Available online at www.sciencedirect.com
SCIENCE DIRECT®
Information and Software Technology 48 (2006) 297–301
INFORMATION AND SOFTWARE TECHNOLOGY
www.elsevier.com/locate/infsof

How large are software cost overruns? A review of the 1994 CHAOS report

Magne Jørgensen*, Kjetil Moløkken-Østvold

Simula Research Laboratory, P.O. Box 134, NO-1325 Lysaker, Norway

Received 11 October 2004; revised 28 June 2005; accepted 5 July 2005
Available online 3 October 2005

Abstract

The Standish Group reported in their 1994 CHAOS report that the average cost overrun of software projects was as high as 189%. This figure for cost overrun is referred to frequently by scientific researchers, software process improvement consultants, and government advisors. In this paper, we review the validity of the Standish Group's 1994 cost overrun results. Our review is based on a comparison of the 189% cost overrun figure with the cost overrun figures reported in other cost estimation surveys, and an examination of the Standish Group's survey design and analysis methods. We find that the figure reported by the Standish Group is much higher than those reported in similar estimation surveys and that there may be severe problems with the survey design and methods of analysis, e.g. the population sampling method may be strongly biased towards 'failure projects'. We conclude that the figure of 189% for cost overruns is probably much too high to represent typical software projects in the 1990s and that a continued use of that figure as a reference point for estimation accuracy may lead to poor decision making and hinder progress in estimation practices.

Keywords: Software cost estimation; Research method; Survey review

1. Introduction

The Standish Group (www.standishgroup.com) claims that the results of their CHAOS research, i.e. their large-scaled surveys conducted in 1994, 1996, 1998, 2000 and 2002 [1], are the most widely quoted statistics in the IT industry. This may very well be true. Quoted with particular frequency are the results described in the 1994 CHAOS report, the results of which have been used in several project reviews, research studies and governmental reports, e.g. the US PITAC 1999 report (President's Information Technology Advisory Committee Report to the President; see www.hpcc.gov/pitac/report).

An important result of the 1994 CHAOS research is the reported 189% average cost overrun of so-called challenged projects, i.e. projects not on time, on cost, and with all specified functionality. The question we try to answer in this paper is whether or not it is wise trust the cost overrun figures described in the 1994 CHAOS report. This question is especially pertinent, since the figures are seen to represent a 'software development crisis'. Our analysis is based on an examination of the quality of the Standish Group's estimation survey and a comparison of the cost overrun results with those of the other estimation surveys completed in the same period.

The paper is organized as follows: Section 2 discusses and exemplifies the use of the 1994 cost overrun figures reported by the Standish Group. Section 3 examines how the CHAOS report interprets cost overrun. Section 4 compares the CHAOS report results with those of similar software cost estimation surveys. Section 5 discusses possible reasons for the disparity in results between the 1994 CHAOS report and similar surveys, which reasons focus on the quality of survey design. Section 6 describes lessons learned from the examination of the 1994 CHAOS report. Finally, Section 7 summarizes the paper.

* Corresponding author. Tel.: +47 924 333 55; fax: +47 67 82 82 01.
E-mail address: magnej@simula.no (M. Jørgensen).

0950-5849/$ - see front matter © 2005 Elsevier Ltd All rights reserved.
doi:10.1016/j.infsof.2005.07.002

2. Users of the cost overrun figures

Among the users of the cost overrun figures in the 1994 CHAOS report are scientific researchers, software process

improvement groups, and government advisors. The main use seems to be to support arguments for more research, better estimation processes and improved project management methods. These are all laudable goals, well motivated by the 'software development crisis' implied by a 189% average cost overrun. Unfortunately, there are several examples of situations where the 189% result may have hindered progress. The following three real-world examples illustrate this.

Example 1. A project had a 146% cost overrun, i.e. the actual cost was about 2.5 times the estimated cost. A report on the project's performance stated that the cost overrun was not that bad, because it was better than the industry average of 189%. We have found several examples of this type of use.

Example 2. A consultancy company claimed to be in the elite class of software development companies, based on a comparison of its own high cost overrun figures with the 189% cost overrun. The company's cost mean cost overrun was high, but lower than 189%.

Example 3. A recent UK study of software projects reports an average cost overrun of 18% (The state of IT project management in the UK 2002–2003, www.computerweekly.com/pmsurveyresults/surveyresults.pdf). Here, an adjusted version of the 189% cost overrun figure from the 1994 CHAOS report is used to support an argument for an enormous improvement in estimation performance since 1994. Readers of that study may get the impression that the improvement of cost estimation processes does not require a great deal of systematic work and focus, whereas, in fact, as the lack of progress reported in [3] suggests, it is a very demanding and difficult enterprise.

3. The meaning of '189% average cost overrun'

Before we compare the CHAOS report results with those of other studies it is important to clarify what the expression '189% average cost overrun' actually means. This turned out to be more difficult than expected. In fact, we were unable to find in the CHAOS reports an explicit definition of the cost overrun measure that is applied. Only informal descriptions were presented; and even they were inconsistent with each other. The following quotations provide typical examples of how the CHAOS reports describe the 189% average cost overrun:

- '… 52.7% of projects will overrun their initial cost estimates by 189%', page 41 in [2].
- 'The average cost overruns for combined challenged and cancelled projects is 189%.', page 42 in [2].
- '… 52.7% of projects will cost 189% of their original estimates', page 14 in [1].

52.7% is identical to the percentage of so-called challenged projects. Yet even the definition of challenged projects is not easy to interpret. It is defined in [1] as 'The project is completed and operational but over-budget, over the time estimate, and offers fewer features and functions than originally specified.' The problem here is the use of 'and' instead of 'or', combined with the following definition of successful projects [1]: 'The project is completed on-time and on-budget, with all features and functions as initially specified.' Consider a project that is on-time, and on-budget, but not with all specified functionality. Is this project to be categorized as challenged or successful? If it is categorized as challenged, this is not consistent with the provided definition of challenged projects; while if it is categorized as successful, it fails to meet the criteria given for a successful project.

To determine whether we were the only ones confused by these descriptions, we conducted a simple survey of the actual use of the 1994 CHAOS report cost overrun figure. We examined 50 randomly sampled web documents applying the search term: ((Standish Group) AND (189% OR 89%)) and the search engine www.yahoo.com. We found the following:

- 50% of the documents described the result as '189% cost overrun', 40% as '189% of original estimate', and 10% as '89% cost overrun'. '189% of original estimate' and '89% cost overrun' seem to reflect the same understanding of the result, i.e. we found two different interpretations of cost overrun that were used with almost the same frequency.
- 70% of the documents related the result to '53% of the projects' (without explicitly pointing out that this 53% referred to challenged projects only), 16% to 'all projects', 8% to 'challenged and cancelled projects', and 6% explicitly pointed out that the average cost overrun is based on 'challenged projects' only.

Generally, in recent reports and press releases the Standish Group seems to apply the interpretation '189% average cost overrun of challenged projects'; see, for example, the press release March 2003 (www.standishgroup.com/press/article.php?id=2). This means that many, perhaps the majority, of users interpret the results differently from the Standish Group.

4. A comparison with other cost estimation accuracy studies

We conducted a systematic search for estimation surveys [3]. This systematic search was based on (i) a manual search in about 100 software development, project management and quality assurance journals (see www.simula.no\BESTweb or contact the authors for a list of the journals included) and (ii) a variety of searches for studies applying terms

Table 1
Cost overrun surveys

Study	Jenkins [4]	Phan [5]	Bergeron [6]
Year	1984	1988	1992
Respondents	23 software organizations	191 software projects	89 software projects
Country of respondents	USA	USA	Canada
Average cost overrun	34%	33%	33%

related to 'estimation survey' in the library database *Inspec*. We identified three surveys of cost estimation accuracy for the relevant period and countries; see Table 1. These surveys suggest an average cost overrun in the range of about 30%, i.e. values far from a 189% cost overrun. The surveys in Table 1 have all been subjected to a scientific review of research method and results, in contrast to the Standish Group CHAOS reports. For a more detailed description of the search process, these surveys and other cost estimation surveys, see [3].

These values are not directly comparable with those in the CHAOS reports. The studies in Table 1 include successful as well as challenged projects, as opposed to the CHAOS report, where the successful projects are excluded from the cost overrun calculation. However, the proportion of successful projects in the 1994 CHAOS report was only 16% and cannot explain the huge difference in the results. The question is therefore: are there other differences between the three studies in Table 1 and the CHAOS 1994 survey that can explain the huge difference in average cost overruns? Are there, for example, reasons to believe that the cost accuracy performance was so much worse in 1994 than in the period 1984–1992, or that the three studies in Table 1 are biased towards cost overruns that are too low? We can find no such explanations for the difference in results.

Interestingly, the Standish Group's CHAOS surveys for the years 1996, 1998, 2000, and 2002 report strongly decreasing numbers, i.e. 142, 69, 45, and 43% average cost overrun. Adjusted for differences in how cost overrun is measured, we find that the figures for 2000 and 2002 correspond well with the average cost overrun of about 30% in the studies in Table 1, i.e. it seems as if it is mainly the early (1994, 1996 and 1998) CHAOS report cost overrun figures that are unusual. The strong decrease in average cost overrun, as measured by the Standish Group, is a reason to doubt the research method in itself. Are we to believe, for example, that the average cost overrun improved from 142 to 69% in only two years?

5. Why are the cost overruns in the 1994 CHAOS report so high?

To investigate reasons for the high 1994 cost overrun number, we examined the design of the survey. This design was described poorly in the CHAOS reports. In particular, we were interested in how the projects to be included were selected and how 'cost overrun' was defined. We asked the Standish Group about these issues. Their response to our survey method questions was that providing this type of information would be like giving away their business free of charge, and we received no response on how to interpret 'cost accuracy'. This unwillingness to reveal research method and measurement definitions would have been an unacceptable response in an academic context, but is, in our experience, not uncommon in commercial companies conducting research studies.

This lack of transparency concerning research method and measurement definition leaves us with no choice but to speculate about potential reasons for unusual results and potential design weaknesses. We have identified the following potential reasons to explain the '189% cost overrun' reported in the 1994 CHAOS research report:

- Non-random sampling of projects. Unusual results are sometimes caused by non-random, strongly biased samples. A close reading of version 3.0 of the CHAOS report provides some support for this explanation. On page 13 the selection process of the 1994—study is described as follows: "We then called and mailed a number of confidential surveys to a random sample of top IT executives, asking them to share failure stories. During September and October of that year, we collected the majority of the 365 surveys we needed to publish the CHAOS research." Notice the request for *failure stories*! The decreasing average cost overrun numbers of more recent CHAOS research may therefore be a consequence of an increasingly more representative (and less failure-story related) selection of projects.
- Incorrect interpretation of own results. Recalculations of the average cost overrun based on the Standish Group's 1994 distribution of cost overrun per overrun category results in a cost overrun close to 89%, but far from 189%. There seems, consequently, to be an inconsistency between the two presentations (average cost overrun and cost overrun distribution) of the overrun data. When the Standish Group presents the 1994 study as revealing a 189% instead of 89% cost overrun, it may have been misled by its own confusing description of the results. However, even a reduction from 189 to 89% does not lead to results on the level of the comparable studies.
- No category for cost underrun. In most studies on software cost accuracy there is a proportion of projects with cost underruns. For example, in the UK study on project performance referred to in Section 2, as many as 15% of the projects were completed *under* budget. Even challenged projects may have cost underruns, since they may be challenged only regarding time or functionality. We find no reference to, or description of, treatments of

cost underrun in the Standish Group's reports. It is, therefore, possible that cost underruns are not included as cost underruns, but perhaps as 0% cost overrun.

- Unusual definition of cost overrun. The Standish Group may have used an unusual definition of cost overrun, e.g. the definition may include cost on cancelled projects, as indicated in one of the three informal descriptions of cost overrun.

6. What should we learn from this?

This paper does not *prove* that the 189% average cost accuracy reported in the 1994 CHAOS report is biased and unrepresentative for the situation in 1994. Such a proof would require the availability of information that the Standish Group will not release. It is possible that the results are valid, and merely very difficult to interpret, given the lack of measurement definitions and description of research method. Bearing in mind this cautionary note as to what we are able to establish, we have attempted to provide reasons that cast into doubt the veracity of the 189% average cost overrun value as it is interpreted by most of its users. In particular, we believe that the unusually high average cost accuracy figure, the lack of a description of the research method, and an enormous disparity with the results of studies that have been subjected to a scientific review of research method and results are valid reasons for doubting that figure. As software cost estimation researchers, we (the authors of this paper) and many others have applied the 1994 CHAOS Report cost overrun figures uncritically to motivate several studies, e.g. in [7]. We believe, therefore, that there are lessons to be learned:

Lesson 1. When something does not correspond with one's own experience and other studies, doubt it. A 189% *average* cost overrun, as reported by the CHAOS research, is an extremely high figure in relation to results reported in other studies. Consequently, we should at least require and examine a detailed description of the research method applied, or require independent studies that replicate the result, before believing such extreme results.

Lesson 2. The number of observations, which is higher in the CHAOS report than in the comparable studies, is not always a good indicator of the validity of the results. We should be just as concerned about the selection process as with the number of observations. If the selection process is not described properly, we should doubt the results regardless of the number of observations. Increasing the number of observations does not eliminate error caused by a biased selection process.

Lesson 3. Great care should be taken when interpreting the results of studies that do not define their measures precisely, as is the case with the CHAOS research. The confusion about the interpretation of the 189% average cost overrun (see Section 3) illustrates this.

7. Summary

The Standish Group reported in 1994 that the average cost overrun of software projects was as high as 189%. This figure for cost overruns is still regarded by many as being of great importance and is used to support the view that there is a 'software development crisis'. It has, in addition, been used as input in recent governmental reports, as a benchmark for the estimation performance of recent projects, and to support the claim that there has been an immense improvement in cost estimation performance over the last 10 years. We found several reasons to believe that an average cost overrun of 189% is much too high to reflect the situation in 1994. In particular, the number is not consistent with cost overrun results of other surveys in that period, and there seem to be serious problems with how the Standish Group conducted their research. Unfortunately, the Standish Group provides an incomplete description of how they conducted their studies, regarding, for example, how they selected the projects to be included in the study, and fails to provide a description of how they measure 'cost overrun'. This makes it difficult to evaluate the validity of their 1994 study. Even worse, the lack of a precise definition of the term 'cost overrun' seems to have created much confusion. Many, perhaps the majority, of users interpret the cost overrun results differently from the Standish Group. Our main conclusion is that we should doubt the validity of the 189% average cost overrun reported by the Standish Group in 1994 until such time as they disclose how they measure cost overrun and how they conduct their research. Currently, the validity and comprehensibility of their results are highly questionable and may foster the erroneous impression (a) that the IT industry has improved strongly since 1994 and (b) that even very inaccurate projects are 'better than average'.

References

[1] The Standish Group: Chaos Chronicles Version 3.0. 2003: West Yarmouth, MA.
[2] J. Johnson, CHAOS: the dollar drain of IT project failures, Application Development Trends, (January) (1995) 41–47.
[3] K. Moløkken-Østvold, M. Jørgensen, A review of software surveys on software effort estimation International Symposium on Empirical Software Engineering, Simula Res. Lab. Lysaker Norway, Rome, Italy, 2003 (pp. 223–230).
[4] A.M. Jenkins, J.D. Naumann, J.C. Wetherbe, Empirical investigation of systems development practices and results, Information and Management 7 (2) (1984) 73–82.
[5] D. Phan, D. Vogel, J. Nunamaker, The search for perfect project management, Computerworld (1988) 95–100.
[6] F. Bergeron, J.Y. St-Arnaud, Estimation of information systems development efforts: a pilot study, Information and Management 22 (4) (1992) 239–254.
[7] M. Jørgensen, D.I.K. Sjøberg, Impact of experience on maintenance skills, Journal of Software Maintenance and Evolution: research and practice 14 (2) (2002) 123–146.

Magne Jørgensen is a Professor at the University of Oslo and Simula Research Laboratory. His research interests include software cost estimation, uncertainty assessments in software projects, expert judgment processes, and learning from experience. He received a D. Scient. degree in Informatics from the University of Oslo, Norway in 1994. He has 10 years industry experience as a consultant and software project manager. His publications on software cost estimation and related topics are listed at: www.simula.no. magne.jorgensen@simula.no.

Kjetil Moløkken-Østvold is a Postdoctoral Researcher at Simula Research Laboratory. His research interests include software cost estimation, process improvement and public management of software projects. He received his Masters degree from the University of Oslo in 2002, and his PhD from the University of Oslo in 2004. kjetilmo@simula.no.