Productivity Analysis of Japanese Enterprise Software Development Projects

Masateru Tsunoda
Nara Institute of Science and Technology
Kansai Science City, 630-0192 Japan
masate-t@is.naist.jp

Akito Monden
Nara Institute of Science and Technology
Kansai Science City, 630-0192 Japan
akito-m@is.naist.jp

Hiroshi Yadohisa
Doshisha University
1-3 Miyakodani, Tatara, Kyotanabe,
Kyoto, 610-0394 Japan
hyadohis@mail.doshisha.ac.jp

Nahomi Kikuchi
Software Engineering Center
Information-technology Promotion Agency
2-28-8, Hon-Komagome,
Bunkyo-ku, Tokyo, 113-6591 Japan
n-kiku@ipa.go.jp

Ken-ichi Matsumoto
Nara Institute of Science and Technology
Kansai Science City, 630-0192 Japan
matumoto@is.naist.jp

ABSTRACT
To clarify the relation between controllable attributes of a software development and its productivity, this paper experimentally analyzed a software project repository (SEC repository), consisting of 253 enterprise software development projects in Japanese companies, established by Software Engineering Center (SEC), Information-technology Promotion Agency, Japan. In the experiment, as controllable attributes, we focused on the outsourcing ratio of a software project, defined as an effort outsourced to subcontract companies divided by a whole development effort, and on the effort allocation balance among development phases. Our major findings include both larger outsourcing ratio and smaller upstream process effort leads to worse productivity.

Categories and Subject Descriptors
D.2.9 [**Software Engineering**]: Management – *Cost estimation, productivity*; K.6.1 [**Management of Computing and Information Systems**]: Project and People Management – *Strategic information systems planning;*

General Terms: Management, measurement, economics

Keywords
Software productivity, subcontract, upstream process, custom software, software project repository

1. INTRODUCTION
Estimation of software development effort is required throughout a software development lifecycle to set up, evaluate and revise a project plan including resource allocation and scheduling. Software productivity is one of the key factors in drawing up the early effort estimation in a project, although the productivity greatly varies according to project attributes such as business area, application type, programming language, team size and experience, etc [8]. Past researches have shown that some attributes do affect the productivity [2][7][8][9]. Unfortunately, many of these attributes, e.g. business area and programming language, are usually not controllable by a software development company.

This paper focuses on the project attributes that are controllable by a software development company. Through the analysis of the year 2005 version of the SEC repository consisting of 253 samples of enterprise software development in Japanese companies, this paper seeks to clarify the relation between project attributes and software development productivity. The repository has been developed and maintained by Software Engineering Center (SEC), Information-technology Promotion Agency, Japan [10].

In the analysis, as controllable project attributes, this paper focuses on the outsourcing ratio of a software project, and on the effort allocation balance among development phases. While high outsourcing ratio is one of the major distinctive properties of Japanese enterprise software development, there is no past researches focused on the relation between outsourcing ratio and productivity of software development in Japan.

The reminder of this paper first describes details of the SEC repository we used (Section 2). Next, describes analyses we conducted to clarify the relation between project attributes and productivity (Section 3). Afterward, related works are described (Section 4); and in the end, conclusions and future topics will be shown (Section 5).

2. THE SEC REPOSITORY
The year 2005 version of the SEC repository consists of 1009 software projects held in 15 Japanese companies. These projects are custom enterprise software, which is the majority in Japan [4][5]. Each project is characterized by about 400 attributes (metrics); however, 87.7% of them were unrecorded on average. Since the project size is a necessary attribute to calculate the productivity, we excluded projects having missing value in the function

MSR'06, May 22–23, 2006, Shanghai, China.

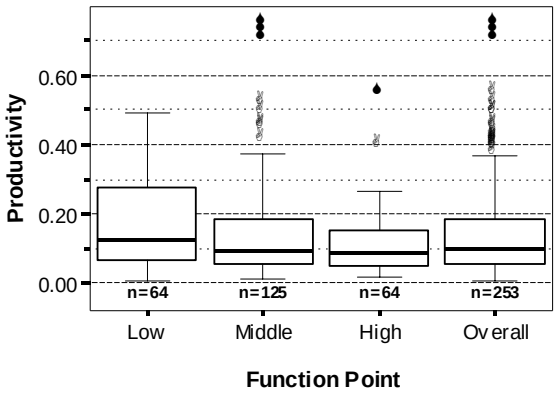

Figure 1. Boxplots of productivity for project groups classified by FP

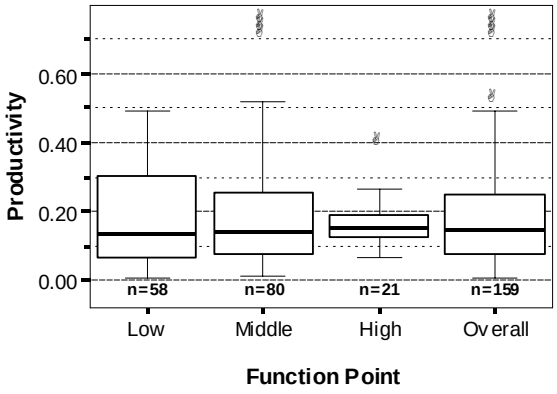

Figure 2. Boxplots of productivity for project groups classified by FP (Outsourcing ratio is zero)

Table 1. P-values for project groups classified by FP

Group 1	Group 2	All	Outsourcing ratio is zero
Low	Middle	0.072	0.789
Low	High	*0.017*	0.610
Middle	High	0.296	0.744

point (FP) attribute from our analysis. Also, we excluded maintenance projects and enhancement projects, and focused to new development projects and re-development projects because maintenance/enhancement processes are often very different from new/re-development processes. As a result, 253 projects were selected for the analysis. All these projects were waterfall process development.

Productivity is defined as FP divided by (total) development effort (person-hour). FP methods consist of IFPUG (34%), SPR (26%), NESMA (3%), others (31%), and missing (5%). Productivity of the upper quartile project is about 3.3 times larger than the lower quartile project. Because of biased distribution of productivity and other metrics, we used the median of metrics instead of the average in our analysis. We also used non-parametric tests in the analysis.

In the analysis, we mainly focused to two factors to analyze productivity. One is the outsourcing ratio, defined as an effort outsourced to subcontract companies divided by a whole development effort. The other is the upstream process ratio, defined as the sum of requirement analysis process effort and design process effort divided by a whole development effort. In many cases, these two factors are controllable by a software development company.

3. PRODUCTIVITY ANALYSIS

There are several factors that may influence software productivity. One of the most considerable factors is the project size [1][3] since development processes (e.g. human resource allocation) vary according to the project size. Therefore, we also focus to the FP, which is one of project size metrics, in addition to the outsourcing ratio and the upstream process ratio.

3.1 Project Size

Before focusing on the outsourcing ratio and the upstream process ratio, we analyzed the influence of FP (project size), which is the most basic factor in productivity analyses.

In this analysis projects were classified into three groups by their FP. Projects whose FPs are equal to or less than lower quartile were classified as the “low” group. Projects whose FPs are equal to or greater than upper quartile were classified into the “high” group. The rest projects were classified into the “middle” group.

To visually explore the difference in productivity among three groups, boxplots were used. Figure 1 shows boxplots of productivity for the three project groups (FP=low, FP=middle, and FP=middle) plus “overall” group for whole projects. Circles indicate outliers, and stars indicate extreme outliers. Upper quartile of FP is about 4.1 times larger than lower quartile. The figure shows there are several high productivity projects in the low group. Median of productivity of the low group is about 1.4 times larger than the high group. Median of productivity of the middle group is only about 1.1 times larger than the high group. Using Mann-Whitney U test, we confirmed that the difference of productivity is statistically significant at the 0.05 level between the low group and the high group. P-values are shown in the column “All” of Table 1. The column “Group 1” and the column “Group 2” indicate paired group when testing. Italic indicates p-value < 0.05.

Next we excluded the influence of outsourcing ratio since there was a high correlation between FP and outsourcing ratio (Spearman's rank correlation was 0.51). Figure 2 shows boxplots for projects whose outsourcing ratios were zero. Median of productivity is almost same among the three groups. The differences in productivity among the three groups are not statistically significant at all. P-values are shown in the column “Outsourcing ratio is zero” of Table 1.

Above all, the SEC repository showed that the project size alone do not directly influence the productivity. However, it can be said that large size projects have high outsourcing ratio; and thus, it indirectly leads to worse productivity.

3.2 Outsourcing Ratio

Figure 3 shows boxplots of productivity for three project groups classified by the outsourcing ratio. (Note that the number of projects of the low group is different from the high group because there are many projects whose outsourcing ratios are zero and all

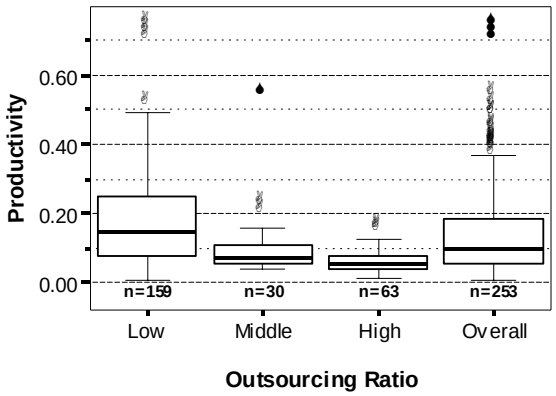

Figure 3. Boxplots of productivity for project groups classified by outsourcing ratio

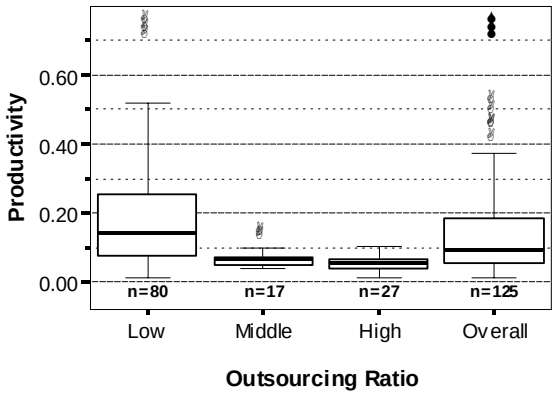

Figure 4. Boxplots of productivity for project groups classified by outsourcing ratio (FP is middle)

Table 2. P-values for project groups classified by outsourcing ratio

Group 1	Group 2	All	FP is middle
Low	Middle	*3.2E-04*	*3.9E-04*
Low	High	*8.8E-14*	*1.8E-07*
Middle	High	*0.002*	0.104

of them are classified as the lower quartile group.) The figure shows there are high productivity projects in the low group. Median of productivity of the low group is about 2.6 times larger than the high group. The differences of productivity are statistically significant at the 0.05 level among the three groups. P-values are shown in the column "All" of Table 2.

Next we excluded the influence of FP. As described in Section 3.1, high FP projects tend to have high outsourcing ratio. Figure 4 shows boxplots for projects whose FPs are in the "middle" group. Figure 4 shows similar tendency to Figure 3. The differences in productivity are statistically significant at the 0.05 level between the low group and other groups. P-values are shown in the column "FP is middle" of Table 2.

These results suggest that lower outsourcing ratio projects have higher productivity. We consider that higher outsourcing ratio introduces communication overhead between companies. However, these results do not mean that the outsourcing ratio must be suppressed because the development cost would be higher if a company stopped outsourcing. (Unfortunately, because the SEC repository does not record cost factors, we were not able to analyze relations between cost and outsourcing ratio.) In addition, if the software size is too large to be developed in-house, a company needs to outsource a part of software development. Anyway, a company must be aware of the trade-offs between increase of effort and decrease of cost.

3.3 Upstream Process Ratio

Figure 5 shows boxplots of productivity for three project groups classified by upstream process ratio. The figure shows there are high productivity projects in the high group. Median of productivity of the high group is about 1.8 times larger than the low group. The differences in productivity are statistically significant at the 0.05 level between the high group and other groups. P-vales are shown in the column "All" of Table 3.

Next we excluded the influences of FP and outsourcing ratio. Figure 6 shows boxplots for projects whose FPs are in the "middle" group. The differences of productivity are statistically significant at the 0.05 level between the high group and other groups. P-values are shown in the column "Function point is middle" of Table 3. Similarly, Figure 7 is boxplots of projects whose outsourcing ratio is zero. The differences in productivity are not statistically significant at the 0.05 level among the three groups. P-values are shown in the column "Outsourcing ratio is zero" of Table 3. Although there was no strong significance among these groups, significance p=0.083 (< 0.10) was observed between middle and high group.

From these results, high upstream process ratio alone has some influence to the productivity. This suggests that high upstream process ratio may avoid additional effort (reworks) on downstream processes.

4. RELATED WORK

Maxwell et al. [8] and Premraj et al. [9] analyzed the influence of the business sector type on productivity, using Finnish software development project dataset collected by Software Technology Transfer Finland (STTF). Lokan et al. [6] also showed productivity analysis focused on the business sector using a dataset of International Software Benchmarking Standards Group (ISBSG). In these researches, projects in the manufacturing sector have the highest productivity, and projects in the banking/Insurance sector have the lowest productivity. Same tendency was observed in the SEC repository.

Blackburn et al. [2] analyzed the influence of requirement analysis process ratio on the productivity. They found that requirement analysis process ratio leads to higher productivity of coding process. Our finding about the influence of upstream ratio follows their result.

5. CONCLUSIONS

We analyzed the influences of outsourcing ratio and upstream process ratio on the development productivity of Japanese enterprise software development projects. We found that both lower outsourcing ratio projects and higher upstream process ratio pro-

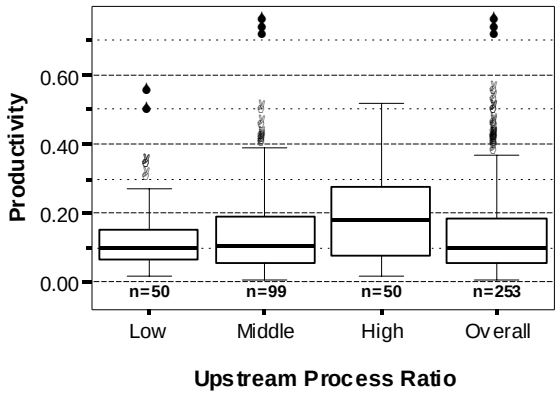

Figure 5. Boxplots of productivity for project groups classified by upstream process ratio

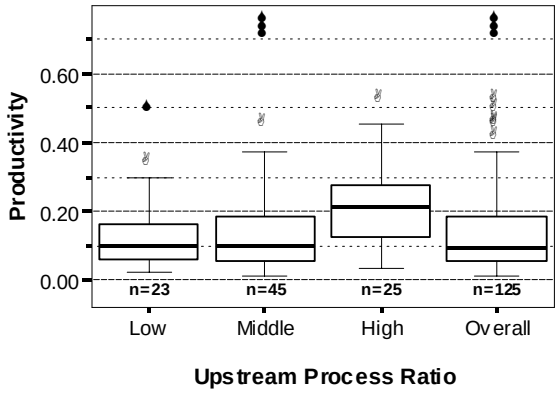

Figure 6. Boxplots of productivity for project groups classified by upstream process ratio (FP is middle)

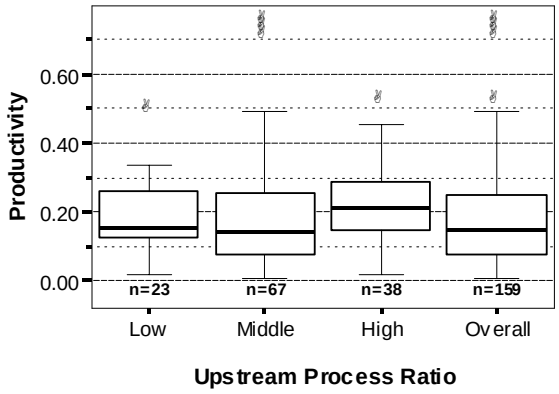

Figure 7. Boxplots of productivity for project groups classified by upstream process ratio (Outsourcing ratio is zero)

Table 3. P-values for project groups classified by upstream process ratio

Group 1	Group 2	All	FP is middle	Outsourcing ratio is zero
Low	Middle	0.819	0.974	0.595
Low	High	*0.009*	*0.016*	0.222
Middle	High	*0.012*	*0.011*	0.083

jects have significantly higher productivity. Our future work is to analyze the influences of other factors thoroughly.

6. ACKNOWLEDGMENTS

The authors would like to thank Software Engineering Center, Information-technology Promotion Agency, Japan for offering the SEC Repository. This work is supported by the EASE (Empirical Approach to Software Engineering) project of the Comprehensive Development of e-Society Foundation Software program of the Ministry of Education, Culture, Sports, Science and Technology of Japan.

7. REFERENCES

[1] Banker, R., Chang, H., and Kemerer, C. Evidence on economies of scale in software development. *Information and Software Technology, 36*, 5 (1994), 275-282.

[2] Blackburn, J., Scudder, G., and Wassenhove. L. Improving Speed and Productivity of Software Development: A Global Survey of Software Developers. IEEE Trans. on Software Eng., 22, 12 (1996), 875-885.

[3] Boehm, B. *Software Engineering Economics*. Prentice Hall, 1981.

[4] Cusumano, M., MacCormack, A., Kemerer, C., and Crandall, B. Software Development Worldwide: The State of the Practice. *IEEE Software, 20*, 6 (2003), 28-34.

[5] Duvall, L. A study of software management: The state of practice in the United States and Japan. *Journal of Systems and Software, 31*, 2 (1995), 109-124.

[6] Lokan, C., Wright, T., Hill, P., and Stringer, M. Organizational Benchmarking Using the ISBSG Data Repository. *IEEE Software, 18*, 5 (2001), 26-32.

[7] Maxwell, K., Wassenhove, L., and Dutta, S. Software Development Productivity of European Space, Military, and Industrial Applications. *IEEE Trans. on Software Eng., 22*, 10 (1996), 706-718.

[8] Maxwell, K., and Forselius, P. Benchmarking Software-Development Productivity. *IEEE Software, 17*, 1 (2000), 80-88.

[9] Premraj, R., Shepperd, M., Kitchenham, B., and Forselius, P. An Empirical Analysis of Software Productivity over Time. In *Proceedings of 11th IEEE International Software Metrics Symposium (METRICS'05)* (Como, Italy, September), 2005, 37.

[10] Software Engineering Center, Information-technology Promotion Agency, Japan *The 2005 White Paper on Software Development Projects (in Japanese)*. Nikkei Business Publications, 2005.