The 10+/-2 Factors For Estimate Success

ICEAA Symposium – October 2016

Andy Nolan – Chief of Project Estimating – andy.nolan@rolls-Royce.com

Olimpia Vlad – Technical Analyst – olimpia-diana.vlad@controlsdata.com

Andrew C Pickard – Fellow of Systems – Andrew.C.Pickard@rolls-royce.com

Abstract

Success leaves clues and successful estimates are no accident. Success is defined here as an accurate estimate, one where the final project costs and schedules come in acceptably close to the estimate.

If we were to ask estimators for their views on the factors that lead to an accurate estimate, we might hear answers like; clear requirements, accurate historic data, and robust estimation techniques. But how important are these and is it possible to quantify their value, their contribution, towards a successful estimate?

This paper summarises the research of several hundred estimates across Rolls-Royce. The research aims were to determine what matters when estimating, what techniques really work and what are the top Factors that assure success.

Based on BAE Systems' Estimate Maturity Assessment (EMA) method (BAE Systems 2014), we developed a calibrated scoring mechanism to quantify the "maturity" of an estimate, to help leaders understand the "goodness" of an estimate before they make a commitment. The concept is called Estimate Readiness Level and is a score between 1 and 9 representing the maturity of the estimate. Like Technology Readiness Levels, 1 represents a low maturity estimate and 9 represents a high maturity estimate.

This paper summarises how Rolls-Royce calibrated the Estimate Readiness Level assessment so that project managers can now predict the uncertainty in an estimate before the project launches.

Introduction

Creating an accurate estimate can be a challenge, especially when estimating the costs to develop new or novel technologies. Projects exceeding costs estimates and completing later than scheduled is a long-standing problem widely recognised (Evans 2005, GAO 2009, Hammer and Champy 2003, Nikitrina, Kajko-Mattsson, and Nolan 2015, Roush 2005, The Standish Group 2013, Tishler et al. 1996 and Zhang et al. 2003).

> It may not be possible to prevent a bad estimate, but it is possible to prevent it from being a surprise.

We live in a rapid changing environment in which the need for, and benefits of, effective cost estimating has become even more evident with more complex customer requirements and expectations of increased accuracy (Boardman 2013 and Greves and Joumier 2003).

For several years effort has been devoted by The Standish Group (2013) to collecting real-life case information with the purpose of showing organisations how to optimise their projects successfully. Their figures show that it is very unlikely for very large projects to 'finish on time, on budget, and within scope', which is The Standish Group (2013: 4) definition of a successful project.

There is an understandable push by any business to improve performance and reduce costs. Many estimators face budget reductions which increase project risks and may actually be a reason for longer term project problems. The very act of trying to reduce the cost may be the reason for longer term cost overruns and performance issues.

Many estimators unwittingly commit themselves to a troublesome future. Like someone who lives an unhealthy life-style, they do not experience the consequences today. They may continue believing that in their case, this time, there are no consequences. The business encourages this thinking by rewarding low estimates and challenging estimates that are perceived to be high. Opinions may trump data especially if the data suggest a more expensive outcome.

There can be more of politics than science in estimation. Estimation is a passionate subject in that, after presenting a scientifically analysed estimate, the passion may emerge. There are few other analytical disciplines that invoke so much emotion in a business. It is unlikely that a mechanical stress analysis would provoke the emotional responses that an estimate can, especially when personal and business success hinges on a low cost or short schedule..

The estimate warns us today what will happen tomorrow as a consequence of decisions we make. Estimators can be like a nagging doctor constantly telling us to eat better and drink less alcohol and coffee. It can be annoying because in the back of our minds, we know the doctor is right and we would prefer to live in hope rather than having to make a tough decision to actually do what is right.

There are sources of research analysing the root causes for poor projects and poor estimates. But whilst it all seems logical in hindsight, is it possible to predict the problems before the fact rather than after? Is it possible to actually predict how much a project is likely to overspend or how much risk it contains? And if we could predict it, could we sell this message convincingly to our business leaders in a way that they will listen?

With over 120 identified root causes for estimation problems (identified through a Process Failure Modes and Effects Analysis inside Roll-Royce) and with many of these relating to culture and behaviour, estimation is a challenging discipline – not just for the science involved but because of the way those estimates are sold to, and accepted by, the business.

We need a convincing way to assess the risks involved in an estimate and then present this in a simple way that people can understand and accept. The paper presents the concept of Estimate Readiness Level as a way to assess an estimate's maturity and to communicate this in a simple way that will make it difficult to ignore.

Human behaviour can neither be legislated nor regulated, but understanding which behaviours introduce errors means that measures can be put in place (Bassford 2012). It may not be possible to prevent a bad estimate, but it is possible to prevent it from being a surprise!

Why Estimate?

In our home, there are few of us that would commit work to a builder without an understanding of the likely costs of the project. We estimate because most of us want a good understanding of what is likely to happen in the future, so we can make informed decisions. An estimate suggests a range of possible futures that are most probable. If we do not like the future, then we can either change what we want or change what we do.

We estimate:

- To make good business decisions;
- To understand and challenge costs;
- To make better plans and commitments;
- To optimise technical solutions for cost and benefit;
- To allocate budgets;
- To plan resource load and capability;
- To identify and quantify risk;
- To identify or validate the benefits from improvements.

We estimate because it is good for business.

Robust estimates become imperative to realistically assess the commercial benefit of taking the project on, and the associated risk involved. Robust estimates can also be used to objectively explore ways to reduce costs.

We estimate because it is good for business. So if estimating is so valuable to any business, why do estimates fail?

Why Do Estimates Fail?

An estimate can fail either through under-estimation or over-estimation. An over-estimate can cause the loss of business or the cancellation of a valuable project. However, on the whole, the data from across Rolls-Royce shows that 65% of the time we under-estimate i.e. we estimate too low. Figure 1 is a summary of a Root Cause Analysis of 118 projects across Rolls-Royce conducted in 2012: the chart is the weighted contributions to project overspend. The dominant reasons for issues were:

- Lack of data and/or calibrated tools. This means the estimator did not have access to good data or validated and calibrated estimation tools.
- Lack of or ineffective estimation process. The estimator did not follow a structured approach for estimate development, buy off, maintenance and close.
- Failing to factor for risk & uncertainty. The estimator did not perform sufficient risk/uncertainty analysis.
- Culture and behaviour. There are a multitude of problems including biases (optimism), rushing estimates, and so on.

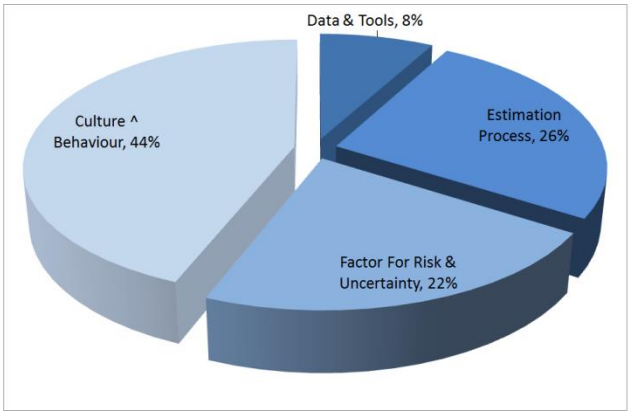

Figure 1: Reasons Estimates Fail. Analysis of 118 projects

Interestingly, lack of data and tools was the lowest contribution to estimation failure. This means that if we were to use the world's best data and tools, it would have a negligible effect on our estimate accuracy. However, the research did reveal that a good estimation tool can drive the right process and behaviours and will minimise many of the dominant root causes for failure.

What was needed was an assessment method to detect the right behaviours and practices and to warn the business in the event of unhealthy practices. This assessment method would need to cover all dominant root causes for estimate failure and present the results in a simple standard way that both the estimator and business leaders would understand. For this reason we adopted from BAE Systems the concept of Estimate Readiness Levels.

Estimate Readiness Levels

The Estimate Readiness Level (ERL) is a score between 1 and 9 representing the "maturity" of the estimate and its readiness for use. 1 represents a low maturity estimate and 9 represents the highest. The range 1 to 9 was chosen to mimic existing scoring mechanisms like Technology Readiness Levels or Manufacturing Readiness Levels.

The ERL concept is based on the principle that an estimate needs to be fit for purpose, no more and no less.

The ERL assessment was developed to quickly assess the health of an estimate before handing it over to the business. It can also assess the health of estimates that we receive from our suppliers.

It is not always necessary to achieve a high ERL score. For example, if we are exploring trade options, we want the ability to quickly generate estimates without the burden of the full formal process. In contrast, when committing the business to a critical deadline or price, we would need a mature estimate.

The following are a number of scenarios where the ERL assessment has prevented a problem:

- The business needs a high maturity estimate, the ERL assessment will ensure the estimate is of sufficient maturity.
- We are asked to make a guess (rather than an estimate) which then becomes our budget. Using the ERL assessment, the guess would be accompanied with a low ERL score as a warning to the budget holder not to use it in this way.
- An estimator may be rushed to develop an estimate. Again, the estimate would result in a lower ERL and a warning to the customer.
- An estimator may be asked to generate an estimate with little information about the project. The corresponding ERL would reflect the risk. This implies that it is difficult to get a good ERL score for novel technologies.
- The customer expects a high maturity estimate but the ERL score shows why it is not possible at that time.

The problem is that many customers of estimates do not recognise that the estimate "quality" can vary. They may treat all estimates as equal. A low maturity estimate could unwittingly be used for critical business decisions.

The ERL concept is based on the principle that an estimate needs to be fit for purpose, no more and no less. There is little point investing thousands of hours in an estimate that is destined only for a brief presentation. As Peter Ducker put it 'There is nothing quite so useless, as doing with great efficiency, something that should not be done at all.'

Likewise, it is unwise to under invest in an estimate that is used to commit a business. A nervous estimator will likely over invest in an estimate and in contrast a confident estimator is likely to underinvest. With ERL we can warn against both of these extremes.

Table 1 is a summary to the purpose of each ERL level. We have introduced the term Desired ERL to demote the maturity the estimate needs to be in order to be considered fit for purpose.

Desired ERL	Precision	Description
ERL9	-5% / +5%	Annual budgeting and post contract budget baseline
ERL8	-9% / +10%	Contract Signature / Memorandum of Understanding
ERL7	-17% / +20%	Request For Quotation
ERL6	-29% / +40%	Request For Proposal
ERL5	-44% / +80%	Request For Information
ERL4	-62% / +160%	Exploring Trade Options
ERL3	-76% / +320%	Exploring Strategy
ERL2	-86% / +640%	Exploring Expectations
ERL1	-93% / +1280%	Do not use this estimate

Table 1: ERL Definitions

The "+" values shown in Table 1 double with each level down the table. The "-" values were calculated as *(1 - 1 / (1 + x))* where x is the +%. Thus the "-" values tend to -100% as ERL decreases.

The challenge now was to develop a way that could quantify the ERL scores and create an objective way to assess maturity. For this reason we chose to calibrate the ERL assessment.

Calibrating The ERL Assessment

The research was to understand what contributes to estimate success. It was not intended to simply gather people's opinions or to capture lists of common root causes. The research was to statistically find the relationship between estimation practices and estimate success.

The research was not to understand what happens but what matters.

The research was not to understand what happens but what matters. For example, if we did a root cause analysis for a Bar-Be-Q, we might find that rain was a root cause for an unhappy event. Wider analysis may show that many Bar-Be-Q's in England are afflicted by bad weather. But regularity is not necessary a guide to impact. If some Bar-Be-Q's were successful on rainy days then we may find that rain is not the dominant factor. It might be statistically possible to show that over-indulgences is a stronger reason for problems.

The authors conducted research over a four-year period on 300+ projects of which 134 were completed projects where the final estimate accuracy was known. The research was on Rolls-Royce projects, both small and large covering civil, aerospace and defence projects.

In each case, the research consisted of scoring the estimate for its application of a number of estimation practices (referred to as Factors in this paper). An example Factor might be the estimation techniques, use of risk management, estimate documentation and so on.

The Factors were gathered from many years of analysis on the reasons estimates fail. They also included insights from a Process Failure Modes and Effects Analysis of the estimation process.

The estimators were asked to score themselves for how robustly they applied each Factor. The estimator was given a pre-defined range of Options to select from. For example, a Factor could be "How much experience did the estimator have" and the Options would be 3 months, 6 months, 1 year, 3 years and 6 years. We used pre-defined Options to minimise errors and biases. For each Factor, the selected Option was then turned into a numerical score to aid analysis.

The research was to find the relationship between Factors and estimate %Error. %Error was measured as *(Estimate / Actual) -1*.

But there was an immediate problem. Imagine we were to guess the number that would come up on the roll of a normal 6 sided dice. Assuming the dice was balanced, we would expect our guesser to be accurate 1 in 6 times. We may not be able to confidently predict the score that would come up on a roll of the dice but we could confidently predict its range.

The same affect applies to estimates. If we were to develop 100 estimates using the exact same estimation practices, we would expect a range of results rather than the exact same %Error.

The analysis had to calibrate to a range rather than a point value – for this we used ERL. For example, for ERL 9, we defined the estimate %Error to lie in the range -5% / +5%.

So, with 134 completed projects, the qualitative results for 50+ Factors and the final estimate %Error, it was then possible to perform analysis to understand the dominant Factors for success. There are many ways to perform multi-way regression but for this paper, the following equation was used:

$C_1F_1 + C_2F_2 + \ldots\ldots C_nF_n <=$ ERL

ERL = derived from table 1 based on the actual estimate %Error

$C_1, C_2, \ldots\ldots C_n$ = Constants (to be derived)

$F_1, F_2, \ldots\ldots F_n$ = Factor Score (Taken from the estimator assessments)

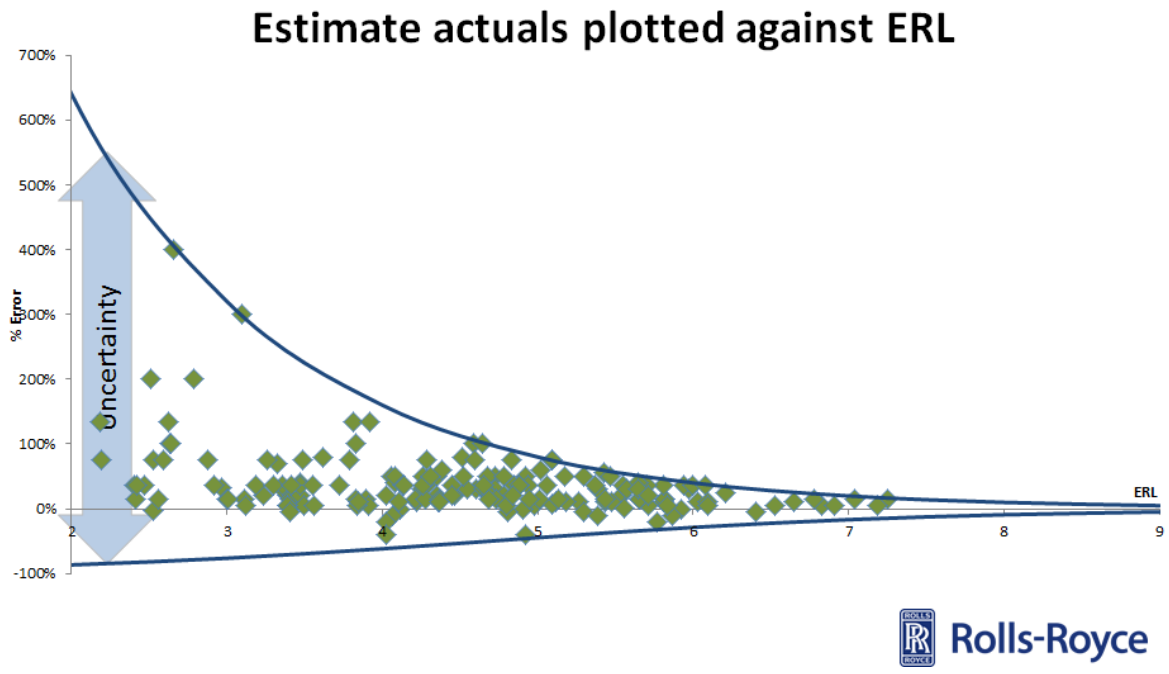

Figure 2: 134 completed projects. The chart shows the relationship between ERL score and final estimate %Error.

The research was to find the constants so that all estimates lay inside the cone (Figure 2) and to the furthest right of the cone (taken from Table 1). Figure 2 shows the results of the 134 projects. It suggests that if a project were to score itself using the ERL assessment, then if the project has a high ERL, it will be more accurate than a project with a low ERL. But based on the example of the dice, a low ERL project could still, by chance, develop an accurate estimate.

These paper summaries how the ERL assessment was calibrated to help give an objective, calibrated, understanding of estimate maturity. The paper gives an overview of the ERL assessment tool and how it can be used to benchmark the business. The paper concludes with some evidence of the benefits from improved estimating and the benefits from high maturity estimates.

The 10+/-2 Factors For Estimate Success

Figure 3 shows, in ranked order, what matters when estimating. Each Theme represents one or more Factors taken from the survey. The Themes was calibrated to a maximum score of 9 i.e. the sum of the constants $C_1 + C_2 + \ldots + C_{11} = 9$. The scores show the contribution (ERL points) each Theme makes to an accurate estimate.

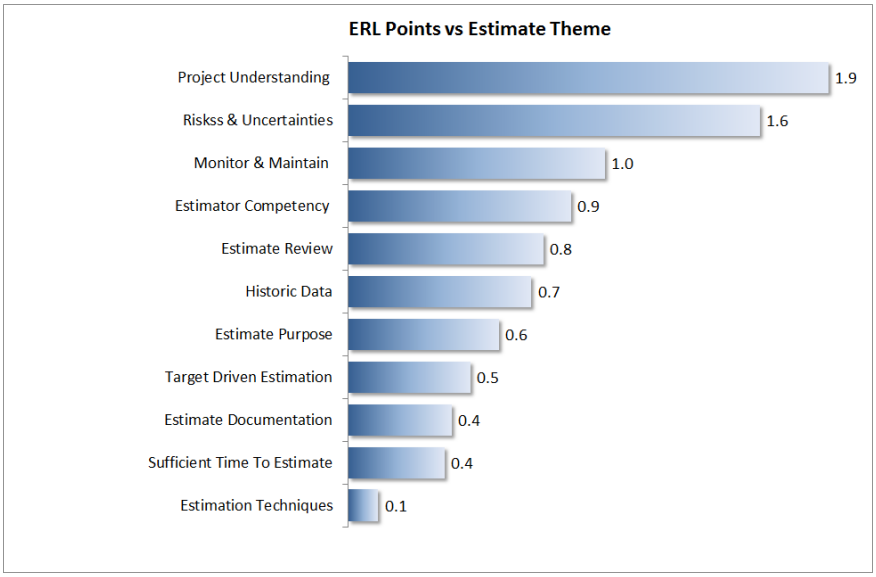

Figure 3: The 10+/-2 Factors For Estimate Success. Scores are shown in ERL points earned with a maximum score of 9 ERL points.

#	Theme	Description
1	Project Understanding	Known-Knowns: A score of maturity of understanding about the project e.g. requirements, scope, environment etc.
2	Risks & Uncertainties	Known-Unknowns: A risk maturity score. A measure of the breadth and depth of risk analysis.
3	Monitor & Maintain	A maturity score of how well the estimate was monitored, governed and maintained through the life of the project
4	Estimator Competency	A maturity score of the estimator's capability in both estimating and the domain of the project.
5	Estimate Review	The maturity of the independent review: robustness, formality, attendees, capability of reviewers etc.
6	Historic Data	A measure of data maturity. Using historic data that is credible, relevant, current and complete.
7	Estimate Purpose	A measure of the maturity of understanding by the estimator and customer for the purpose of the estimate.
8	Target Driven Estimation	The maturity of practices used to explore low risk solutions to meet a customer's cost/schedule challenge.
9	Estimate Documentation	A measure of the maturity of the documentation, content, clarity, detail etc.
10	Sufficient Time To Estimate	A measure of the effort, resource and schedule need to develop the estimate.
11	Estimation Techniques	The maturity of the estimation techniques used to develop the estimate e.g. top down, bottom up etc.

Table 2: Definition of each Theme

We will now discuss 4 of these themes in more detail in the following sections.

Theme 2: Risks & Uncertainties

An understanding of the risks and uncertainties was the second most important Theme for estimation success. Further analysis was undertaken to understand the impact of various classes of risk on the estimate. Figure 4 was calibrated to a maximum score of 9 i.e. the sum of the constants $C_1 + C_2 + \ldots + C_{10} = 9$.

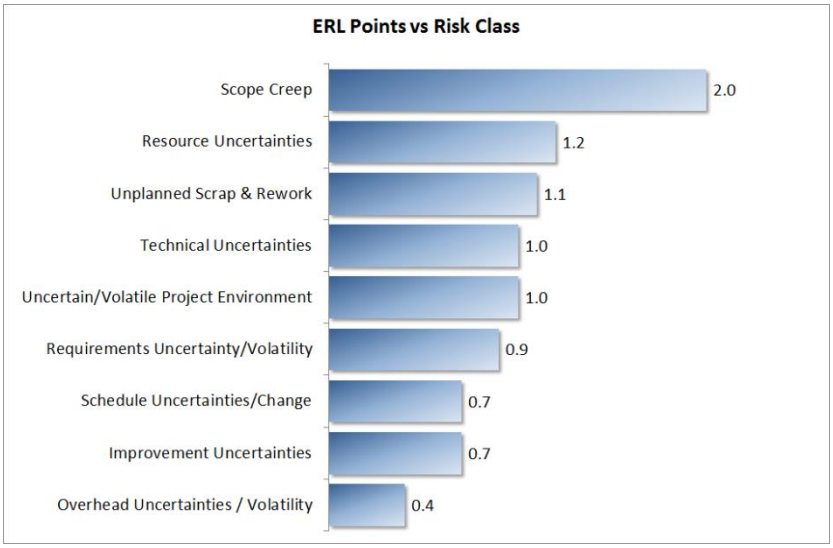

Figure 4: The Impact of Risk Classes. Scores shown in terms of ERL points. C1 + C2 + … + C10 = 9.

What became apparent from the research was that estimators have their favourite classes of risk that they will analyse and include in the estimate.

However, they will tend to overlook other classes of risk. This is a good reason why we may want several people doing the risk assessment to broaden the range of the classes of risk considered.

The results of the research are to ensure the estimators have given appropriate consideration to sufficient classes of risk in order to meet their Desired ERL. For example, if our Desired ERL is 5, we would only need to consider the top 4 classes of risk (or a combination of risk classes that added up to 5). To achieve a Desired ERL of 9 we would need to consider all classes of risk.

Although there was no surprise to see Scope Creep as the dominant risk class, it was a surprise to see that Requirements Uncertainty & Volatility was not a dominant issue. In the example of the Bar-Be-Q in the previous section, regularity is not necessarily a guide to impact. We all experience the frustrations of requirements uncertainty but this does not mean it is our biggest issue.

Risk Class	Description
Scope Creep	The boundary, inclusions and exclusions of project scope.
Resource Uncertainties	People and resources: the commitment acceptance process.
Unplanned Scrap & Rework	Design iteration caused by errors or unexpected results
Technical Uncertainties	Technical risks e.g. Technology Readiness Level
Uncertain/Volatile Project Environment	The stability and maturity of people, processes, tools, suppliers etc.
Requirements Uncertainty/Volatility	The maturity and stability of the project requirements.
Schedule Uncertainties/Change	Milestone dates shifting mid project.
Improvement Uncertainties	Taking on novel improvements to save cost and time.
Overhead Uncertainties / Volatility	Unexpected support, management, meetings, technical oversight etc.

Table 3: Definition of risk classes.

Theme 3: Monitor & Maintain

When a project launches, what should they monitor to improve the success of meeting the estimated cost and schedule? Figure 5 was calibrated to $C_1 + C_2 + \ldots + Cn =$ ERL. Because people pick and mix monitors, instead of making the constants add up to 9, we correlated to the ERL score. Making the scores add to 9 means that an estimator can only have a good estimate by using all monitors when this is not the case.

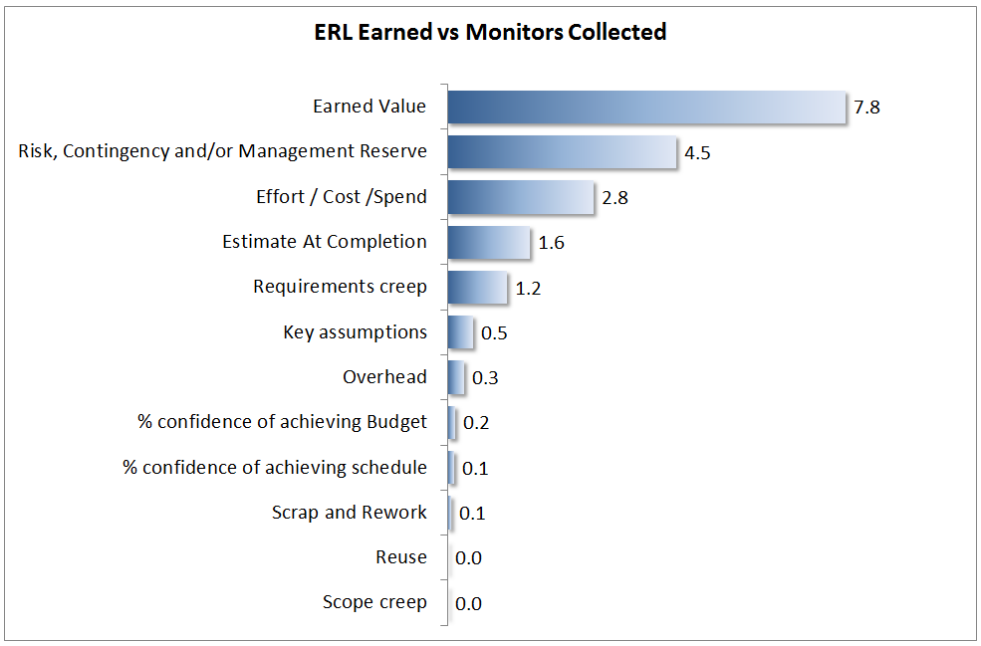

Figure 5: The benefit from project monitors. Scores shown in ERL points. C1 + C2 + … + Cn = ERL.

Monitors by themselves do not make a project successful. The monitors need to be combined with good estimating, planning and change control. The results of Figure 5 may be showing an indirect relationship i.e. the projects that are most successful also use Earned Value. But, to apply Earned Value to a project will not necessarily make it successful.

A project should be collecting a cocktail of monitors such that the total ERL score matches their Desired ERL. For example, if our Desired ERL is 7, the project could:

- Monitor Earned Value (7.8 points)
- Monitor Risk, Contingency and/or Management Reserve (4.5 points) and Effort / Cost /Spend (2.8 points) giving a total ERL score of 7.3.

It is interesting to see that monitoring risks and uncertainties is the second highest score (similar to Figure 3) but the monitoring of requirements, scope and scrap and rework all scored low. This could be suggesting that these monitors are not a differentiator between the successful and unsuccessful projects.

Like the example of the rain on an English Bar-Be-Q (previous section), if all people collect the same monitor then it cannot explain the difference between projects that succeed and fail. However, Earned Value would appear to be the dominant differentiator, exclusive to successful projects.

Theme 12: Sufficient Time To Estimate

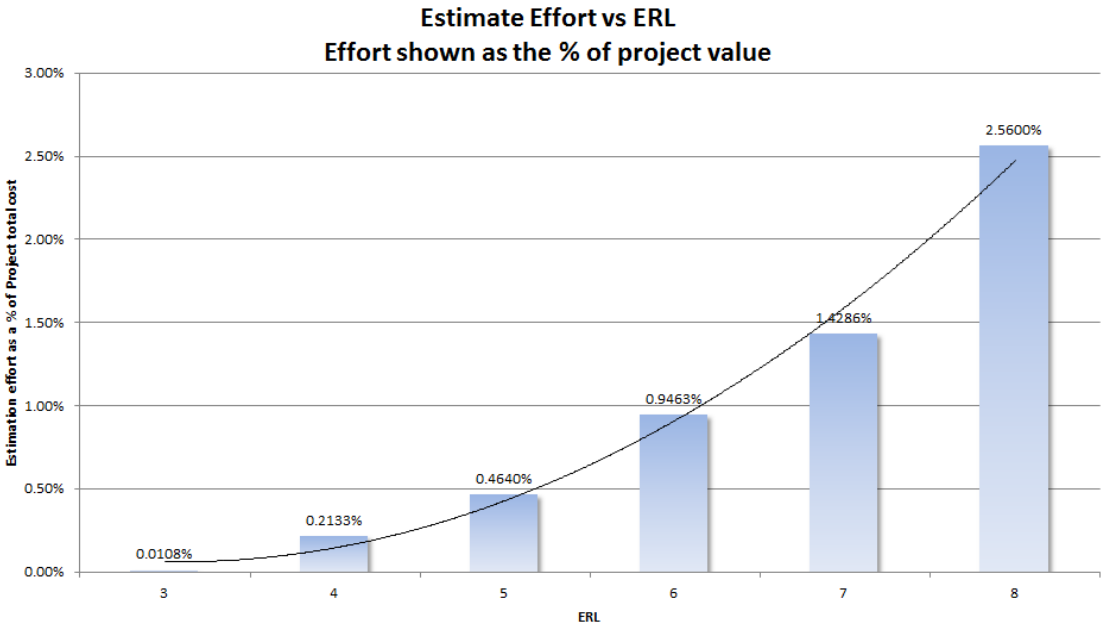

Figure 6: The effort to develop an estimate. Effort shown is a % of project value.

The research asked for the level of effort invested in developing the estimate. The effort was then normalised by the project size (value) to come up with a project "tax". For example, if the project cost 100 hours, and the estimator spent 1 hour developing the estimate, this meant the estimation effort was 1% of the projects value. Figure 6 shows the average scores for each ERL level. If we want a robust estimate, assuming all other things are equal, we are going to have to invest more effort.

Theme 11: Estimation Techniques

Although estimation techniques did not score highly by themselves, it was still interesting to see if there was a relationship between technique used and ERL.

Figure 7 was calibrated to $C_1 + C_2 + \ldots + Cn = ERL$. As for figure 5 (Monitors), we recognise that people tend to pick-and-mix techniques and we would not expect an estimator to use all techniques. The estimators were asked to indicate which of the following techniques they used.

Technique	Description	Use when	Good for	Key reliance
Judgement	The subjective view of an individual	There is a lack of data and tools	Quick estimates or when reviewing estimates	Domain experts
COCOMO	A software parametric cost estimating tool	Software dev. environment is changing	Characterising the dev. environment	Knowledge of the dev. environment
Monte-Carlo	Statistical simulation technique	Complex risk and uncertainty	Balancing a large number of risks and uncertainties	Understanding of risk, uncertainty and opportunities
Multi – People	Using multiple independent people	Avoid common mode failure of estimators	Drawing out assumptions and risk	Many competent estimators
Forecasting	Future costs are a function of the past	Many relevant historic data points	Quick estimates in predictable domains	Historic data + corrections factors

Technique	Description	Use when	Good for	Key reliance
Comparative	Future project is scaled from a past project.	Relies on a relevant historic data point	Quick estimate and justifying difference from past projects	Good data
Bottom up	Creating a WBS and costing each element.	we have detailed plans	Resource planning and budgeting	Experience in process, domain, etc.
Parametric	A calibrated parameter tool.	Calibrated and credible tool for the project	Exploring options and trades. Good for sensitivity analysis.	Knowledge of the parameters
Beans	Project expressed in "standard units".	Work can be expressed in units of work	Runner / repeater projects	All work can be expressed in terms of beans
Top down	Top level number – no breakdown.	Only high level concepts are known	Quick estimates	Historic data or calibrated tools

Figure 7 shows the results of the analysis. What seems disappointing is to find that Bottom Up did not score as high as hoped and Judgement scored the highest!

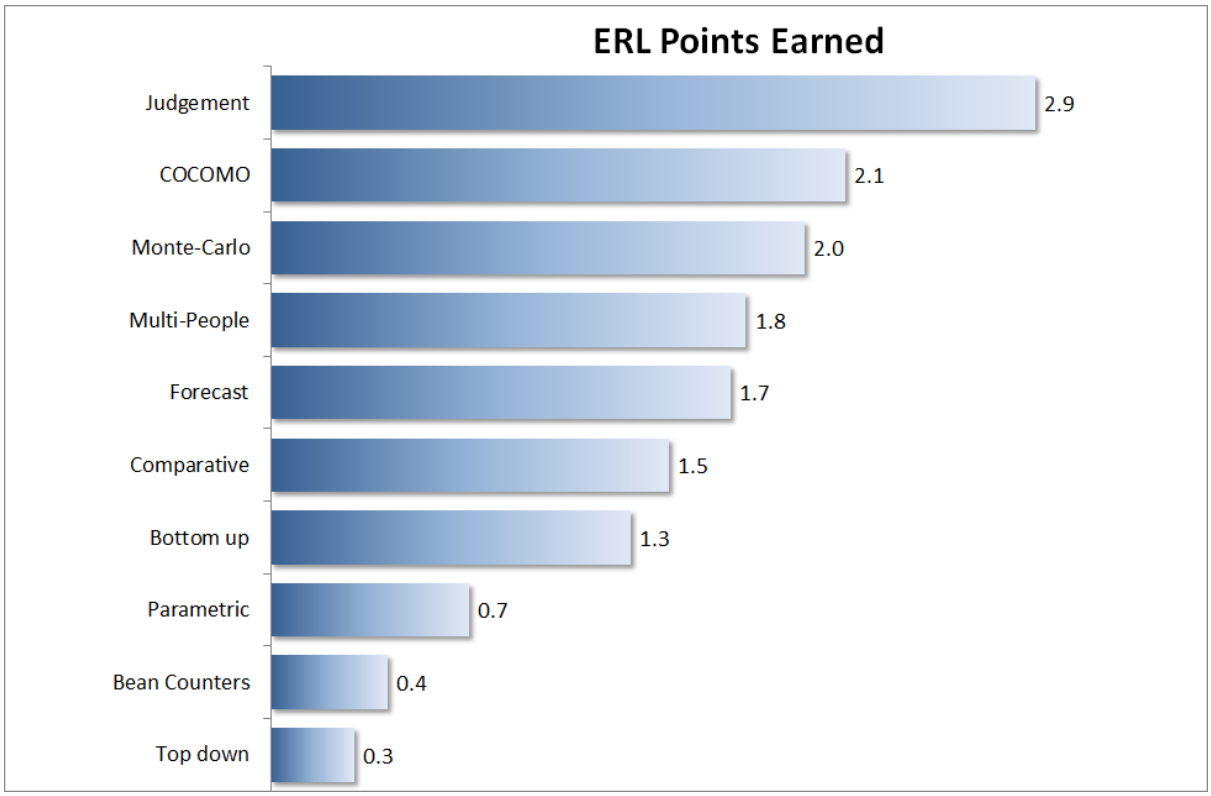

Figure 7: Estimation Techniques and ERL points earned. C1 + C2 + … + Cn = ERL.

Although judgement is probably used in all cases, in some form or other, we have interpreted that judgement in this case relates to deriving numbers from guesswork. However, take the example of a "tick box engineer" who flicks switches without applying any judgement to the input parameters or output results. 'A fool with a tool is still a fool' and the estimator is unlikely to recognise they have created a poor estimate.

It might also be indicating that some people are not experienced enough to use their judgement, and a lack of experience is affecting the estimate accuracy. A competent estimate may be more comfortable to use Judgement. Further research inside Rolls-Royce has found a correlation between the experience of an estimator and the accuracy of their judgements.

The hypotheses for why Bottom-Up did not score better include: (1) people overlook items from their estimate leading to scope creep (largest risk in Figure 4) (2) in contrast it is easy to bloat an estimate by rounding up each element of the estimate (3) leaders may have a tendency to nit-pick through bottom up estimates, trimming the costs and creating an optimistic estimate.

The estimator can use Figure 7 to create an ideal "cocktail" of methods to meet a Desired ERL. For example:

- If we use the Judgement (2.9 ERL points) of a single individual doing a Top-Down Estimate (0.3 ERL points) we would get a net ERL score of 3.2 i.e. it cannot be trusted.
- If instead, we did a Top-Down estimate (0.3 points) using Judgement (2.9 points) but with Multiple People (1.8 points) we get a net ERL of 5.
- Finally, if we used Multiple-People (1.8 points), using Judgement (2.9 points) on a Bottom-Up estimate (1.3 points), we get an ERL of 6.

The ERL Assessment

Based on the research described in this paper, it was then possible to create the ERL assessment.

There are 10 Themes in the ERL assessment and each Theme has a different weighting (or value) towards developing a mature estimate. The previous sections of this paper covered the weightings for each Theme. The Theme of Monitoring & Maintenance was not included as this is post estimate approval. The ERL assessment is shown in Figure 8.

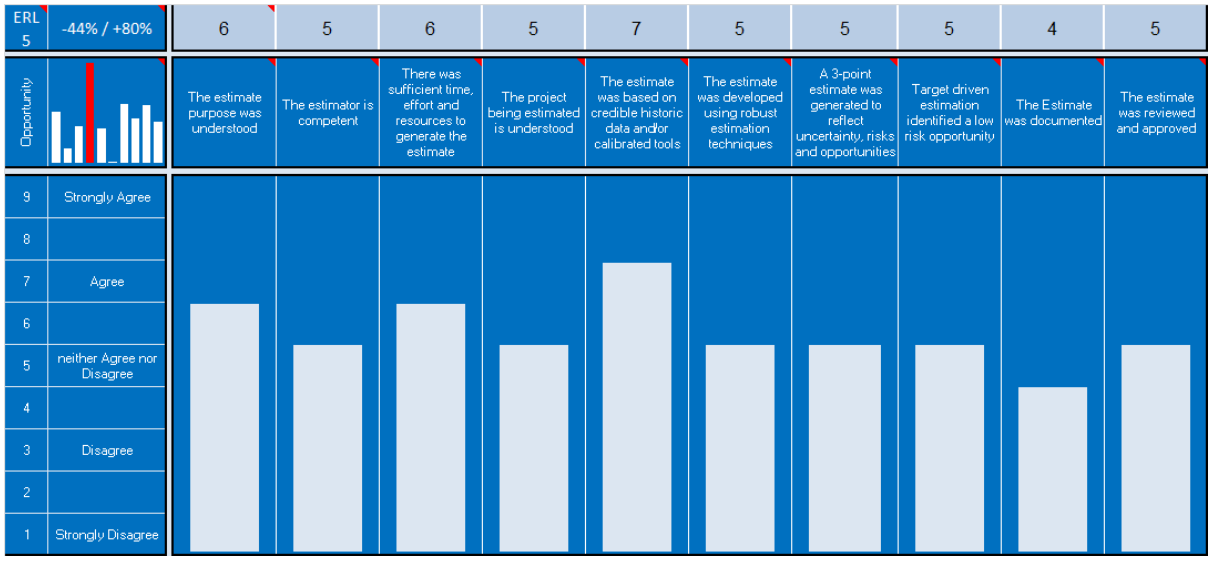

Figure 8: ERL Assessment

The first step is to score all 10 Themes in the range 1 to 9. To make this objective, the help pages used the same scoring mechanisms (Factors and Options) used to score the 134 projects.

On completing the assessment, a weighted ERL for the project is displayed in the top left corner. The +/-% values shows the level of uncertainty in the estimate based on Table 1.

If the estimate falls short of the Desired ERL then further work is required. The ERL assessment gives a bar chart (top left) showing where further ERL points are available by improving each Theme. The

estimator may want to focus on the most sensitive Theme (the bar shown in red) or to improve several of the less sensitive Themes.

There are hundreds of combinations of Themes to achieve a Desired ERL. This gives the estimator flexibility to find the optimal balance of Themes to meet the Desired ERL. Some combinations of Themes will be better suited to some estimates than others.

There are hundreds of combinations of Themes to achieve a Desired ERL. This gives the estimator flexibility to find the optimal balance of Themes to meet the Desired ERL.

In all cases, the estimator should only do what is necessary to meet the Desired ERL, no more and no less. This ensures the estimates are fit for purpose and the estimation activities are optimised for effort.

On completion, either the ERL assessment or the resultant ERL score is submitted with the estimate. By submitting the completed assessment, this provides audit evidence of the ERL.

So, if the ERL assessment could measure the maturity of a single estimate, could it help us plan our improvement strategies to help develop the organisations overall estimate maturity?

Benchmarking The Business

The ERL assessment can be used to assess the maturity of a single estimate or a business.

By analysing many estimates from a project, department or business it is possible to derive an overall picture of maturity.

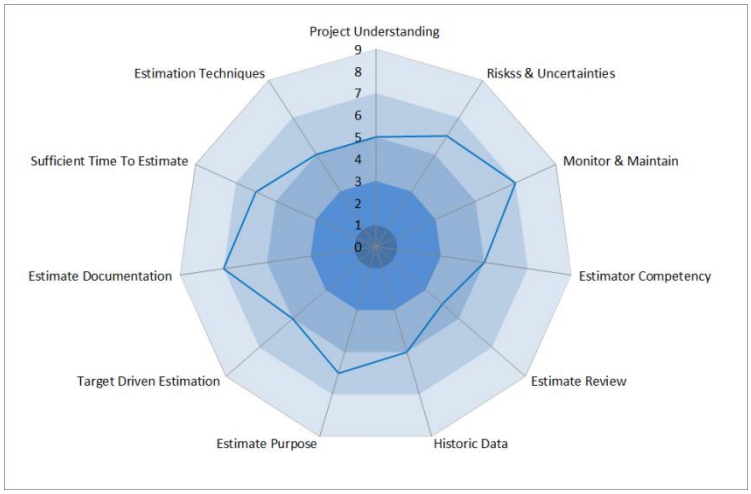

Figure 9: An example of a maturity assessment

For example, one department could not understand why some estimates were successful and yet, they would experience occasional problems. By scoring 70 projects, it was found that the average ERL for the department was 6. This means that the estimation capability deployed by the department left them exposed to an uncertainty of -29% / +40%.

When improving the businesses estimation capability, we could develop plans to fix what mattered rather than fixing what we thought was important. The Theme weightings helped us understand what mattered most.

With ERL, we can now optimise the estimation activities and even mature whole organisations. What other benefits have we experienced from the improved estimating?

The Benefits Of Estimating

The research also wanted to understand if the estimators experience any benefits from improved estimating. The Estimation Capability Owners may be challenged to show the benefits of estimating because for some, it is considered a burden to their teams doing "real work!"

We have kept a log of improvements that can be traced directly to estimating:

- A software department reported an average cost reduction of 11% after the introduction of COCOMO. A better understand of cost led to better cost management.
- One project reported a 45% cost reduction through improved estimating, planning and monitoring.
- A major project reported milestone achievement rose from 20% to 70% simply by applying schedule estimation practices.
- A survey of 118 Rolls-Royce projects in 2014 showed that productivity rose between 6% and 14% for any project that used estimating.
- One department developed an estimation tool to become a "smart customer". They showed problems with many suppliers' estimates and managed to reduce costs in most cases.

The better we estimate, the more accurate our estimates get AND the more benefits we experience.

So why should we expect benefits from good estimating? Here are a few proposed reasons:

- Making better decisions;
- Cancelling bad projects, launching good projects;
- Informed trades to optimise a project for cost and schedule;
- Better understanding of risks;
- Knowing what factors are sensitive and therefore need managing;
- Better control and understanding of cost.

As part of the research, we wanted to understand what benefits the estimator experienced as a result of the estimation practices. Figure 10 shows a summary of the 8 most commonly reported benefits. Figure 11 shows the average number of these benefits reported for each ERL level. Figure 11 shows a strong relationship between ERL and the number of benefits. The better we estimate the more accurate our estimates get AND the more benefits we experience.

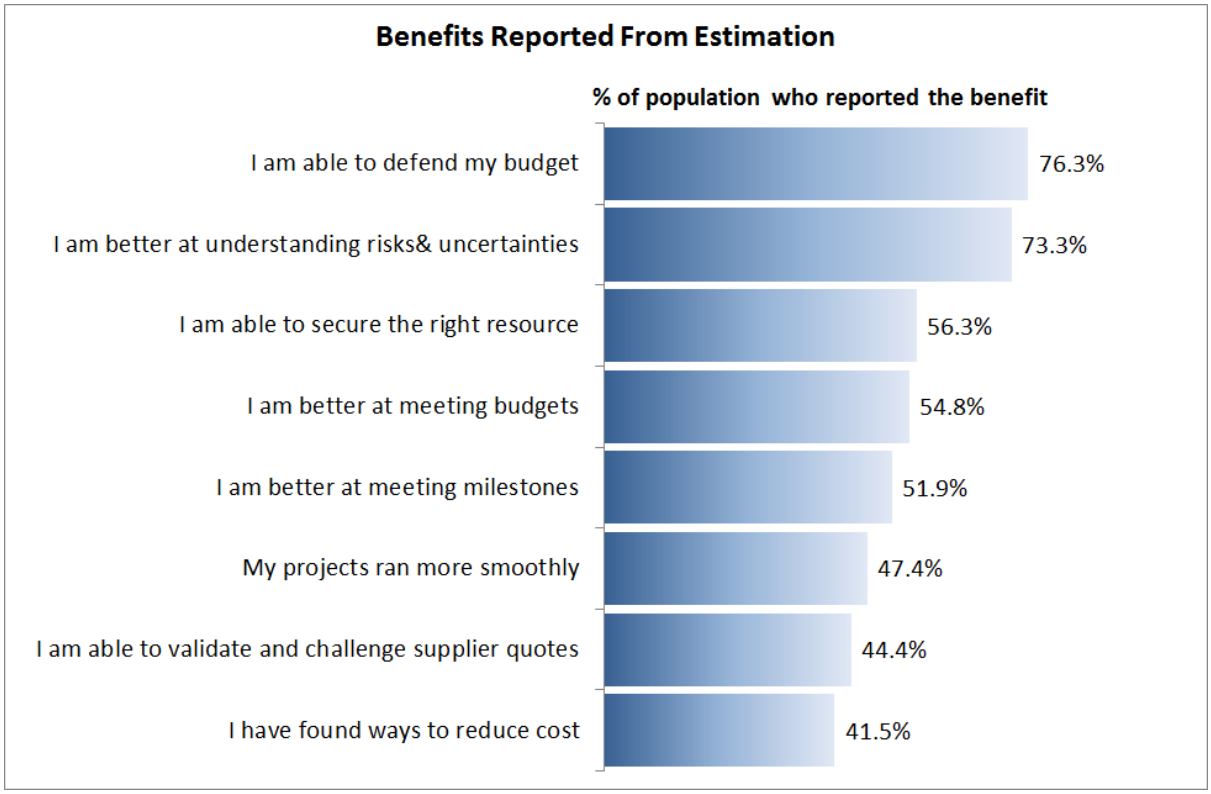

Figure 10: Benefits reported by estimators

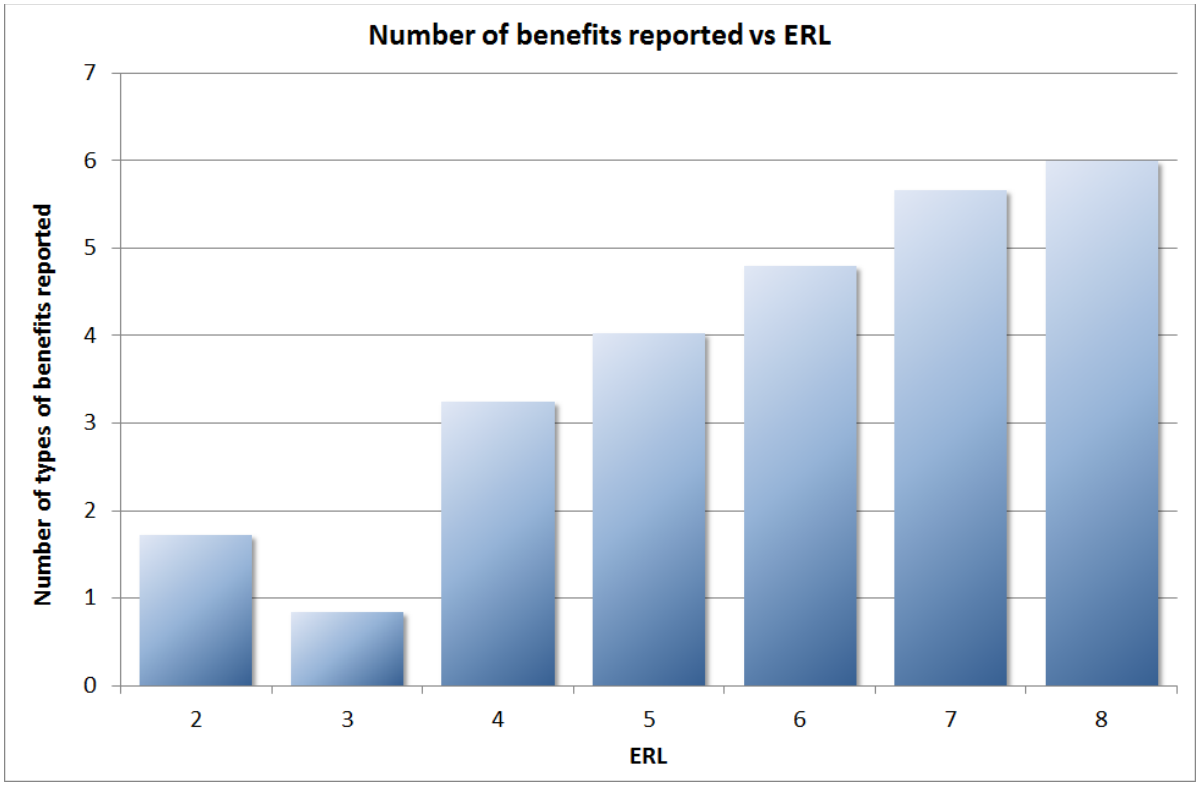

Figure 11: The number of benefits reported vs the ERL score

Conclusions

It would seem that estimate accuracy is no accident. It can be pre-determined from the estimating practices applied when developing and maintaining an estimate. This is good news otherwise industry maturity models, processes and practices would have been invalidated by the research.

Also, maturity is not binary i.e. good or bad, but is analogue, in a range from low to high. We can have a refined understanding of fitness for purpose rather than one-size-fits all approach. This means that estimation activities may be tailored (optimised) depending on the purpose of the estimate and what is available to the estimator.

Although ERL cannot predict by how much an individual project will over/under spend, it can propose that it will lie in a range. If the range is unacceptable, the estimator will need to work towards a higher ERL.

One advantage of using ERL is that it is easy to communicate the maturity of an estimate to the customer. All estimates should be accompanied with its associated ERL to prevent them being used for a purpose other than intended.

The ERL assessment can be used to quantify the maturity of a single estimate or can be used to quantify the maturity of an organisation. We now use the ERL assessment to benchmark teams, projects, departments and the whole business. The weightings used in the ERL assessment ensure that we focus on what matters most.

A company can still be vulnerable to the estimates produced by suppliers. The ERL assessment can be completed with the supplier (or on behalf of the supplier) to understand any residual risks in the supplier estimate.

We were not looking for perfection but progress. The ERL assessment will require further clarifications as we refine our understanding of the key Factors for success. However the ERL assessment has introduced the concept of estimate maturity. Calibrated or uncalebrated this has proven to be valuable.

Whist this paper illustrates the Factors that affect Rolls-Royce estimates, it does not propose that these same Factors will affect other organisations. Broader research should be repeated, crossing multiple companies and domains. It is then possible to develop a parametric estimating tool to estimate the maturity of estimates across industry. If you would like to be involved, please contact the authors of this paper.

References

BAE Systems (2014) *Cost Estimating within BAE Systems.* UK: Rolls-Royce

Bassford, M. (2012) *Where Do We Get It Wrong (and Why)?.* 'SCAF Annual Conference on Cost Estimating ' held on 17th March 2012 at London. London: SCAF

Boardman, C. (2013) *Cost Really Does Matter* 'SCAF Annual Conference' held on 17th September 2013 at London. UK: SCAF

Evans M. (2005) 'Overdue and over budget, over and over again'. *The Economist*

GAO (2009) GAO Cost Estimating and Assessment Guide: Best Practices for Developing and Managing Capital Program Costs [online] available from <http://www.gao.gov/new.items/d093sp.pdf>

Greves, D. and Joumier, H. (2003) Cost Engineering for Cost-Effective Space Programmes. ESA Bulletin 115 [online] available from <http://www.esa.int/esapub/bulletin/bullet115/chapter11_bul115.pdf>

Hammer, M., Champy J. (2003) Reengineering the corporation: a manifesto for business revolution. New York: Harper Collins

Nikitrina, N., Kajko-Mattsson, M., and Nolan, A. (2015) *Success is no Accident: The Checklist Contributing to Successful Process Improvements.* Stockholm: KTH Royal Institute of Technology

Roush, W. R (2005) 'deficit = D inflation'. *Technology Review*

The Standish Group (2013*) CHAOS Manifesto 2013* [online] available from <http://www.versionone.com/assets/img/files/ChaosManifesto2013.pdf>

Tishler A., et al. (1996) 'Identifying critical success factors in defense development projects: a multivariate analysis'. Technological Forecasting and Social Change 51(2), 151–71

Zhang, L., Lee, K.O., Zhang, Z., Banerjee, P. (2003) 'Critical success factors of enterprise resource planning systems implementation success' in China. (ed.) *Proceedings of the 36th Hawaii international conference on systems sciences.*