Benchmarking the expected loss of a federal IT portfolio

P. Kampstra and C. Verhoef

VU University Amsterdam, Department of Computer Science,
De Boelelaan 1081a, 1081 HV Amsterdam, The Netherlands

{pkampst,x}@cs.vu.nl

Abstract

After a stream of media attention about failing federal IT investments, their proper IT governance was put on the Dutch political agenda. Pending a congressional hearing, investigations were carried out, among which the publication of data on 73 large and IT intensive federal investments. Based on industry averages that we encapsulated in formulas for general use, we quantified that the expected financial loss for this 5.8 billion euro portfolio is between a conservative 400 million and a potential 2.4 billion euro. We compared the federal portfolio with 6000 projects gathered from the private sector, totaling to 1.7 billion euro. It turned out that the portfolio of 73 large federal projects has an associated risk of abortion that is more than twice as much as for the 6000 private projects.

Keywords: Risk management, Software management, Economics, IT portfolio management.

IT waste has a long international history. After a string of billion dollar failures, in 1996 the Clinger Cohen Act was passed in the United States of America. More recently, a UK report published by the British Computer Society and the Royal Academy of Engineering [9] warned that both the private and public sector are spending billions of pounds every year on failed IT projects. The report recommended better project management, risk management, adoption of best-practices, professionalism and education in the IT sector. A recent review estimated that information system failures in Europe adds up to about 140 billion dollar in any given year, while the USA is reckoned to lose 150 billion per year [3].

Also, in the Netherlands, the amount of wasted millions on IT has gained quite some attention in the Dutch media. After inquiries from the parliament, the government released a portfolio of 73 large government IT projects. The portfolio contains limited data. We will use industry averages, our own data, and the published portfolio to illustrate how to conduct a best-effort risk estimation that requires little input data.

We will do the risk estimation by using benchmark data that is publicly available. Next to that, we use our own data to compare the federal portfolio with data from the private sector. For instance, we will present a loss-benchmark: for a group of aborted projects, 92% of their aggregated original budgets is spent at abortion time. The risk calculations are available on the internet as an Excel spreadsheet[1] that we instantiated with the published portfolio of large Dutch federal IT projects. This spreadsheet is reusable for other IT portfolios, where a distinction between in-house, outsourced, and military projects is incorporated.

[1] See http://www.cs.vu.nl/~x/fed/fed.xls

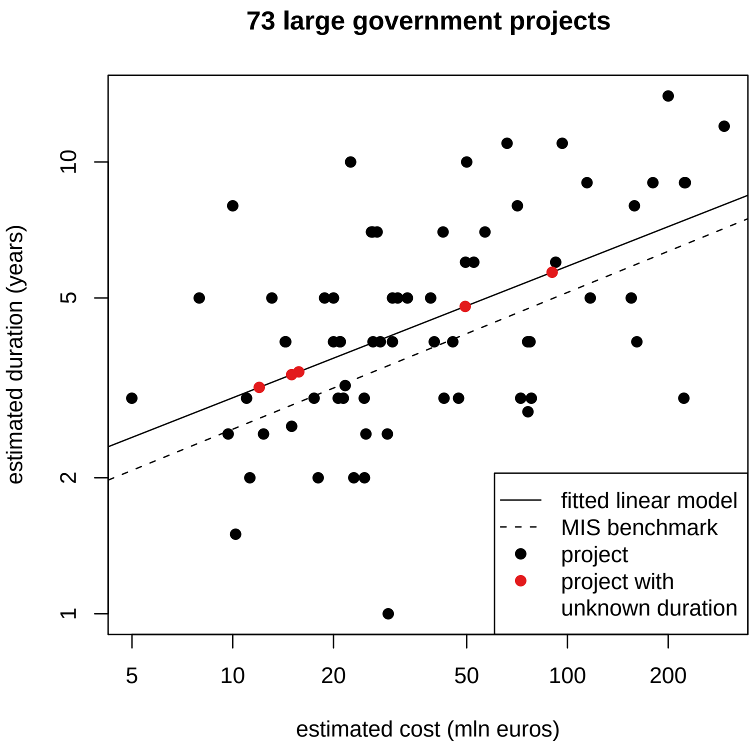

Figure 1: A log-log plot between duration and cost for the federal portfolio. The dotted line is a benchmark as described in the first panel.

The federal portfolio

The portfolio contains information on the total investment of a project, the time-frame, and the percentage of IT cost. There are 73 projects costing 5841 million euro altogether with an IT component of 4175 million. All large projects had to be reported, at least when their size exceeded 20 million euro or their IT component was crucial to the project. Therefore, also some smaller projects were reported.

Correlating cost/duration It is already known for decades that the effort of software projects is correlated to duration. For instance, Rayleigh curves are used as a means to correlate effort and duration. So usually, you would expect such a correlation. Roughly, projects with high costs take longer than small and inexpensive ones. In Figure 1 we show a scatter plot of the 73 federal projects on a log-log scale. The black dots used the federal data as-is, the five red dots are estimates where no actual timeframe of a project was reported. The solid line represents the weak correlation between cost and duration.

Recall that Figure 1 uses a log-log scale so the widely scattered data are even more scattered on a linear scale. Some cost and duration estimations in the 73-project portfolio are very large. A partial explanation for the large deviation is that duration is interpreted in more than one way. For instance, one project's duration is 18 years; this might be true, but it is likely that in this case the entire life cycle is estimated. Other likely explanations are that the estimation process lacks professionalism, which was also found in the UK [9].

We want to assure that our situation is consistent with benchmark data. There are multiple data collections maintained about software projects, of which Capers Jones' collection is probably the most reliable and comprehensive [5], especially for larger projects. The data stems from 1999, and does not solely contain Dutch projects, but there is no indication that it is not applicable in this case. That idea is being confirmed by Figure 1, in which we plot the Capers Jones benchmark data as the dotted line. The benchmark fits the federal data. Capers Jones also published best-in-class benchmark data, which if compared to multiple low-risk portfolios in Table 1, also show that the benchmark is in agreement with the actual data. These portfolios gathered by us contain data from private sector IT projects. By combining many portfolios we obtained a large private sector portfolio with 5842 projects that have size and duration data, which is listed in Table 1, too. In this table we also show that the federal portfolio has a correlation coefficient of 0.23, whereas perfect explanation would be 1.0, so this is not too strong. However, compared to the correlation coefficients in the private sector, which also have low correlation coefficients, this is consistent.

portfolio	#	R^2	Conform to benchmark?
federal	73	0.23	In agreement with benchmark
private	5842	0.13	Location corresponds, but slope is a bit less
org A	$\simeq$ 150	0.40	In agreement with benchmark for best-in-class data
org B	$\simeq$ 250	0.37	Between best-in-class and normal benchmark
org C	$\simeq$ 250	0.35	Best-in-class but slope is less steep

Table 1: Fitted models between cost and duration for different portfolios. Because of confidentiality we do not give all details (# = number of projects).

Comparing project costs It is known that the project sizes of a portfolio of IT projects usually show a normal distribution when plotted on a logarithmic scale (e.g. [14, 7]). In Figure 2 we show this phenomenon for 5965 IT projects from the private sector. We combined the IT portfolios of multiple large private organizations to obtain an idea of the distribution in the private sector. We also show the federal cut-off portfolio in this plot. In both cases, only IT costs are considered. Each green line in Figure 2 represents a project, while the distributions of the different portfolios are shown in lilac.

FP	$/FP: MIS	dur MIS	$/FP: OUT	dur OUT	$/FP: MIL	dur MIL
100	352.46	5.75	459.21	5.25	-	-
1000	1035.46	15.85	1104.31	13.80	1949.84	19.50
10000	2851.90	43.65	2904.62	36.31	6055.47	52.48
100000	-	-	-	-	16950.45	141.25

Table 2: Cost per function point and durations in months according to Capers Jones [5, pp. 191, 274, 498].

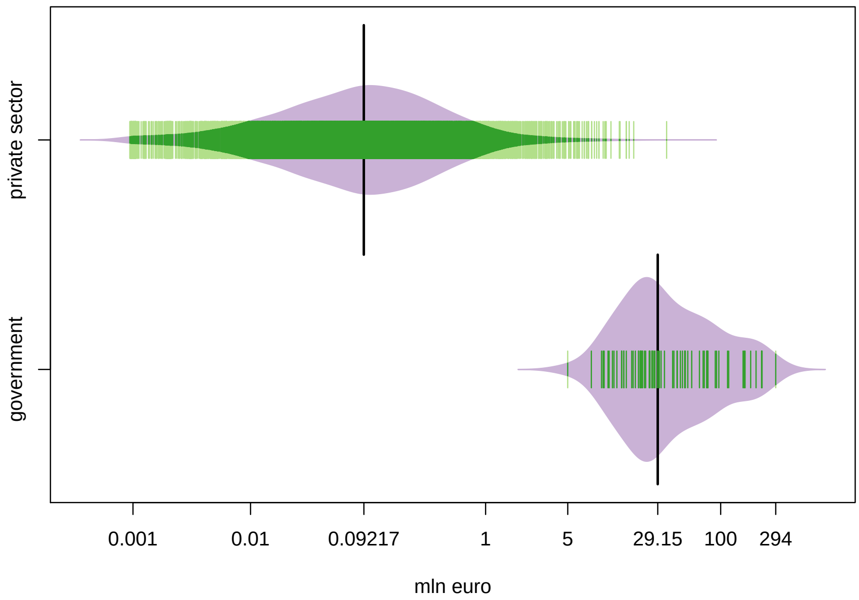

Figure 2: Beanplot [6] of the project size distribution of a private 5965 project portfolio compared to the 73 project federal portfolio.

The private sector portfolio almost shows a normal distribution, as found in the literature. It appears that the largest federal projects are of an entirely different scale than the projects found in the private sector. On geometric average, private sector projects cost 92.170 euro, while the federal projects cost 29.15 million, which is more than 300 times as much.

Private organizations mainly invest in projects with a size of 10.000 to 1 million euro. The federal cut-off portfolio contains projects with a size of 20 to 1000 million euro. There is only one private project that falls in this category. Even if one keeps in mind that only 73 of the largest federal projects are shown, the projects are still significantly larger than in the private sector.

Benchmarking failure

The best understood metric by executives is money. Therefore, we will express the risk of this cut-off portfolio in euros. We used public benchmarks to construct a derived benchmark that directly calculates wasted money. We used our own data and publicly available data from Capers Jones [5].

For the cost of a project, we plot the probability of failure in Figure 3. The exact formulas used are given in the two panels. It is clearly visible that when a project becomes larger, the probability of failure increases. Especially projects costing more than a million euro become risky.

Duration benchmark

Capers Jones [5] lists relations between duration and cost in dollars. We converted dollars to euros using a conversion rate of 1.0658 given by the European Central Bank for 1999. We also have to compensate for inflation, which has been 17.8% according to Statistics Netherlands, the Dutch Census. Therefore, cost in dollars (c) in terms of current euros (c_e) is:

$$c = c_e \cdot 1.0658/1.178206706$$

Jones lists different data for the categories Management Information Systems (MIS or I), Military (MIL or M) and Outsourced (OUT or O) software. The benchmark for duration derived from Table 2 is per category:

$$d_I = 0.02007679 \cdot c^{0.3027108}$$

$$d_O = 0.02120133 \cdot c^{0.2888007}$$

$$d_M = 0.02336047 \cdot c^{0.2925596}$$

Where d stands for duration in years, and c stands for cost in dollars. In millions of euros (c_m) the relation for MIS projects, shown as the dotted line in Figure 1, is:

$$d_I = 1.275784 \cdot c_m^{0.3027108}$$

The slightly different relation we found for our Federal case study (F), shown as the solid line in Figure 1, is:

$$d_F = 1.537648 \cdot c_m^{0.2911873}$$

In the federal portfolio we mostly have MIS projects. The MIS benchmark lies in the 95%-confidence interval for the first parameter [0.9762, 2.4219] and the second parameter [0.1695, 0.4129]. Therefore, we cannot reject the hypothesis that the Capers Jones benchmark fits the data perfectly. The data from the Federal portfolio is consistent with the benchmark.

In these benchmarks we converted cost in dollars into functionality in function points, but this relation is not always observed for federal projects. For example, in 1993, the Dutch version of the GAO [8] found that for similar systems in terms of functionality, some departments spent 3–5 million, while two departments needed more than 100 million each. As another example, in 2007 [2], it was reported that on one project 1493 function points were produced at a cost of 39 million euro. Either the data is invalid, or more than 20 times as much money was spend than industry benchmarks indicate. Such unexplained expenditures are in accordance with the work of Paul Strassmann [12], who found over and over again that IT spending does not necessarily lead to more productivity. For projects with unexplainable expenditures, our benchmarks will of course be very conservative in terms of estimated wasted money.

Applying our benchmark Using the cost data from the portfolios displayed in Figure 2 and the benchmark of failure displayed in Figure 3, we calculated the weighted risk of failure. For the projects from the private sector, we benchmarked that 17.9%

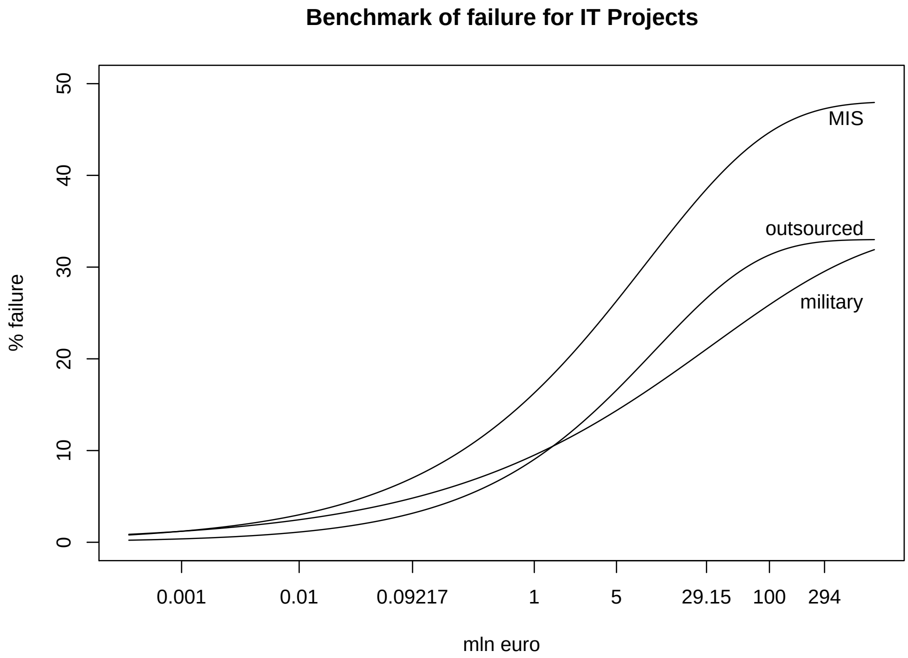

Figure 3: The larger the project, the higher the chance of failure.

Failure benchmark

Capers Jones [5] lists relations between function points, size and cost in dollars. Cost in dollars (c) are related to function points (f) as follows (derived from Table 2):

$$f_I = 0.07406812 \cdot c^{0.6877308}$$
$$f_O = 0.04744419 \cdot c^{0.7139649}$$
$$f_M = 0.05190356 \cdot c^{0.680408}$$

For function points (f) we inferred the probability of failure of the project (p) (from [14] and Table 3). A failure is defined as an aborted project.

$$p_I = 0.480552258 \cdot (1 - \exp(-0.007488657 \cdot f_I^{0.587382996}))$$
$$p_O = 0.330077910 \cdot (1 - \exp(-0.003296691 \cdot f_O^{0.678428650}))$$
$$p_M = 0.34066666 \cdot (1 - \exp(-0.01631659 \cdot f_M^{0.47035413}))$$

We plot the above three formulas in Figure 3.

of the money was allocated to aborted projects. This 17.9% is allocated to 7.8% of the projects. This is lower than 18 out of 156 projects (11.5%) reported in 2007 by clients from Cutter Consortium [4]. However, 18 out of 156 projects is not statistically

FP	failure MIS (%)	failure OUT (%)	failure MIL (%)
1	1	1	1
10	1	1	2
100	5	2	6
1000	17	10	10
10000	39	27	25
100000	48	33	33

Table 3: Relation between function points and canceled projects. Capers Jones [5, p. 192, 275, 499].

different from our 7.8% (p-value of 0.0985 using a binomial test). Our 7.8% is also not statistically different from 38 out of 412 projects (9%) reported by experienced managers on relatively large projects in 2002 [11] (p-value of 0.27). Unfortunately, the real failure rates in our dataset are mostly unknown. For some subportfolios in our dataset, we found 4–6% aborted projects, but we suspect that not all aborted projects were marked as such. In conclusion, we consider the benchmarked 17.9% loss on aborted projects as normal in the private sector.

The published federal list does not contain information on whether the reported costs comprise one or more projects. Since government did not report on this aspect, we were forced to use scenarios. In these scenarios we also list the loss that occurs if a project is aborted. In this article we present a new benchmark giving the loss at time of abortion.

Budget spent at time of abortion

In Figure 4 we show a group of 86 projects that have been aborted in private organizations. For the whole group, at abortion 91.6% of the total budget was spent. In this group, there were only projects with an IT component of 100%. Therefore, our best-effort guess for the federal portfolio is that on abortion 91.6% of the IT costs have been spent. For the non-IT costs (28.6% of the federal portfolio) we propose, based on personal experience, to take one third of that cost budget. So for the total budget, we estimate that $0.916 \cdot 0.714 + 0.333 \cdot 0.286 = 75\%$ of the total budget has been spent at the time of abortion in our best-effort scenario, but we also use other scenarios.

example In the federal portfolio, one project that should have been finished has been aborted after spending 87 million [10] out of a budget of 127 million. This project had an IT component of 60%. Assuming that 91.6% of the IT component, and one third of the non-IT component has been spent before abortion, we obtain exactly 87 million in this case. It is pure coincidence that our calculation exactly matches the reported loss. Due to the large spread in the numbers, normally we should only do this at the portfolio level, as it takes more than one project for the variation to average out.

Three scenarios

For the federal portfolio, we opted for three scenarios, which we list in Table 4. In our just-mentioned best-effort, or 50/75 scenario, the largest subproject for each federal project is 50% and 75% of the total investment is lost if the largest subproject fails. In

Figure 4: Failed projects of which 91.60% of the total budget was spent.

scenario	normal loss	%	risk excess	%	total loss	%
50/75	780 m€	17.9	855 m€	19.5	1637 m€	37.4
20/20	208 m€	17.9	169 m€	17.9	378 m€	32.4
100/100	1040 m€	17.9	1318 m€	22.6	2362 m€	40.4

Table 4: Scenario results for the federal portfolio.

this scenario, the abortion risk amounts to of 37.4%, of which 19.5% is excess risk. We define excess risk as the risk above the normal risk of 17.9% we observed in the private sector. Here, the expected loss because of excess risk is 855 million euro. In total, 1637 million euro is expected to be spent on failing projects, also taking into account the normal risks associated with IT projects.

In a conservative, or 20/20, scenario, the largest subproject for each federal project is 20% and 20% of the total investment is lost if the largest IT project fails. In this conservative scenario we obtain 378 million euro that will be wasted, of which 169 million is excess risk.

In a worst case, or 100/100 scenario, we obtain 2.4 billion euro wasted, of which 1318 million due to excess risk.

Our scenarios are conservative due to the following simplifications.

- We did not consider the cost of unnecessary budget and schedule overruns.
- The additional cost of business failure, which for example occurs because old systems have to be kept in place and opportunities are lost, was assumed to be zero.
- Correlations between the projects were not accounted for; for example, if many projects are simultaneously starting the risk of failure is compounded due to lack of experienced personnel at all levels, from programmer to program manager.
- In our 50/75 and 20/20 scenarios, we have only taken into account the case were the largest subproject fails, but sometimes other, smaller subprojects will fail.
- Projects that should be terminated but continue nevertheless are not taken into account (e.g. [1]).

Reusable components that are often claimed after project failure were considered to be of no value, because in practice we hardly ever see actual reuse.

Updated portfolio

In 2009, the government published an update on the portfolio [13], which now only includes 34 active projects, mostly because projects from non-departmental government bodies were taken off the list. Eleven projects have been finished, one was put on hold and one was canceled. In the current version, cost overruns of 390 million and two stopped projects with a total investment of 210 million are reported, indicating that the order of size of hundreds of millions depicted by our scenarios is solid.

There is only one new project, indicating the success that smaller, less risky projects are being done. However, based on earlier results [15], it is likely that many projects just below 20 million are being launched. To keep projects on the radar, it would be better to ask for the top 5 projects for each department, including all government bodies.

Portfolio approach

While there are no simple solutions to avoid risky projects, it is best practice to avoid large projects and split them into smaller ones. This is best done by investigating the project size distribution on a portfolio level. In the United States the Clinger Cohen Act already mandates a portfolio approach. For such a portfolio, one can asses the risk of the complete portfolio, and calculate past performance. A portfolio gives an overview in which excess risk is spotted, and projects can then be split up or delayed to regain control. Such a portfolio analysis, which can prevent huge losses, is not often done currently (e.g. [9]).

By using a portfolio approach and very minimal data, we assessed the risk of a published Dutch federal IT portfolio. It turned out that this portfolio contained many large projects. This lead, in a conservative scenario, to an abortion risk that is already more than twice as high as the normal risk found in the private sector of 17.9%.

References

[1] Barry Boehm. Project termination doesn't equal project failure. *IEEE Computer*, 33(9):94–96, 2000.

[2] Peter Brussaard, Frank Niessink, Erik Stel, Jan Turk, and Gerrit Vogelzang. Information baseline measurement for the policy domain for PSB, 2007. In Dutch, Available via http://docs.minszw.nl/pdf/35/2007/35_2007_3_10929.pdf.

[3] D. Dalcher and A. Genus. Introduction: Avoiding IS/IT implementation failure. *Technology Analysis & Strategic Management*, 15(4):403–407, december 2003.

[4] Khaled El Emam and A. G. Koru. A replicated survey of IT software project failures. *IEEE Software*, 25(5):84–90, 2008.

[5] C. Jones. *Software Assessments, Benchmarks, and Best Practices*. Information Technology Series. Addison-Wesley, 2000.

[6] Peter Kampstra. Beanplot: A boxplot alternative for visual comparison of distributions. *Journal of Statistical Software, Code Snippets*, 28(1):1–9, 10 2008.

[7] K.D. Maxwell. *Applied Statistics for Software Managers*. Software Quality Institute Series. Prentice-Hall, 2002.

[8] General Chamber of Auditors. Budget administration systems, 1993. The Dutch House of Representatives 23 265, 2, in Dutch, Available via www.rekenkamer.nl/9282000/d/tk23265_2.pdf.

[9] The Royal Academy of Engineering and The British Computer Society. The challenges of complex IT projects, 2004. Available via www.bcs.org/upload/pdf/complexity.pdf.

[10] Dutch House of Representatives. Kamerstuk 26448, no 371, 2008.

[11] Chris Sauer, Andrew Gemino, and Blaize Horner Reich. The impact of size and volatility on IT project performance. *Commun. ACM*, 50(11):79–84, 2007.

[12] P.A. Strassmann. *The Business Value of Computers – An Executive's Guide*. The Information Economics Press, New Canaan, Connecticut, USA, 1990.

[13] G. ter Horst. Report on large IT projects, 2009. In Dutch, Available via www.bzk.nl/onderwerpen/openbaar-bestuur/kamerstukken/120874/2009.

[14] C. Verhoef. Quantitative IT portfolio management. *Science of Computer Programming*, 45(1):1–96, 2002. Available via: www.cs.vu.nl/~x/ipm/ipm.pdf.

[15] C. Verhoef. Quantifying the effects of IT-governance rules. *Science of Computer Programming*, 67(2–3):247–277, 2007. Available via www.cs.vu.nl/~x/gov/gov.pdf.