Productivity of Software Projects by Business Sector: An Empirical Analysis of Trends

Rahul Premraj, Bhekisipho Twala and Carolyn Mair
Bournemouth University
Poole House, Talbot Campus,
Fern Barrow, Poole,
Dorset - BH12 5BB, United Kingdom
{rpremraj, btwala, cmair}@bournemouth.ac.uk

Pekka Forselius*
Software Technology Transfer Finland Oy
Tekniikantie 14,
2.floor FIN- 02150 Espoo,
Finland
pekka.forselius@kolumbus.fi

Abstract

Software organisations continuously adopt new technologies and improve their business processes to increase their productivity and hence, be abreast with or ahead of their competitors. This emphasises the importance of investigation into the results of such improvements over productivity over a period of time. In this paper we compare productivity of projects that commenced in and pre-1994 with those that commenced in and post-1997 (the latter a larger sample). Our results reveal very interesting trends in which companies within some business sectors have become very productive due to their past investments, while others are revamping their business processes to help increase productivity in the future. The analysis also reveals important features that play a crucial role in determining productivity within individual business sectors.

Keywords: *software metrics, software productivity, benchmarking*

1 Introduction

In this paper we aim to investigate whether software organisations within business sectors have become more productive over a period of time. We expect there to be a change owing to recent technological developments and improvement in business processes [3].

The measure we use for comparison is Productivity ($ProjectSize/Effort$). In competitive markets productivity is used for self-assessment by organisations, and measuring their performance against their competitors is crucial for survival. Well established metrics protocols facilitate crucial and accurate business decisions [9].

In the current analysis, we attempt to calculate productivity of recent software projects within business sectors and compare our results with those from Maxwell and Forselius [10]. Their analysis was based on a previous version of the data set used in this analysis, and hence comprised older projects. This comparison will help reveal an expected trend in productivity within business sectors and we attempt to explain the causes for the same.

The organisation of the paper is as follows: Section 2 describes the data set used in the analysis and the data editing performed on it. Section 3 elaborates upon the design of our analysis. Then, we explore our results from the analysis in sections 4, 5 and 6. Lastly we summarise our results in section 7 as conclusions and in section 8 we indicate possible directions for future research.

2 Data Set Description

The data set used for analysis in this paper is a subset (a result of data editing) of the latest version of the Experience data set [4]. This section briefly introduces the Experience Pro initiative and the data editing performed to make the data set suitable for our analysis.

2.1 Experience Data Set

The Experience data set is a result of commercially driven initiatives by Software Technology Transfer Finland (STTF). The on-going efforts on the project have resulted in a comprehensive data set which includes software projects that commenced as far back as 1978 and those completed as recently as 2004. In its current form the data set comprises 622 projects. Organisations pay an annual fee to gain access to the data via a tool called Experience Pro. The same

*Pekka Forselius also represents a registered association titled Finnish Software Measurement Association (FiSMA) and undertakes research at the University of Jyväskylä, Finland.

tool can be used to submit companies' own project data for which they are entitled to a discount on the annual fee. The use of the tool for project data submission ensures standardisation of features included. Also the project data are carefully assessed at STTF by experts before being added to the data base. More information about Experience Pro is available at the website [4].

2.2 Data Editing on the Data Set

We first removed all projects in the data set that commenced prior to 1997. This was necessary for two reasons. Firstly, we aim to compare the productivity of business sectors against their performance in the period covered by projects included in Maxwell and Forselius [10]. Secondly, the co-author (Pekka Forselius), who is also the *Repository Manager* at STTF, recommended the use of projects that commenced in and post-1997 for added reliability.

The data set comprises a mixture of new development and maintenance projects. Due to the differences in the nature of the two types of projects, they have been described using some features that are different from each other. Hence, in the current analysis we only consider analysing the productivity of new development and enhancement projects, while we removed maintenance projects which are characterised as having the value $MT22$ (Modifying Maintenance) in the $SituationAnalysisModel$ feature.

Further, to ensure that the results were more robust, it was vital to remove outliers from our analysis. This was achieved by calculating $DeliveryRate$ ($effort/ProjSize$ or $Productivity^{-1}$), which was chosen because it was easier to set cut-off points using the resultant values in comparison to those of $Productivity$. We removed projects that had values < 1 and > 30 of $DeliveryRate$ because these projects had unusually high or low delivery rate. It is difficult to ascertain the reasons behind such unusual productivity. Projects with delivery rate < 1 may have possibly made substantial amounts of reuse of code, while those with deliver rate > 30 must have encountered several obstacles during implementation. Due to this vagueness, we chose to remove 8 abnormal projects from our analysis (4 projects having $DeliveryRate < 1$ and the other 4 > 30).

Some of the values of the $ProductivityFactors$ (T variables) had the value -1 which symbolises that they had not been recorded. A total of 102 such values were replaced by the average value of 3 as approved by the data collecting organisation. In addition to these, there were 42 blank values belonging to 2 cases. These have been confirmed as erroneous by the data collecting organisation, but in our analysis, we have changed these values to the average value of 3 in order to retain the maximum number of projects.

In the latest version of the Experience data set (March 2004), projects from 14 unique business sectors have been included. We removed projects from 8 business sectors from our analysis for a variety of reasons. Firstly, we aim to compare the performance within the 5 business sectors covered in Maxwell and Forselius [10]. Secondly, 7 of these business sectors had 5 or fewer projects and hence, constrain us from making reliable deductions. Lastly, a set of projects whose business sector was classified as 'Not Defined' was removed, since without more information, our judgement on their performance from the analysis may be fallible. One option would be to group all these projects as 'Other' and continue our analysis, thus maximising the number of projects. However, due to marked differences in the nature of these projects arising from environmental specificity the analysis might not be very fruitful.

3 Experimental Setup

To investigate the trend in productivity within business sectors over a period of time, we undertook three separate experiments.

In the first experiment we construct a simple linear regression model between effort and project size to analyse the degree of relationship between to the two variables. Next, we calculated productivity of software projects according to business sector and compared them with their corresponding values in Maxwell *et al.*'s analysis and explain the causes behind observed the trend. In the third analysis, we look into projects from individual business sectors exclusively and investigate the most important features that affect their productivity.

In the second and third analyses, we perform an ANOVA [12] to investigate which features strongly influence productivity for the whole data set and within each business sector. With the results, we are able to find out if different factors contribute significantly towards productivity in comparison to those that affected productivity in Maxwell's analysis. Tables summarizing features that were found to be significant (at $\alpha = 5\%$) along with the respective variance explained[1] [5] by the feature are shown under the appropriate following sections.

4 Introductory Regression Analysis

As our dependent variable $Productivity$ is a function of effort and project size, this section examines the relationship between the two variables in our data set. A scatter plot of the two variables is shown in figure 1. It shows that the relationship between effort and project size is non-linear.

[1] The proportion of variance explained in ANOVA is calculated by dividing the sum of squares between groups by the sum of squares total. This ratio represents the proportion of variance explained.

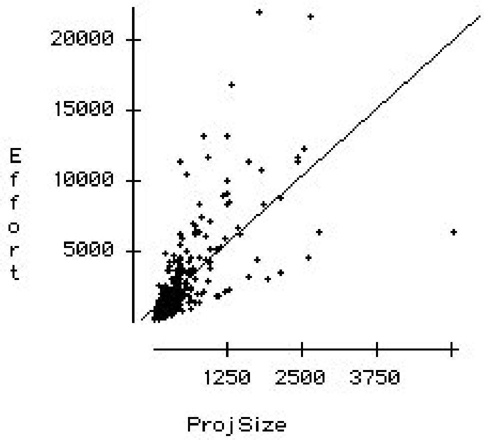

Figure 1. Scatter Plot of Effort Vs. ProjSize (n = 305)

Also, a greater degree of scatter is noticeable for projects requiring larger effort.

Boehm [1] (and others, e.g. [2]) suggested that effort is related to size as follows:

$$Effort = \alpha\ (ProjectSize)^{\beta} \tag{1}$$

Equation 1 takes into account the non-normality in the distribution of effort and project size and hence, is a multiplicative model. Taking the log on both sides of equation 1, we derive the familiar regression equation as follows:

$$ln(Effort) = \alpha + \beta\ ln(ProjSize) \tag{2}$$

The parameters α and β are determined using linear regression on the natural logs of the two variables. Typically, due to diseconomies of scale, β is > 1 [1]. Thus, larger projects need more effort (or are less productive) in comparison to smaller projects. However, some researchers have reported β values < 1. For more information on this area, we direct you to reference [7].

Due to non-linear relationship of the variables in our data set (figure 1), we derive the relationship between effort and project size by taking the natural log transformation of both variables and then using Boehm's model (eq. 1) to derive values of α and β. Equations 2 and 1 take the form of equations 3 and 4 respectively for our data set. The mean predicted value of $ln(effort)$ is:

$$ln(Effort) = 2.2473 + 0.89352\ (ln(ProjSize)) \tag{3}$$

or the traditional multiplicative form is:

$$Effort = 9.462\ (ProjectSize)^{0.893523} \tag{4}$$

From equation 4, we can see that for every unit increase in project size, effort increases linearly since the exponent of $ProjSize$ drives its value to be very close to itself ($x^1 = x$). Importantly, we observe that the dependent variable ($Effort$) moves in the same direction as the independent variable ($ProjSize$). The value of R^2 between $ln(Effort)$ and $ln(ProjSize)$ was 65%, i.e. 65% of the variation in effort is explained by size, but other factors play a role in the determination of effort and hence, need to be explored. The relationship is visualised in figure 2 which also suggests that the relationship between the two variables is linear.

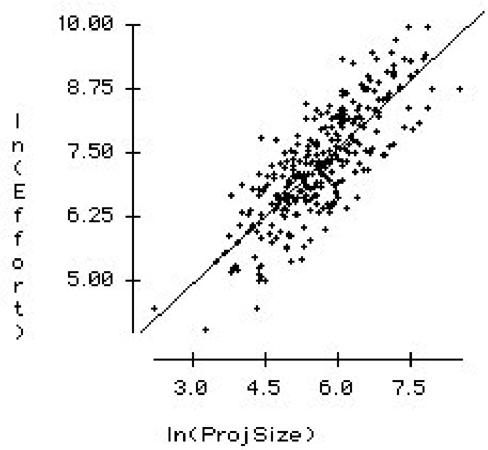

Figure 2. Scatter Plot of ln(Effort) Vs. ln(ProjSize) (n = 305)

The plot of residuals and the predicted values in figure 3 indicates that the assumption of normality holds and that the model is heteroscedastic. Figures 2 and 3 may also suggest that project size is a crucial factor for cost or effort estimation and knowing its approximate value at an early stage in the project may help in more accurate allocation of resources to the project.

5 Inter-Business Sector Analysis

A productivity analysis was conducted using the setup as described in section 3. Table 5 lists all the features that were found to be significant at the 5% level.

Our analysis shows that Banking and Insurance sectors are the most productive business sectors. This is in striking contrast to Maxwell's analysis in which these two sectors were the least productive. The reason for poor performance

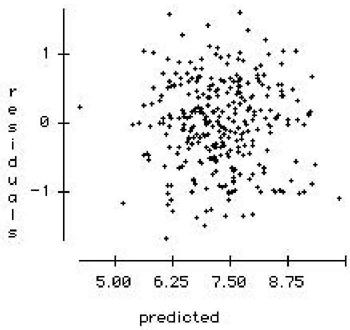

Figure 3. Scatter Plot of Residuals and Predicted values using Equation 3

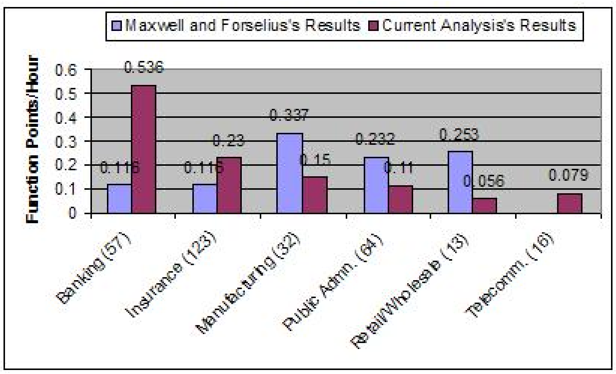

Figure 4. Comparison of Average Productivity ($Effort/ProjSize$) across Business Sectors

Variable	Obs.	Variance Exp. (%)
BusinessSector	305	85.1
Company	305	77.4
Devp.Lang.	305	41.1
CaseTools1	50	36.9
TechnicsMethods1	195	29.6
Proj.Mgmt. Tools1	188	24.5
HardwarePlatform	305	10.3
Product Type	305	7.0
Productivity Factors		
ExpOfProj.Mgmt.	305	5.9
ToolSkillsOfStaff	305	5.8
UsabilityReqs.	305	5.5
NumOfStakeHolders	305	4.9
ImpactOfStandards	305	4.5
AvailOfITStaff	305	4.1
PorabilityReqs.	305	4.1
AnalysisSkillsStaff	305	4.1
ImpactOfMethods	305	3.9
MaintainabilityReqs.	305	3.1

Table 1. Productivity Variance in All Sectors due to Significant ($\alpha = 5\%$) Individual Factors

of Banking and Insurance sectors was partially explained in [10] to be due to in-house development of software projects. In our case, the results are explained by the projects undertaken in the Banking sector in the early 90s. During this period, the sector invested heavily in time and money to develop very integrated and complex 2 or 3-tier architecture systems. This has paid off in recent years since it has made new development and enhancement projects more productive than in the past. This is also partially true in the Insurance sector, who have invested considerably into development process improvement. (This partially explains the reason for the large number of recorded Insurance sector projects recorded.)

The productivity of Manufacturing and Wholesale/Retail sectors has decreased in recent times due to their move from separate applications to integrated ERP and other systems. Since the sectors are concentrating upon such large and complex ERP systems, their productivity has reduced by nearly 50% for the Manufacturing sector and as much as 78% for the Wholesale and retail sectors.

The nearly 50% slide in productivity of Manufacturing sector organisations can be explained by recent efforts to integrate their large Manufacturing databases. In the early 90s, separate and smaller applications were being developed (or mostly such projects that were recorded into the Experience data set in its earlier versions). Manufacturing sector systems are not based on commercial ERP products, but they are now migrating several old public registers from mainframe to 3-tier architecture to make it possible to develop all kinds of web-services in future. Perhaps, in the next 5 years we could see the Manufacturing sector following the footsteps of the Banking and Insurance sectors.

The Telecom sector was not available in the version of the Experience data set during Maxwell's analysis. However, its relatively poor performance is currently explained by the fact that the sector is developing more commercial software packages than the other sectors. A more credible analysis of this business sector can be undertaken in due course when more projects are added to the Experience data set.

In this paper we only cover the understanding of the few significant variables from table 5. Most interestingly, in contrast to [10] in which Company was the most important variable to explain productivity variance, we now find Business Sector to be the highest determinant of variance in productivity. This is perhaps a result of competition in

Variable	Obs.	Variance Exp. (%)
ProjectType	57	71.3
HardwarePlatform	57	25.2
LevelOfChangeMgmt.	57	20.2
ImpactOfMethods	57	14.6
WorkBreakDownStruct	57	2.7

Table 2. Productivity Variance in Banking due to Significant ($\alpha = 5\%$) Individual Factors

the business sector in which all competing firms are abreast in technology and hence their productivity levels are comparable. This finding warrants more research in comparing the benefits of using company-specific and multi-company databases, for e.g. [13] and [8]. Similarly [10], we also find Development Language, Hardware Platform and Product Type to be crucial determinants of productivity. For productivity factors, we found Experience of Project Management (5.9%) to be the highest determinant of variance in Productivity. In [10], Requirements Volatility was the highest at 16%. Further, our analysis showed that Tools Skills of Staff and Usability Requirements are other important determinants of productivity.

6 Intra-Business Sector Analysis

This section examines productivity within business sectors exclusively and points out possible causes for the observed results.

6.1 Banking

From Table 2, we can see that Project type in the Banking sector is the most crucial determinant of Productivity. Grouped by Project Type, the average productivity of enhancement projects in the Banking sector is $0.74 FP/hr$, while the same for new development tailored projects is $0.48 FP/hr$. Next, Hardware platform is an important determinant which accounts for 25.2% of the variance in productivity. This is not surprising because new projects were undertaken to revamp the system's architecture, while other projects must have been simultaneously undertaken to keep current systems in working order.

The level of change management measures the stability and predictability of the functional requirements in the project. Figure 5 illustrates the effect that this feature has on productivity. The boxplot $\#4$ represents only 8 projects, but its median lies considerably higher than that of its peers. This shows that during the course of the project, fewer changes made to specifications that were laid down in the beginning of the project increases its productivity considerably.

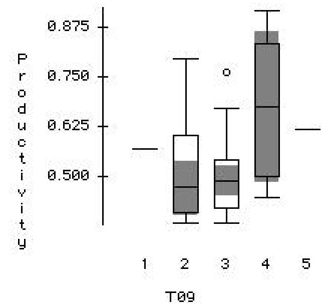

Figure 5. Comparison of Average Producivity ($Effort/ProjSize$) across Business Sectors

Variable	Obs.	Variance Exp. (%)
ProjectType	123	75.6
HardwarePlatform	123	11.3
MaturityOfSWDevp.Process	123	11.3

Table 3. Productivity Variance in Insurance due to Significant ($\alpha = 5\%$) Individual Factors

6.2 Insurance

The Insurance sector (table 6.2) has evolved on the same lines as the Banking sector. With recent change in the system's architecture, it is not surprising to see new systems being developed to cater for the new architecture, while numerous enhancement projects have been undertaken to keep the current systems running. Hence, the project type accounts for as much as 75.58% of variation in productivity. Hardware platform is a crucial determinant of productivity for the very same reason.

Lastly, our analysis revealed the importance of Maturity of Software Development Process (figure 6) that accounted for the same amount of variation (11.3%) in productivity as the choice of hardware platform, i.e. 11.3%. This feature measures the stability and conformance of the software development processes and lifecycle related activities in the project. From the medians of boxplots 2, 3 and 4, we observe that increasing trend in productivity with more stability and early defect detection.

6.3 Manufacturing

In the Manufacturing sector, Development Language was found to be the most influential factor accounting for 58% of variance in productivity. From the 32 projects in the sector, 21 used C++ with an average productivity of $0.15 FP/hour$. Most of the other development languages

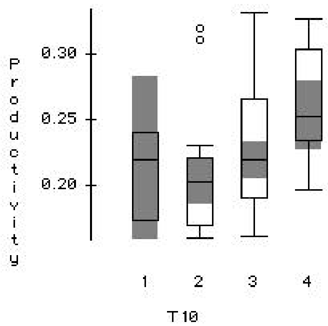

Figure 6. Effect of Maturity in Software Development Process on Productivity in the Insurance Sector

Variable	Obs.	Variance Exp. (%)
DevpLanguage	32	54.8
WorkBreakDownStruct	32	46.0
HardwarePlatform	32	43.3
ProjMgmtTools1	26	35.4
ProjectType	32	32.5
Company	32	22.7

Table 4. Productivity Variance in Manufacturing due to Significant ($\alpha = 5\%$) Individual Factors

included by 1 or 2 projects in the Manufacturing sector. The authors in [10] found Hardware Platform to be the most significant feature accounting for productivity variance. Interestingly, this is the only sector in which we found the Company to be a significant feature affecting productivity.

6.4 Public Administration

Techniques and Methods were found to be the most influencing feature in the Public Administration sector. This may perhaps be explained by the remodelling taking place in the sector recently. Besides other factors that were found to be significant in the previous sectors, we also found Reliability Requirements and Availability of IT Staff to be significant. This may be due to the vast amount of human effort involved in handling the massive data bases involved in Public Administration sector. Oddly, we found that the average productivity within the sector decreases with more staff dedicated to a single or few projects (figure 7).

Variable	Obs.	Variance Exp. (%)
TechnicsMethods1	28	74.4
DevpLanguage	64	66.5
ProjectType	64	54.6
HardwarPatoform	64	53.4
DevpMode1	21	40.4
ReliabilityReqs	64	16.3
AvailITStaff	64	14.0

Table 5. Productivity Variance in Public Sector due to Significant ($\alpha = 5\%$) Individual Factors

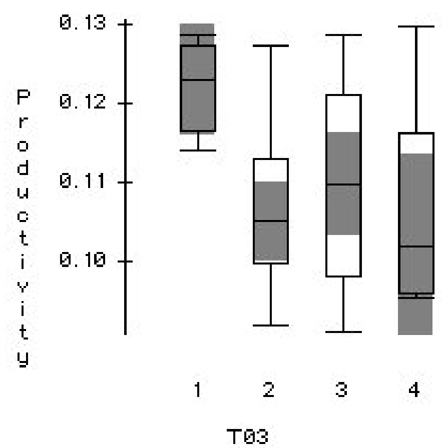

Figure 7. Effect of Availability of IT Staff on Productivity in the Public Administration Sector

6.5 Wholesale and Retail

Besides Techniques and Methods, we found Involvement of Customer Representatives to significantly affect productivity. Figure 8 fails to show us any conclusive trends due to lack of projects. However, the significance of the feature is not surprising due to their shift to ERP systems when developers and customers may need to be in frequent contact for the success of such large and complex systems.

6.6 Telecommunications

Upon viewing the contents of the significant features, we discovered that Techniques and Methods1 explains such high variance due to lack of data points. Hence, we skip this feature to move on to Project Type. The telecommunication projects have 3 different types of new developments including enhancements, new development product and new development tailored. Figure 9 shows the distribution of pro-

Variable	Obs.	Variance Exp. (%)
TechnicsMethods1	10	94.1
InvolvementCustReps	13	71.8
ProjectType	13	59.7
EfficiencyReqs	13	57.0

Table 6. Productivity Variance in Wholesale and Retail due to Significant ($\alpha = 5\%$) Individual Factors

Variable	Obs.	Variance Exp. (%)
TechnicsMethods1	8	99.3
ProjectType	16	95.1
MaturitySWDevpProcess	16	53.6
ToolSkillsStaff	16	49.4

Table 7. Productivity Variance in Telecommunications due to Significant ($\alpha = 5\%$) Individual Factors

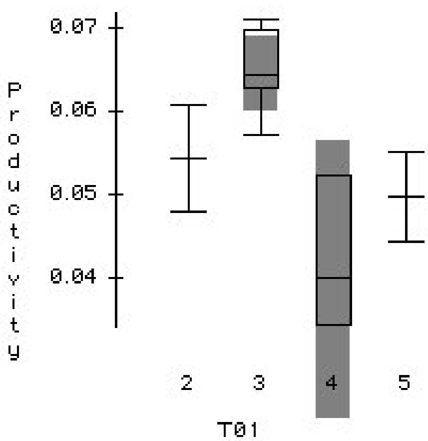

Figure 8. Effect of Involvement of Customer Representatives on Productivity in the Wholesale and Retail Sector

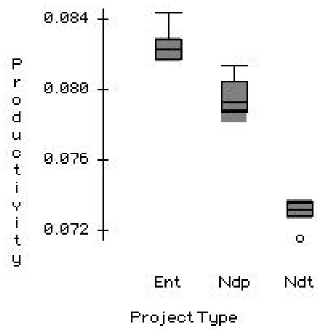

Figure 9. Effect of Project Type on Productivity in the Telecommunications Sector

ductivity within each project type. The first boxplot representing enhancement projects shows that they are most productive. The other two types of projects are considerably less productive. It is important to note that on the whole, in its current state, the telecommunications industry is very less productive.

Unsurprisingly, in a technologically oriented sector, the maturity of software development process and tool skills of staff have an important role to play. Unfortunately, due to few data points, we are unable to establish any certain trends in productivity. This will make an interesting study in the future so as to understand and compare the productivity of technical and non-technical business sectors.

7 Conclusions

Estimation of software development productivity is a difficult and key problem in software engineering. The credibility of the model depends upon the data, definitions and assumptions that were used to derive the model. This empirical investigation has revealed interesting results and may incite more research in the same direction. We observed that the performance of pre-1995 and post-1996 software projects within business sectors have undergone many changes for a variety of reasons. We found that overall, the productivity of software projects has gone up from $0.177 FunctionPoints/Hour$ (for pre-1995 projects) to $0.247 FunctionPoints/Hour$ (for post-1996 projects).

In section 3, we noticed that the Banking and Insurance sectors are now the most productive sectors between those that were examined in our analysis. This was mainly due to heavy investments made in revamping their architecture in the 1990s which has paid off with more productivity in recent projects. Current efforts made by the Manufacturing sector organisations to revamp their systems may also help increase their productivity in future projects.

Within individual business sectors, we understand that features pertaining to management and planning now play an important role in determining productivity. For e.g. in Banking, impact of methods had strong influence on productivity, whereas the Insurance sector productivity was influenced by the maturity of software development process. Perhaps, in the future with more standardisation and technological advances, we may see a drift away from technical factors to management factors that strongly influence productivity.

An important observation to be noted is that within the individual business sectors, nearly all were significantly affected by Hardware Platform and Project Type (e.g. new development tailored or enhancements). The latter can be partially explained due to revamp efforts in some of the sectors. The significance of the former explains the degree of complexity involved when developing software for different platforms. From our data set, we analysed that the productivity levels for networks and multi-platform projects is much lower than for main-frame and mid-range platforms. However, future projects may help establish more concrete trends in many dimensions.

8 Future Research Directions

As in many data sets, the relationship between effort and project size was non-normal. Current research shies away from exploring methods to handle such data and resorts to the standard method of transforming the variables into their natural logs. Discovering new statistical techniques for this purpose may be well valued in the software engineering community and others.

Investigating the relation between productivity and number of defects in the project will make an interesting study to see if speedy projects, perhaps due to efficiency or pressure, are referred back to the development company for fixing more and more bugs. Unfortunately, due to non-availability of data at the moment, it was not possible for us to explore this area.

Software productivity has been exploited for the purpose of effort estimation previously by Jørgensen *et al.* [6] and Premraj *et al.* [11]. The information in this paper may perhaps help take such research a step further. By understanding how and why productivity varies within a project, methods can be developed to account for the variation and perhaps result in more accurate estimates.

Acknowledgements

The authors would like to thank all anonymous referees for their feedback on the draft version of this paper.

References

[1] B. W. Boehm. *Software Engineering Economics*. Prentice-Hall, 1981.

[2] S. Conte, H. Dunsmore, and V. Shen. *Software Engineering Metrics and Models*. The Benjamin/Cummings Publishing Company, Melno Park, California, 1986.

[3] M. Cusumano, A. MacCormack, C. F. Kemerer, and B. Crandall. Software development worldwide: The state of the practice. *IEEE Software*, 20(6):28–34, 2003.

[4] Experience Pro Internet Site. www.sttf.fi/eng/products/experience/indexexperience.htm.

[5] D. N. Gujariti. *Basic Econometrics*. Economics Series. McGraw-Hill Inc., Singapore, 1988.

[6] M. Jørgensen, U. Indahl, and D. Sjoberg. Software effort estimation by analogy and regression toward the mean. *Journal of Systems and Software*, 2003.

[7] B. A. Kitchenham. The question of scale economies in software - why cannot researchers agree? *Information and Software Technology*, 44(1):13–24, January 2000.

[8] K. Maxwell. Benchmarking software development productivity: Statistical analysis by business sector. In R. Kursters, A. Cowderoy, F. J. Heemstra, and J. Trienkens, editors, *Project Control for 2000 and Beyond*, Maastricht, 1998. Shaker Publishing B.V.

[9] K. Maxwell. *Applied Statistics for Software Managers*. Prentice Hall, New Jersey, 2002.

[10] K. Maxwell and P. Forselius. Benchmarking software development productivity. *IEEE Software*, January-February:80–88, 2000.

[11] R. Premraj, M. Shepperd, and M. Cartwright. Meta-data to guide retrieval in cbr for software cost prediction. In B. Lees, editor, *8th UK Workshop on Case-based Reasoning*, pages 26–37, Cambridge, UK, 2003.

[12] A. Rutherford. *Introducing ANOVA and ANVOVA: A GLM Approach*. Sage Publications Ltd, 2000.

[13] I. Wieczorek and M. Ruhe. How valuable is company-specific data compared to multi-company data for software cost estimation? In *Metrics*, page 237. IEEE, 2002.