Verification of the Adjustment Phase

(Chapter 7 – Software Project Estimation)

Alain Abran

(Tutorial Contribution: Dr. Monica Villavicencio)

Topics covered

1. Introduction
2. The Adjustment phase in the estimation process
3. The bundled approach in current practices
4. Cost drivers as estimation sub models
5. Uncertainty & error propagation

7.1 Introduction

Adjustment factors (cost drivers)

Example of challenges in quantifying properly ‘qualitative’ independent variables and in figuring out their impact quantitatively on the dependent variable:

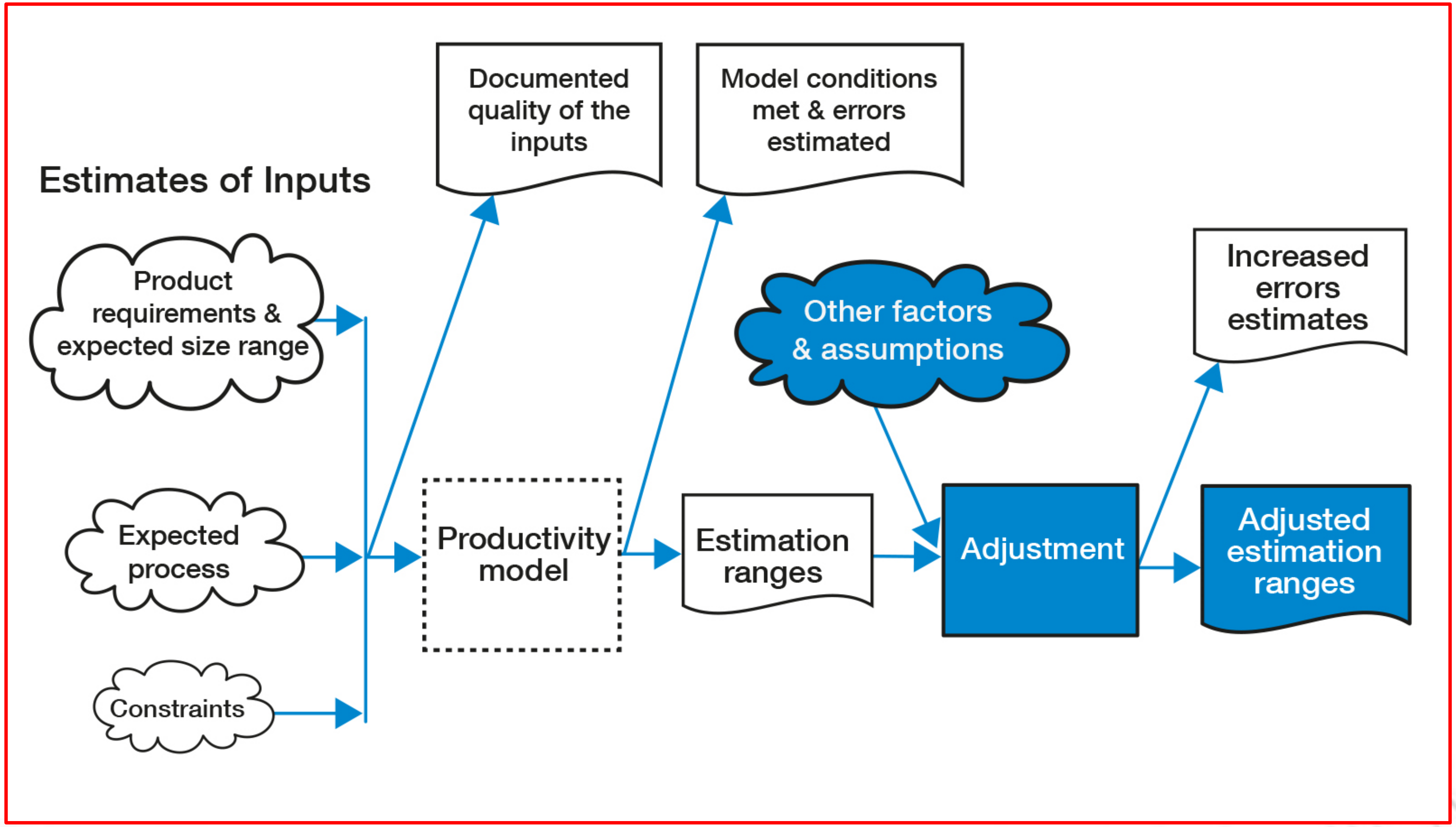
- Staff experience in software development in:
  - specific programming language
  - the database environment
  - the use of support tools for design and testing
  - Etc.

Challenges

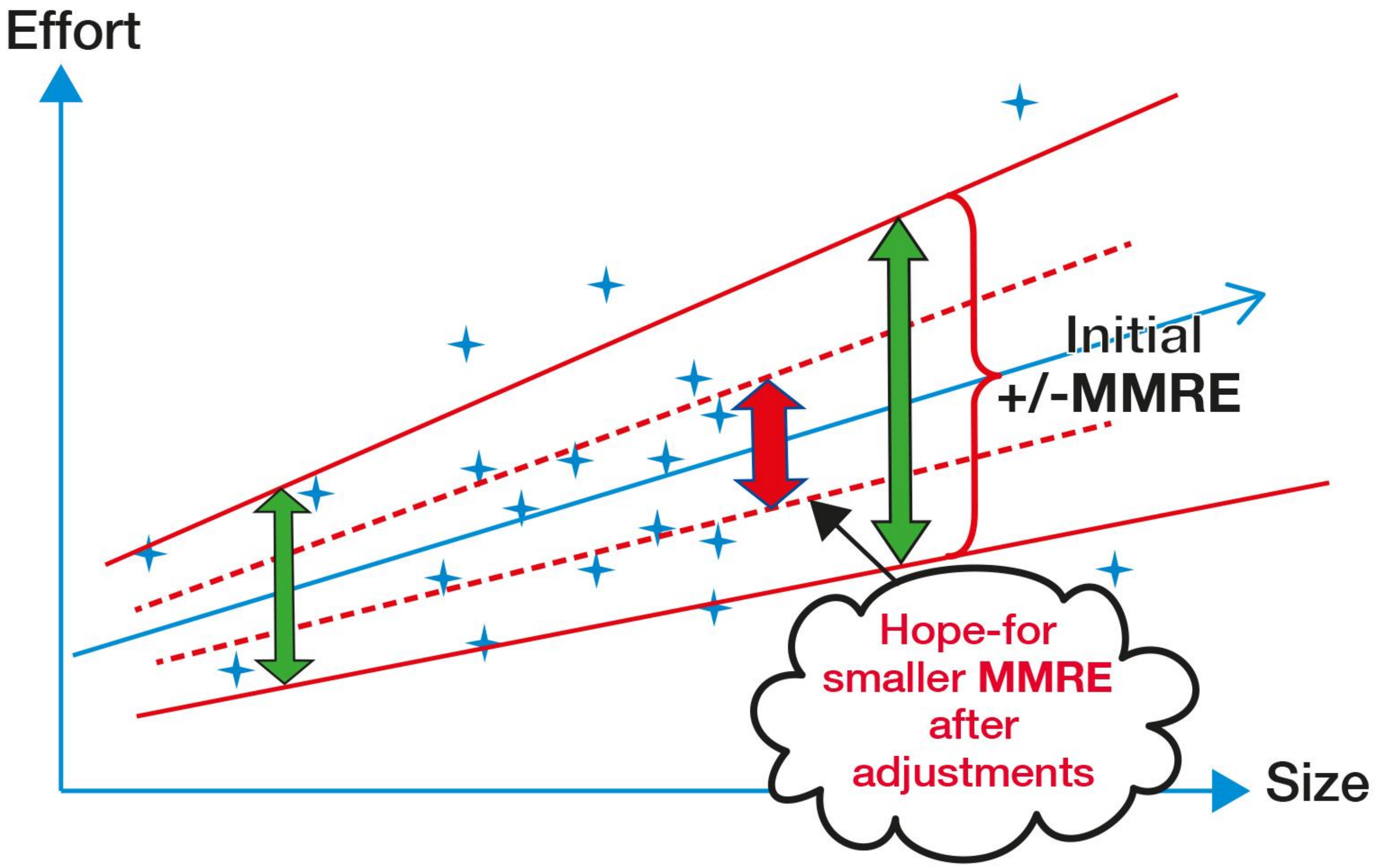
- How do you quantify these cost drivers?
- What is the impact of each of these cost drivers?
- Which ones are the most important in a specific context?
- What is their real contribution in estimation models?
- Does taking them into account using current practices really decrease estimation risks and increase estimation accuracy?

7.2 Adjustment phase in the estimation process

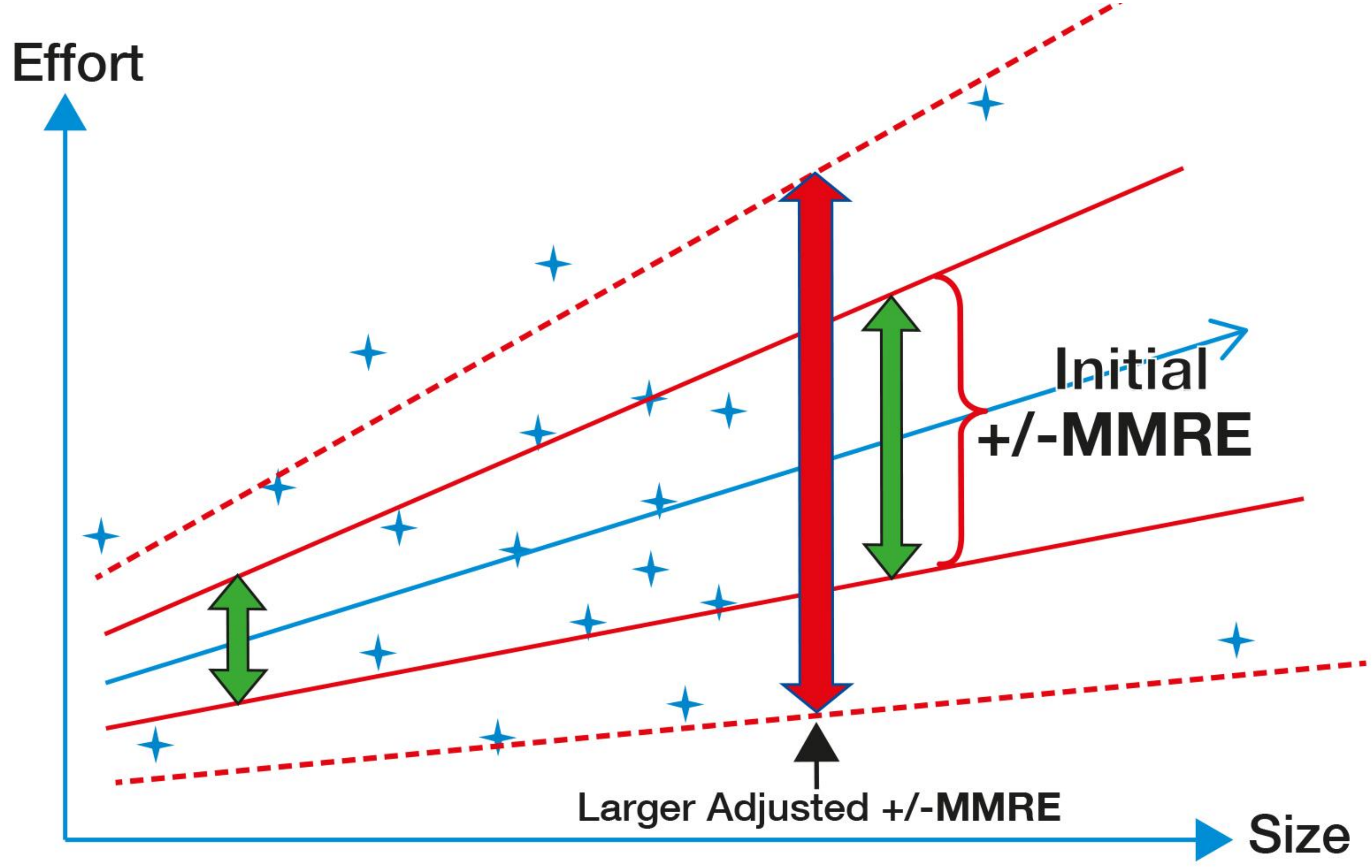
The Adjustment Step



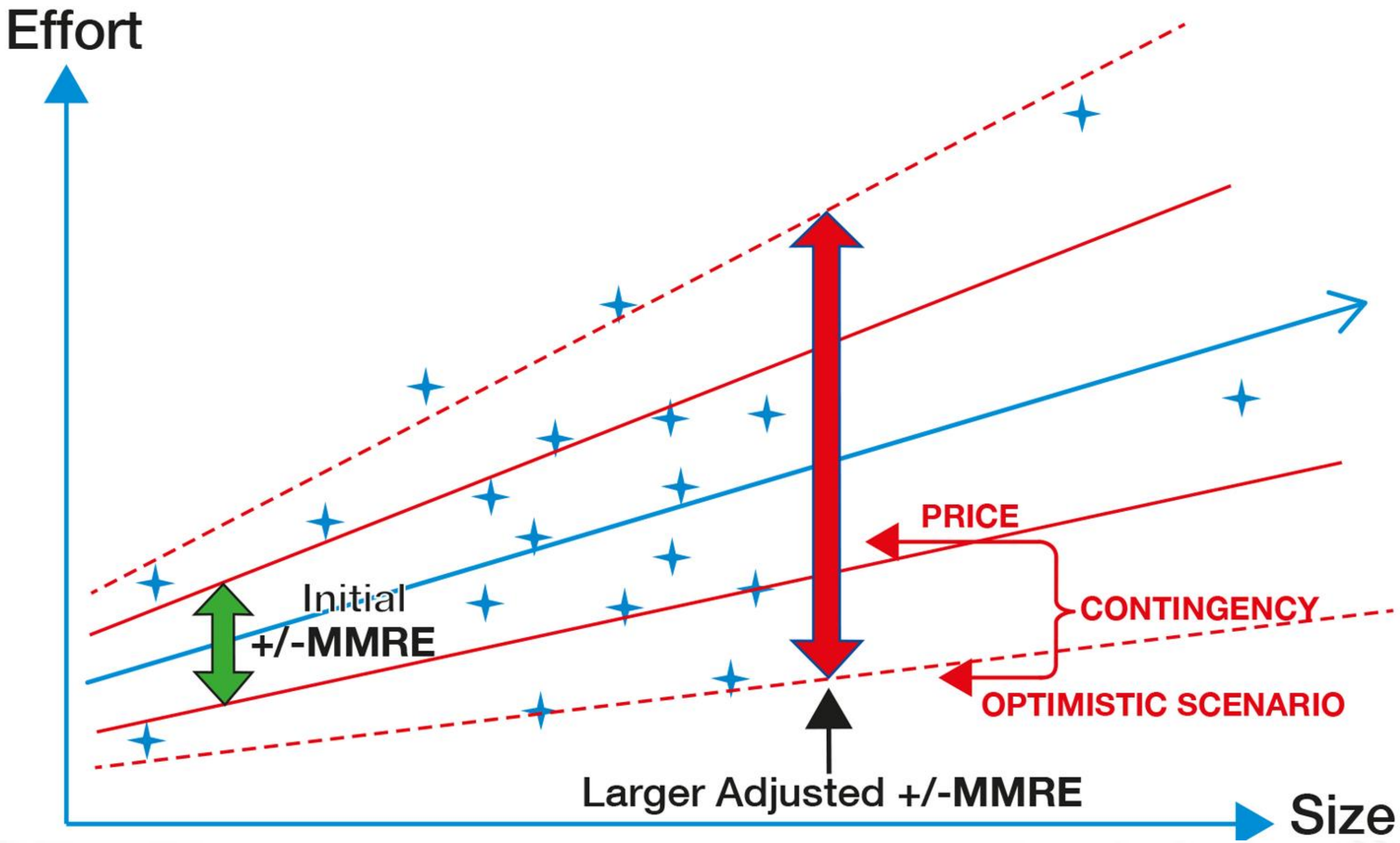
Desired impact
after taking into account multiple costs drivers



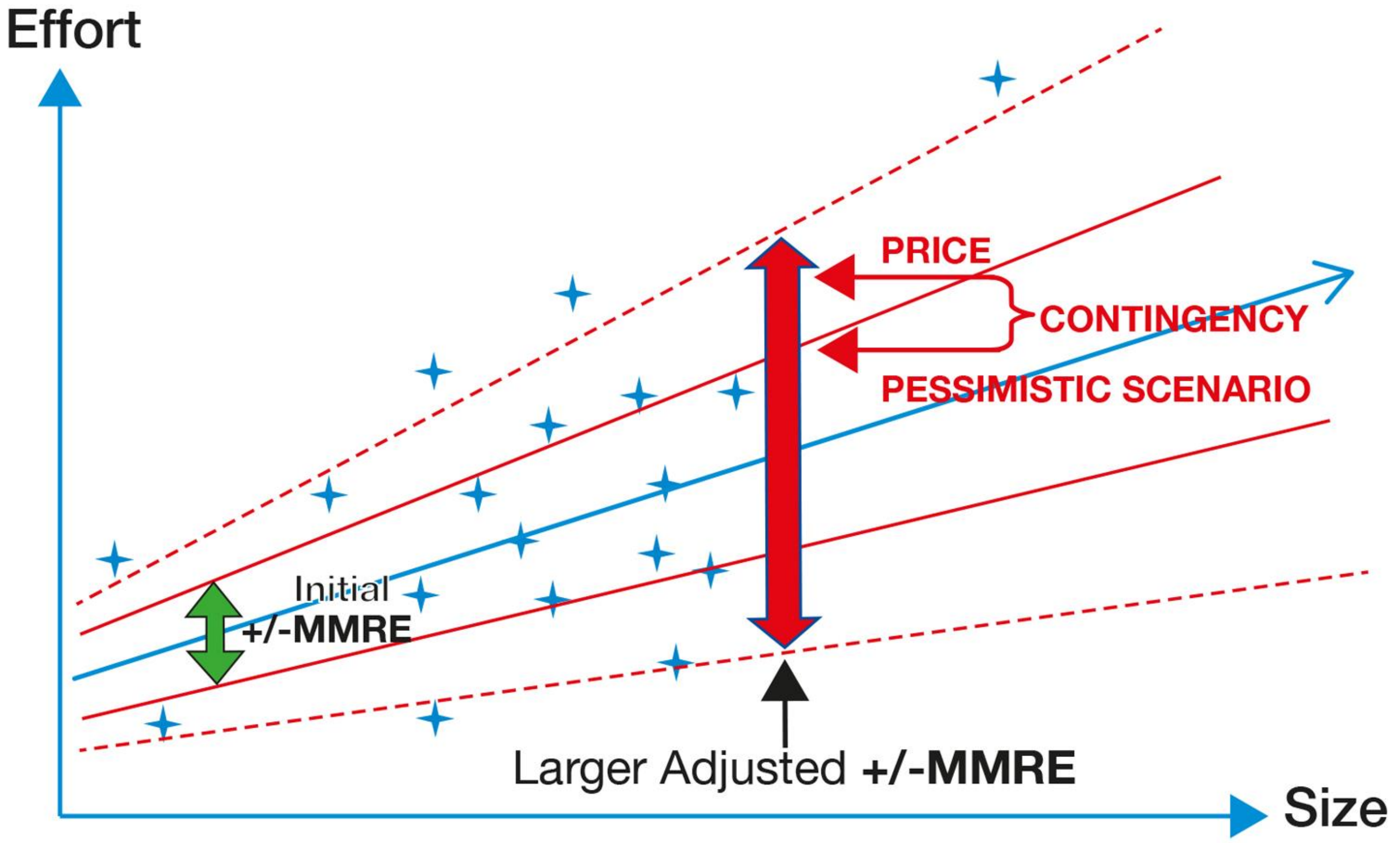
Plausible greater impact of adjustments!



Project budget – optimistic scenario



Project budget – pessimistic scenario



7.3 The bundled Adjustment approach in current estimation practices

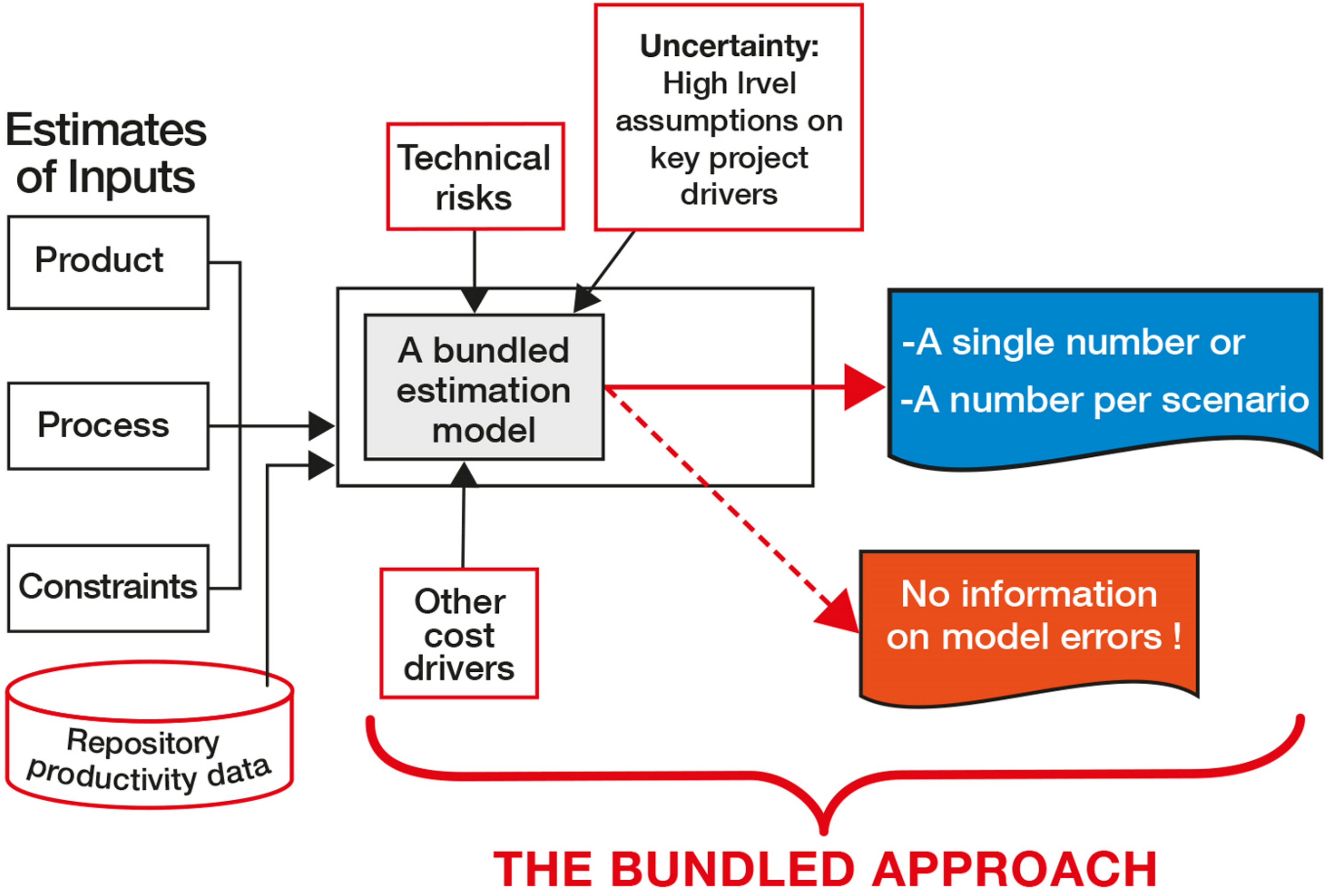
Bundled estimation model

Observed in both:

- Black box estimation models
- White box mathematical models

- Builders of estimation models attempt to include as many cost factors as may be of interest to a client.

Multiple cost drivers are bundled together in an attempt to 'tackle' them concurrently to the estimation model



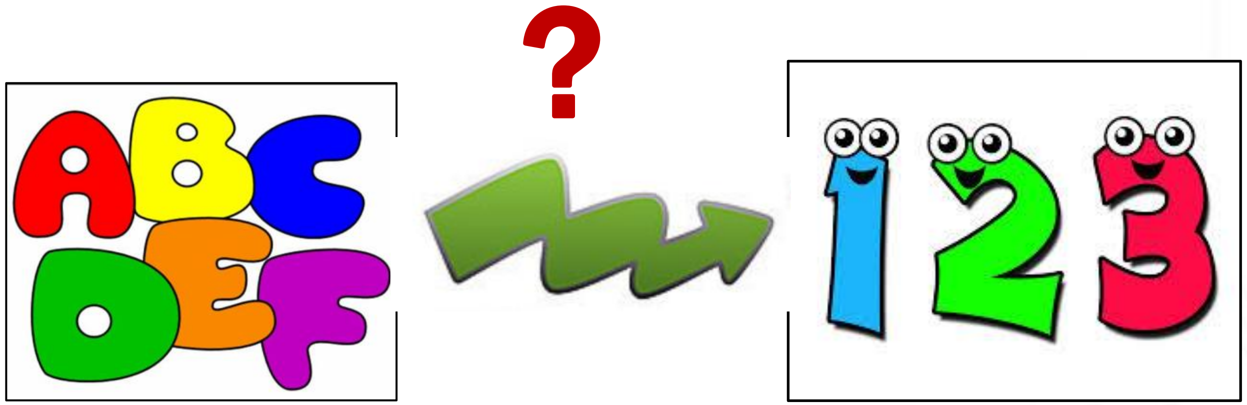
Combining the impact of multiple cost drivers

$$\text{Impact of all cost drivers} = \sum\nolimits^{i}_{n} PF_i$$

$$\text{Effort} = a \times \text{Size} \times \left(\sum\nolimits^{i}_{n} PF_i\right) + b$$

Transformation of a nominal scale type into numbers

Nominal variables → Ordinal categories → Numbers



Categorization of cost drivers: Example

Cost driver	Very Low	Low	Nominal	High	Very High	Ultra High
Project management experience	None	1-4 years	5-10 years	11-15 years	16-25 years	25+ years
Reuse	None	0 to 19%	20-30%	31-50%	50-80%	Over 80%

Assigning impact factors to cost drivers: Example

Effort Driver	Very Low	Low	Nominal (neutral)	High	Very High	Ultra High
F_{1}: Project management experience	+15%	+5%	1	-5%	-10%	-20%
F_{2}: Reuse	+20%	+5%	1	-10%	-20%	-40%

7.4 Cost drivers as estimation sub models

Cost drivers are step functions

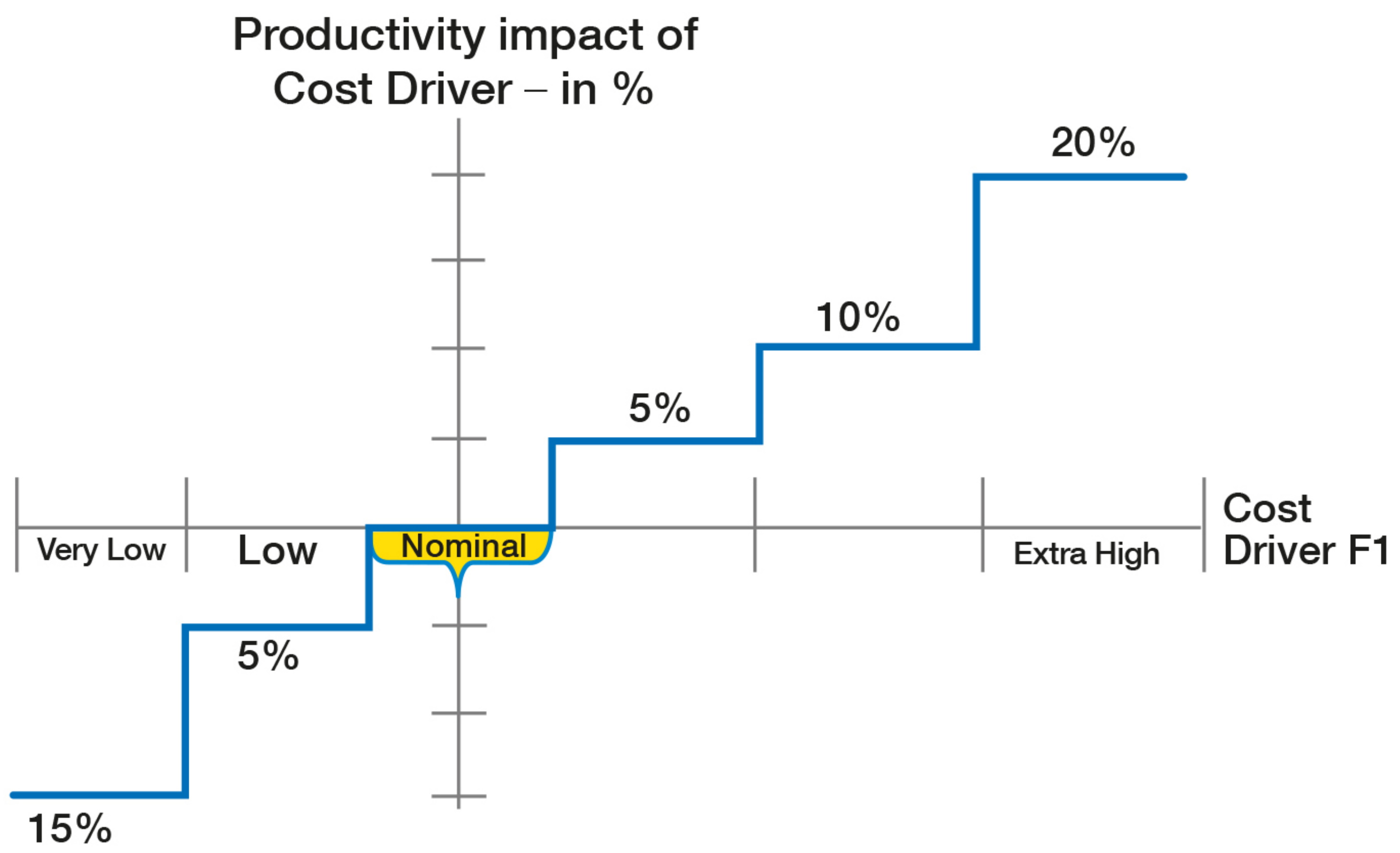
In the bundled approach, the cost drivers are no longer direct inputs to an estimation model:

- each corresponds to a stepwise production function.
- step values are determined in advance by ‘judgments’

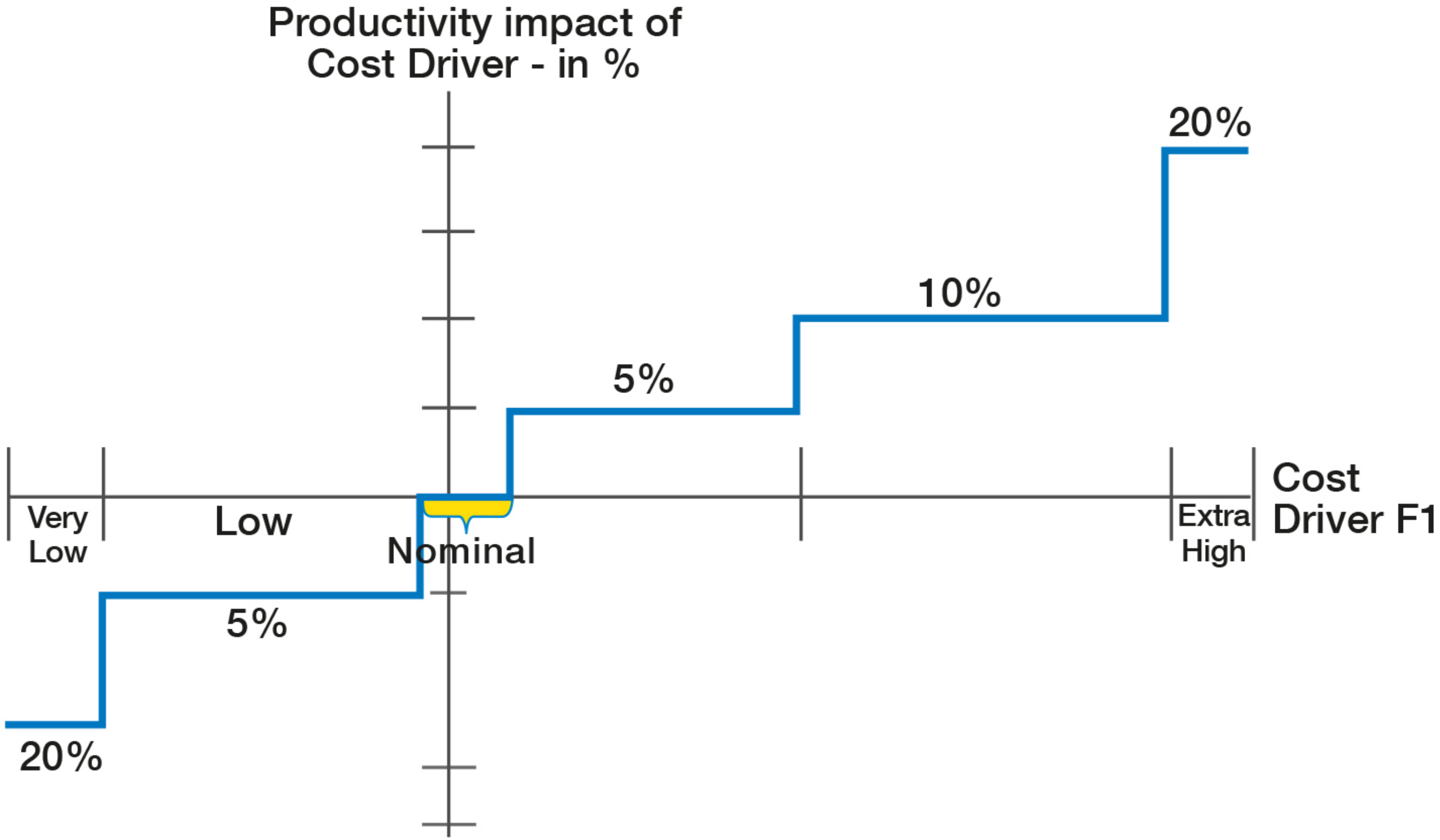
Therefore,

- they are not obtained from a statistical analysis of a dataset.

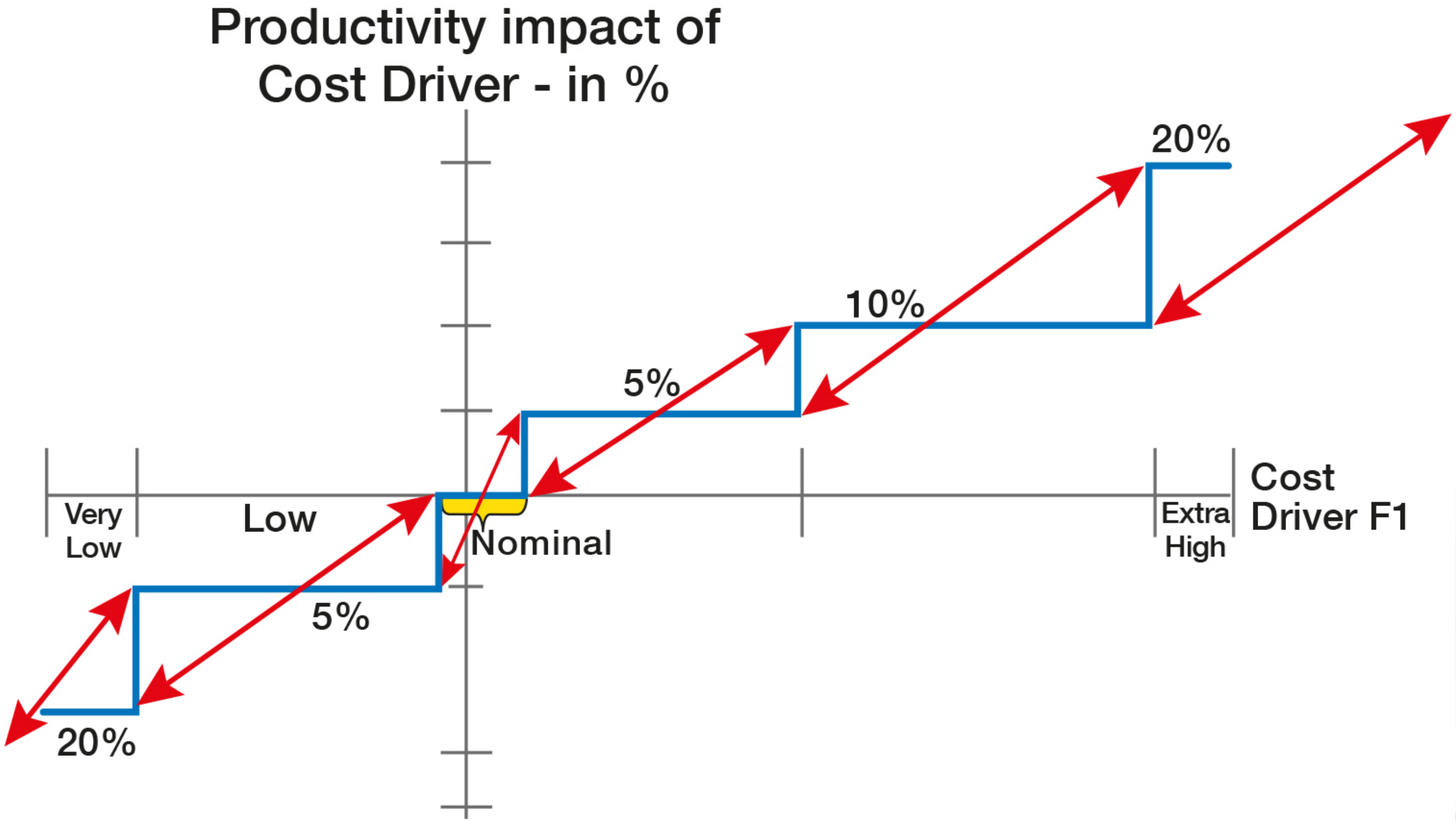
A step function estimation model: Regular intervals



A step function estimation model: Irregular intervals



Approximation of step-function productivity models



Cost drivers as estimation sub models

When an estimation model includes 20 cost drivers built with this approach, this model implicitly has 20 other sub estimation models.

What does this mean?

- The categories selected for each of the cost drivers are estimation sub models (with unknown quality).
- The basis for the transformation is not supported by documented empirical evidence (only by subjective judgment).
- The error range of an impact value is not known.
  - It is not taken into account in the analysis of the errors of the total estimation model.

7.5 Uncertainty an error propagation

Error propagation in mathematical formula

Mathematical Formula	Uncertainty Function
$X = A \pm B$	$(\Delta X)^2 = (\Delta A)^2 + (\Delta B)^2$
$X = cA$	$\Delta X = c\ \Delta A$
$X = c(A \times B)$ or $X = c(A/B)$	$(\Delta X/X)^2 = (\Delta A/A)^2 + (\Delta B/B)^2$
$X = cA^n$	$\Delta X/X = \|n\|\ (\Delta A/A)$
$X = \ln(cA)$	$\Delta X = \Delta A/A$
$X = \exp(A)$	$\Delta X/X = \Delta A$

The relevance of error propagation in models

- The qualities & advantages of any model can be improved or diminished.
- The selection process of the model can be performed, from a theoretical (the form of the method/model) and a practical point of view (application in real cases).

Exercises

1. <u>In your software organization</u>, what is expected when additional cost drivers, uncertainties, and risks are added to models?
2. <u>In engineering</u>, what is expected when additional cost drivers, uncertainties, and risks are added to models?
3. What is the contribution of the Adjustment Phase to decision making in an estimation process?
4. Illustrate how an adjustment phase can help identify an optimistic scenario for decision making.
5. Many models have the following form: $x\ (\sum_{n}^{i} PF_i)$, where PF_i represent the values assigned to the impact of each cost driver. Does the combined impact of these cost drivers help improve the estimate ranges?
6. Cost drivers in many models are often structured as step functions with crisp values for each interval in a cost factor. How is uncertainty taken into account in this structure and use of these cost drivers?

Exercises

7. When you have uncertainty and errors in the variables in input to a model of the linear form, how do you calculate the resulting uncertainty in a model of the form: Effort = a x Size + b?

8. When you have uncertainty and errors in the variables in input to a model of an exponential form, how do you calculate the resulting uncertainty in a model of the form: ?

9. When you have uncertainty and errors in the variables input to a model, how do you calculate the resulting uncertainty in a model of the form: x $(\sum^{i}_{n} PF_i)$?

10. Take the mathematical formula for gravity. In your own words, describe how much time and experimentation was required to quantitatively figure out the relationship among the elements in this formula. Describe how much care is taken by industry to measure the various elements to determine gravity in a specific context (for example, in rocket launches). Next, take the software estimation model used in your

Term Assignments

1. Select one or more of the cost drivers in a model in exponential form in which multiple cost drivers are combined (i.e. bundled models), which are themselves estimation sub models, and document candidate ranges of variation on the one(s) you selected. Then, analyze the impact of these candidate variations on the outcome of the full model.

2. Select two software estimation models (from books or from the Web). Explain how the cost factors included in these models were 'quantified'. Comment on this quantification process.

3. What is the typical approach currently taken to build cost factors into the software estimation models proposed to practitioners: art or engineering? Compare this approach to similar estimation practices in the construction industry, in business, or in medicine. Comment on the software engineering approach.

4. Identify the candidate error range of each cost factor in your organizational estimation model. What would be the rate of error propagation when they are combined in your estimation model?

5. Select one of the software estimation models available free of charge on the Web, and determine its rate of error propagation.

Term Assignments

6. Take one estimation model based on Use Case Points, and determine its rate of error propagation.

7. Take one estimation model based on anecdotal user evidence, and determine its rate of error propagation.

8. Access three software estimation models available on the Web. Document the basis for experimentation provided by the designers of each of these publicly available estimation models. To what point can you trust them? Explain your point of view to your management and customers. Look back at the reasons supporting that point of view and classify them as: engineering-based or opinion-based.

9. Look back at the last three projects you estimated. What assumptions did you document? From what you know now about these completed projects, what assumptions should you have documented?