Exploring the Relation Between Effort and Duration in Software Engineering Projects

Serge Oligny, Pierre Bourque, Alain Abran, Bertrand Fournier
Université du Québec à Montréal
C.P. 8888, Succ. Centre-Ville
Montréal, Québec, Canada H3C 3P8
Email: {oligny.serge, bourque.pierre, abran.alain, fournier.bertrand}@uqam.ca

Abstract

This paper presents a confirmatory analysis of empirical models that predict software engineering project duration based on project effort, based on a more recent and much larger sample. The models are based on the analysis of project data provided by release 4 of the International Software Benchmarking Standards Group (ISBSG) repository. Duration models are built for subsets of projects using personal computer, mid-range and mainframe development platforms.

Keywords

Software Engineering, Project Duration Models, Project Scheduling Models, ISBSG, International Software Benchmarking Standards Group.

1. Introduction and Context

Many parametric models based on project effort have been proposed in the literature [4], [7] to predict the duration of software development projects**.** A wide selection of tools that incorporate software duration estimation models can also be found on the market [3], [8].

In spite of affirmations from vendors [6], studies have shown mixed results regarding the accuracy of these parametric duration models [2], [7].

However, one must note that the data sets used to conduct these studies are often small and not too recent. Seven of the eight studies discussed in [7] were published over ten years ago. Ferens and Daly [2] present a study published in 1990 and discuss three other assessments published in 1989 or 1990.

Of the 21 empirical duration models analyzed in [7], 12 of them are derived from samples containing 20 data points or fewer. Ferens and Daly [2] discuss the results of three other studies that assess parametric duration models. These studies are conducted on samples containing one, eight and twelve projects while their own study is based on a sample of twenty-one defense projects which cannot be described because of "contractor-proprietary data restrictions".

This does not mean that these types of models should not be further investigated by researchers or applied by practitioners. It does imply though, that, software duration estimation is a somewhat complex problem and that applying these models correctly requires much expertise and commitment [11].

Within this context, this paper proposes to re-examine the relationship between overall project effort and duration:

a) based on a relatively large sample of data, using release 4 of ISBSG database [5], containing data on 396 software projects completed between 1989 and 1996;

b) in the perspective of obtaining a "first order" or overall estimate of project duration rather than a detailed breakdown of project duration at the phases or activity level;

c) to verify if the development platform used, as an indicator of the general approach to project development, bears an impact on project duration;

d) to compare the overall characteristics of the empirical relationship based on sample describe in a) above with the overall characteristics of other known and comparable models.

These goals will be reached by analyzing the ISBSG sample (section 2). Three empirical models, for personal computer, mid-range and mainframe development platform, will be presented in section 3. A brief conclusion and further opportunities for research are presented in section 4.

2. Analysis of the Data Sample

Among the 396 projects in release 4 of the ISBSG repository, projects showing the following characteristics were selected:

- No reasonable doubt as to data validity; that is, the ISBSG has not flagged the selected project as

having uncertain data and has retained it for its own analyses;

- Known effort ;
- Known duration;
- Known platform;

312 projects satisfied all these criteria. Basic descriptive statistics are shown in Table 1 and a scatter plot of the data is shown in Figure1.

	Duration (D) in calendar-months	*Effort (E) in person-hours*
N	312	312
Minimum value	1.0	10
Maximum value	78.0	106 480
Mean value	10.5	5 933
Standard deviation	9.0	12 169
Median	8.0	2 228

Table 1 - Descriptive Statistics

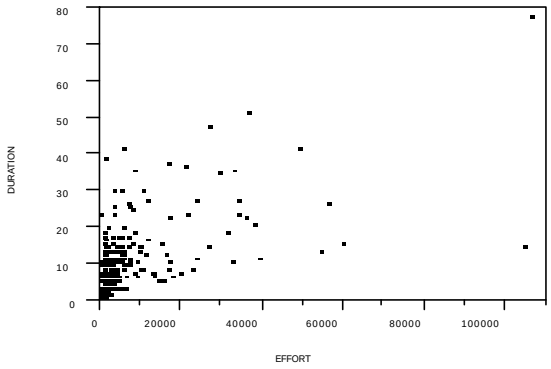

Figure 1 - Scatter plot of project effort vs. duration (n= 312)

Figure 1 suggests that the relation between effort (E) and duration (D) is non-linear. Furthermore, a greater degree of scatter is observable for projects requiring larger effort.

Taken as is, such data distributions offer weak support for linear regression. It was thus deemed appropriate to work on a mathematical transformation of the data. A Log transform was therefore applied successfully in the sense that the Log(duration) and the Log(effort) data both follow a normal distribution. Test of normality confirm this at the 95% level.

As shown in Figure 2, the Log transformations renders the relation more linear and reduces the extent of scatter for high levels of effort. The Log-transformed values will therefore be used to derive an empirical model linking project effort to project duration.

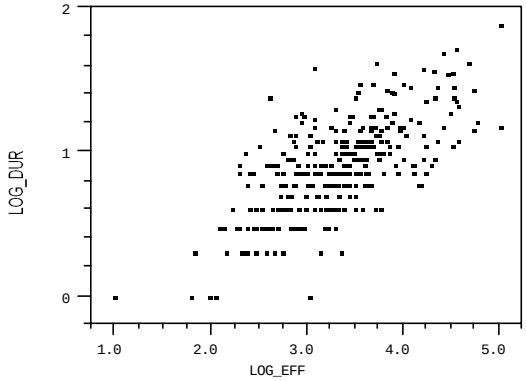

Figure 2 – Scatter plot of project's Log(effort) vs. Log(duration)

3. Deriving duration models for each development platform

Following a recommendation in [10], duration models are built for more homogeneous subsets of data rather than using the entire data sample. From the 312 projects meeting the criteria presented in section 2, 208 (67%) developed software for a mainframe platform, 65 (21%) for a mid-range platform and 39 (12%) for a personal computer platform.

We begin by presenting detailed results for the mainframe platform projects. An identical analysis was performed for projects developing software for the other two platforms and the results are presented subsequently.

A linear regression was performed on the Log-transformed values of effort and duration for the 208 mainframe projects. Selected results from this regression and from the regressions for mid-range and personal computer data are presented in Table 2.

Analysis of residuals against both Log(effort) and predicted Log(duration) shows that the residuals are randomly distributed over the range of the independent variable (Log(effort)) and that the

RESULTS	**MR**	**PC**	**MF**
N:	65	39	208
R^2:	0.434	0.140	0.522
F(1, (n-1)):	48.324	5.970	224.865
Prob > F :	0.0001	0.0194	0.0001
Log(E) coefficient:	0.360	0.201	0.366
Standard error of Log(E):	0.052	0.082	0.024
Constant:	-0.261	0.287	-0.339

Table 2 - Selected results from the regressions (mid-range, personal computer, and mainframe platform projects)

variance of the residuals is constant over the range of the dependent variable (predicted Log(duration)). It is thus held that the resulting linear regression model is acceptable.

The empirical model linking project effort and duration for the mainframe platform projects can thus be characterized by the following equations.

The mean predicted value of Log(duration) is:

$$\text{Log (D)} = (0.366 * \text{Log (effort)}) - 0.339 \qquad (1)$$

or, in the traditional multiplicative form:

$$D = 0.458 * \text{Effort}^{0.366} \qquad (2)$$

Based on this sample, this empirically derived model explains slightly over half the variance in project duration.

Regression analyses were conducted separately for the 65 mid-range projects and for the 39 personal computer projects. The four hypotheses of linear regression were satisfied in both cases. For the mid-range platform projects, the resulting multiplicative equation is:

$$D = 0.548 * E^{0.360} \qquad (3)$$

The data for the personal computer platform projects is much more scattered and the relation is weaker. The resulting multiplicative equation for this data set is:

$$D = 1.936 * E^{0.201} \qquad (4)$$

Given equations (2, 3 and 4) describing the relation between effort and duration for three classes of projects, it is legitimate to ask whether or not there is a significant difference between the models. Figure 3 shows that the relation for projects developing software for mainframe and mid-range platforms appears to be quite similar, while the relation for projects developing software for personal computer platforms seems to be different.

To answer this question in a more formal way, we used the statistical approach described by Neder *et al.* in [9]. This approach is similar to regression modeling using indicator variables and interaction terms to test the equality of equations for different classes of a qualitative variable.

Based on this analysis, identity of the regression equations (at the 95% level) is not rejected for mid-range and mainframe platform data. There is, however, a significant difference (Student's t_{306} = 2.385, p = 0.0177) when comparing the identity of the slopes for projects developing software for a personal computer (PC) and mainframe platform.

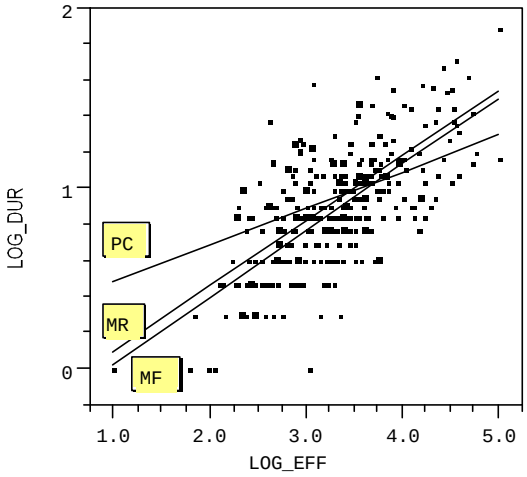

Figure 3 - Regression lines of Log(duration) by Log(effort) for each development platform

4. Conclusion and Further research

Estimation of software project duration is a difficult and key problem in software engineering which can only be solved with an inherently white-box approach. The credibility of an estimate depends entirely on the transparency of the method, data, definitions and assumptions that were used to derive this estimate.

Based upon the recent and relatively large release 4 ISBSG data set, this paper showed that:

- The relation between effort and duration is non-linear and the exponential term is in the range of 0.3 to 0.4 which is comparable to other duration models including COCOMO [1]. Interestingly enough, this is in spite of the roughly 20 years that separate the projects used to develop COCOMO from the ones in release 4 of the ISBSG data set. Secondly, the release 4 of the ISBSG data set to a very large extent comprises business systems, while most other studies are based on defense-oriented projects. The size of the samples of our analyses are much more recent and larger than those of previous studies.
- The duration models developed in this paper show that project managers can derive a «first order» estimate of project duration from a valid estimate of project effort. However, other scheduling methods must be used in conjunction with duration models such as those developed in this paper since much of the variance is not explained by the equations developed in this paper.
- There is a significant difference in the relation between effort and duration based on the development platform of the project. In this regard, project developing software for personal

computers are different from projects developing software for either a mid-range platform or a mainframe platform. However, there is no significant difference in the relation between effort and duration for projects developing software for mid-range and mainframe platforms

In the perspective of these conclusions, many axes of research could be pursued. The size of the ISBSG sample permits the subdivision of the data set into other homogeneous data sets using characteristics like project type (new developments or major enhancements), or business area (banking, manufacturing, telecommunications, etc.). Better context-sensitive project duration models, explaining more variance, for instance, could possibly be derived from these subsets of data. The usage of non-linear models and the incorporation of other variables in the modeling process should also be explored.

A second topic to study would be created by turning this problem around, that is, by developing and assessing empirical models that estimate an acceptable product size and project effort from a given project duration. This is a very worthwhile topic because project duration is often fixed before the project begins in many organizations due to time-to-market or implementation considerations.

Acknowledgments

This research was carried out at the Software Engineering Management Research Laboratory at the Université du Québec à Montréal. The laboratory is supported through a partnership with Bell Canada. Additional funding is provided by the Natural Sciences and Engineering Research Council of Canada. The opinions expressed in this paper are solely those of the authors.

5. References

1. B.W. Boehm. *Software Engineering Economics*, Prentice Hall, 1981.
2. D.V. Ferens, B. A.Daly. *A Comparison of Software Scheduling Methods*. In: Reifer D, ed. Software Management. 4th ed. Washington: IEEE Computer Society Press, 1993.
3. A.E. Giles, D. Barney. “Metrics Tools: Software Cost Estimation”. *Crosstalk The Journal of Defense Software Engineering*, 8(6) 1995.
4. F.J. Heemstra. “Software cost estimation”. *Information and Software Technology* 34(10): 627-639, 1992.
5. ISBSG. *Worldwide Software Development - The Benchmark*. Victoria, Australia International Software Benchmarking Standards Group, 1997. Can be ordered at http://www. isbsg.org.au
6. C. Jones. Determining software schedules. *Computer* 28(2):73-75 1995.
7. B.A. Kitchenham. “Empirical studies of assumptions that underlie software cost-estimation models”. *Information and Software Technology* 34(4): 211-218 1992.
8. O. Mendes, A. Abran and P. Bourque. *Function Point Tool Market Survey*. Université du Québec à Montréal, 1996 Available at http://www.lrgl. uqam.ca/publications/pdf/204.pdf
9. J. Neder, W. Wasserman and M. Kutner, *Applied Linear Regression Model*. Second Edition ed, Richard D. Irwin Inc., Homewood, IL, 1989.
10. S. Oligny, P. Bourque and A. Abran. “An Empirical Assessment of Project Duration Models in Software Engineering”. In: *The Eight European Software Control and Metrics Conference (ESCOM'97)*, Berlin Germany, May 1997.
11. R. E.Park, W. B. Goethert and J.T. Webb. *Software Cost and Schedule Estimating: A Process Improvement Initiative*. Pittsburgh, PA Software Engineering Institute, 1994