How Large Are Software Cost Overruns?
A Review of the 1994 CHAOS Report

Magne Jørgensen[1] and Kjetil Moløkken
Simula Research Laboratory

Abstract: *The Standish Group reported in 1994 that the average cost overrun of software projects was as high as 189%. This cost overrun number is used as input in recent governmental reports on software development and as benchmark for several recent projects' estimation performances. It is therefore important that we can trust this number. More recent cost overrun results reported by the Standish Group and others, however, show much lower cost overrun. Does this mean that software companies have improved their estimation ability strongly the last 10 years? In this paper we argue that there are reasons to doubt the validity of the Standish Group's 1994 cost overrun results and that a continued use of these results may hinder progress.*

Keywords: Software cost estimation, research method, 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, 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 [1]. The results of that report have been used in several recent governmental reports, project reviews, and research studies, e.g., the PITAC 1999 report [2]. An important result from 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 we can trust the cost overrun numbers, implying a 'software crisis', described in the 1994 CHAOS report.

The paper is organized as follows: Section 2 discusses and exemplifies the importance of the correctness of the 1994 cost overrun number reported by the Standish Group, Section 3 compares the CHAOS report results with those of similar software cost estimation surveys, Section 4 discusses potential reasons for differences in results between the 1994 CHAOS report and similar surveys, Section 5 describes lessons learned from the review of the 1994 CHAOS report, and, Section 6 summarizes the paper.

---

[1] Corresponding author: Magne Jørgensen, Simula Research Laboratory, P.O.Box 134, NO-1325 Lysaker, Norway. Telephone: +47 924 333 55. Fax: +47 67 82 82 01. email: magnejorgensen@simula.no.

2 The Importance of the Correctness of 189% Average Cost Overrun Result

Among the users of the cost overrun numbers in the 1994 CHAOS report are scientific researchers, software process improvement groups, and government advisors. The main use seems to be to argue for more research, better estimation processes and improved project management methods. These are all laudable goals, well supported by a 'software cost estimation 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%. There are 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 numbers with the 189% cost overrun number.

Example 3: A recent UK study of software projects [3] report an average cost overrun of 18%. Here an adjusted version of the 189% cost overrun number from the 1994 CHAOS report is used to support an argument for an enormous improvement in estimation performance. Readers of that report may get the impression that the improvement of cost estimation processes does not require a great deal of systematic work and focus.

3 A Comparison with Other Cost Estimation Accuracy Studies

Before we compare the CHAOS Report results with those of other studies it is important to clarify the meaning of '189% average cost overrun'. This turned out to be more difficult than expected. In fact, we were unable to find an explicit definition of the cost overrun measure applied in the CHAOS reports. Only informal, inconsistent descriptions were presented. 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 [4].
    - o **Comment**: 52.7% is identical to the percentage of so-called challenged projects. 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: "*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? Our guess is that it would be categorized as challenged, but this is not consistent with the provided definition of challenged projects.

- "*The average cost overruns for combined challenged and cancelled projects is 189%.*", page 42 in [4].
    - o **Comment**: Here, the cancelled projects are included, i.e., there are two inconsistent descriptions of cost overrun in the same document. The method by which the cost overrun of a cancelled project might be included is not given.
- "… *52.7% of projects will cost 189% of their original estimates*", page 14 in [5].
    - o **Comment**: As we interpret it, if the cost is 189% of an estimate there is an 89% cost overrun, i.e., this is not the same as the first description.

To see 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 number. 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 seams to apply the interpretation "*189% average cost overrun of challenged projects*", see for example the press release March 2003 [6]. This means that many, perhaps the majority, of the users interpret the results differently from the Standish Group.

All surveys of cost estimation accuracy in the relevant period and countries that we were able to find suggest average cost overrun in the range of about 30% (see Table 1), i.e., values far from 189% cost overrun. The surveys in Table 1 have all been subject to scientific review of research method and results, as opposed to the Standish Group CHAOS reports. For a more detailed description of these and other cost estimation surveys, see [7].

Table 1: Cost Overrun Surveys

Study	Jenkins [8]	Phan [9]	Bergeron [10]
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%

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 bias towards too low cost overruns? We can find no such reasons explaining 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 numbers for 2000 and 2002 corresponds 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 numbers 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. For example, do we believe that the average cost overrun improved from 142% to 69% in only two years?

4   Why Are the Cost Overruns in the 1994 CHAOS Report So High?

To investigate reasons for the high 1994 cost overrun number, we asked the Standish Group how they selected the projects to be included in their CHAOS studies and how we should interpret 'cost overrun'. The response to our research method question was that providing this type of information would be like giving away their business for free, and we got 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, as far as we have experienced, not uncommon in commercial companies conducting research studies.

This lack of research method and measurement definition information leaves us with no choice but to speculate about potential reasons. We have identified the following

potential reasons explaining the '189% cost overrun' reported in the 1994 CHAOS research report:

- *Incorrect interpretation of own results*: Re-calculations of the average cost overrun based on the Standish Groups 1994 distribution of cost overrun per overrun category results in a cost overrun close to 89%, i.e., there may be an inconsistency between the two presentations (average cost overrun and cost overrun distribution) of the overrun data. When the Standish Group present the 1994-study as 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 under-run:* In most studies on software cost accuracy there is a proportion of projects with cost under-runs. For example, in the recent UK study on project performance [3] as many as 15% of the projects were completed *ahead* of budget. Even challenged projects may have cost under-runs, since they may be challenged only regarding time or functionality. We find no reference to or description of treatments of cost under-run in the Standish Group's reports. It is therefore possible that cost under-runs are not included as cost under-runs, 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.

- *Non-random sampling of projects*. Unusual results are sometimes caused by non-random samples. A thorough reading of the 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."* 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.

5 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 research study description. Bearing in mind the above cautionary note as to what we are able to establish, we have attempted to provide reasons to *doubt* 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 number and the lack of research method description are valid reasons for doubting that number. As software cost estimation researchers, we (the authors of this paper) and many others have

uncritically applied the 1994 CHAOS Report cost overrun numbers to several studies, e.g., in [11, 12]. We believe therefore that there are lessons to be learned:

Lesson 1: When something does not correspond with own experience and other studies, doubt it. An 189% *average* cost overrun, as reported by the CHAOS research, is an extremely high number in relation to numbers reported in other studies. Consequently, we should require a detailed description of research method applied or independent studies that replicate the result before believing it.

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 properly described, we should doubt the results regardless of the number of observations. Bias is not removed with an increased number of observations.

Lesson 3: Studies that do not precisely define their measures, as is the case with the CHAOS research, should be interpreted carefully. The confusion about the interpretation of the 189% average cost overrun illustrates this. For example, many documents referring to the CHAOS report did not notice that the average cost accuracy only referred to challenged (and maybe cancelled) projects, not to all projects.

6 Summary

The Standish Group reported in 1994 that the average cost overrun of software projects was as high as 189%. This cost overrun number is still of great importance. It is, for example, used as input in recent governmental reports, as benchmark for the estimation performance of recent projects, and to support the claim that there has been an immense improvement in cost estimation performance 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. The number not consistent with cost overrun results of other surveys in that period, and, there may 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, e.g., how they selected the projects to be included in the study, and does not include a description of how the measure 'cost overrun'. This makes it difficult to evaluate the validity of their 1994 study. Even worse, the lack of precise definition of 'cost overrun' seems to have created much confusion. Many, perhaps the majority, of the 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 the Standish Group disclose how they measure cost overrun and how they conduct their research. Currently, the validity and comprehensibility of that number is highly questionable and may create the impression (a) that the IT-industry has improved strongly since 1994 and (b) that even very inaccurate projects are "better than average".

References

1.   Reagan-Cirincione, P., *Improving the accuracy of group judgment: A process intervention combining group facilitation, social judgment analysis, and information technology.* Organizational Behaviour and Human Decision Processes, 1994. **58**(2): p. 246-270.

2.   *The PITAC REport.* www.hpcc.gov/pitac/report, 1999.

3.   *The state of IT project management in the UK 2002-2003.* www.computerweeklyms.com/pmsurveyresults/surveyresults.pdf, 2003.

4.   Johnson, J., *CHAOS: The dollar drain of IT project failures.* Application Development Trends, 1995(January): p. 41-47.

5.   *Chaos Chronicles Version 3.0.* 2003, The Standish Group: West Yarmouth, MA.

6.   Reifer, D.J., *Web development: estimating quick-to-market software.* IEEE Software, 2000. **17**(6): p. 57-64.

7.   Moløkken, K. and M. Jørgensen. *A Review of Surveys on Software Effort Estimation. IEEE International Symposium on Empirical Software Engineering (ISESE 2003).* 2003. Rome, Italy: p. 223-230.

8.   Jenkins, A.M., J.D. Naumann, and J.C. Wetherbe, *Empirical investigation of systems development practices and results.* Information and Management, 1984. **7**(2): p. 73-82.

9.   Phan, D., D. Vogel, and J. Nunamaker, *The search for perfect project management*, in *Computerworld*. 1988. p. 95-100.

10.   Bergeron, F. and J.Y. St-Arnaud, *Estimation of information systems development efforts: A pilot study.* Information & Management, 1992. **22**: p. 239-254.

11.   Jørgensen, M. and D.I.K. Sjøberg, *Impact of experience on maintenance skills.* Journal of Software Maintenance and Evolution: Research and practice, 2002. **14**(2): p. 123-146.

12.   Jørgensen, M. and D.I.K. Sjøberg, *The Impact of Customer Expectation on Software Development Effort Estimates.* Journal of Project Management, 2004. **Vol 22**(4).

Biographies:

Magne Jørgensen is a professor at 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 the Dr. Scient. degree in informatics from the University of Oslo, Norway in 1994. He has 10 years industry experience as consultant and software project manager. His publications on software cost estimation and related topics are listed at: http://www.simula.no/people_one.php?people_id=36. *magne.jorgensen@simula.no*.

Kjetil Moløkken is a Ph.D. student at University of Oslo. His research interests include software cost estimation, expert estimation in groups and surveys of software estimation performance. He received the Cand. Scient. degree from University of Oslo in 2003. kjetilmo@simula.no.