did not correspond to the scale factors/cost drivers in COCOMO II.

Tool usage: four detailed factors were identified, although only one cost driver was identified in COCOMO II.

User's skill levels and commitment: one factor and three possible predictor variables were identified, while no cost driver was found in COCOMO II.

Requirement levels: two factors and 3 possible predictor variables were identified, while 8 cost drivers were found in COCOMO II. Some factors/variables corresponded well to the cost drivers in COCOMO II.

Development staff's skill levels: no variable except for PM skill level was identified as the factor, while five cost drivers were found in COCOMO II.

Comparison of results derived from SEC database and scale factors/cost drivers of COCOMO II (subset)

Category	Results derived from SEC database		Similar-ity**	COCOMO II	
	Qualitative candidate predictor variables	Productivity ratio of typical projects*		Scale factor and cost driver***	Productiv-ity range
Tool usage	Similar project	1.47	～	Precedentedness †	1.33
	Development framework	1.40	～	Use of Software Tools	1.50
	Project management tool	1.63			
	Configuration control tool	#1.56			
	Design support tool	++			
	Document-generation tool	2.21			
	Debug and test tool	1.58			
	Code generator	-			
Requirement levels	Requirement level for reliability	1.85	＝	Required Software Reliability	1.54
	Requirement level for security	2.85	～	Product Complexity	2.38
	Requirement level for performance and efficiency	#1.36	～	Time Constraint	1.63
Development staff's skill level or experience	Staff's business experience	-	～	Application Experience	1.51
	Staff's experience with development platform	#1.43	＝	Platform Experience	1.40
	Staff's experience with languages and tools	-	＝	Language and Tool Experience	1.43

* No mark is significant at 5%. "#" is significant at 5%, but biased. "++" is significant at 10%.

** "=" means " (Nearly) equal to", and "～" means "Similar to.

*** † means a scale factor in COCOMO II formula.

(blue) : Increase productivity (yellow) : Decrease productivity

Summary

- The SEC database keeps more than 3000 enterprise software projects with many more quantitative and qualitative variables than other databases open to the world.
- However, it seems difficult to effectively use this database for constructing cost models because of abundant missing values.
- As the first step, the factors that affect the productivity of developing enterprise software were clarified by analyzing the data for 523 projects. Several interesting results were obtained.

Summary (analysis results)

(1) Productivity was proportional to the root of the fifth power of test case density, and that of fault density in addition to the root of the eighth power of size.

(2) Thirteen predictor variables alone were identified. The most effective top four were

- requirement level for security,
- document-generation tool usage,
- clarity of objectives and priorities, and
- requirement level for reliability.

Summary (analysis results) (continued)

(3) The productivity ratios of typical projects of most factors were less than 2.0 except for two factors: 2.85 of requirement level for security and 2.21 of document-generation tool.

(4) The productivity of projects managed by a person with a high skill level was lower than that of projects managed by a person with a low skill level. One of the reasons was PMs with high skill level tended to manage software projects developing large-scale software with high requirement levels for reliability and security. One of their duties is to run much more test cases per FP to maintain software quality than a PM with low skill level.

Summary (research results) (Continued)

(5) Two-dimensional analysis of variance revealed 10 synergy effects. The following two cases were the most notable.

- The productivity of a project developing software required for high security in a narrow working space was 3.48 times lower than otherwise.
- The productivity of a project developing software required for high security without a development framework was 3.36 times lower than otherwise.

Future challenges

- Multiple regression analysis
- Path analysis
- Quality construction model

  …

Thank you