Software Quality Economics for Combining Defect-Detection Techniques

Stefan Wagner

wagnerst@in.tum.de

Software & Systems Engineering

Technische Universität München

Germany

Outline

1. Motivation
2. Types of Costs
3. Economics Model
4. Characteristics
5. Combination
6. Example
7. Conclusions

Outline

1. *Motivation*
2. Types of Costs
3. Economics Model
4. Characteristics
5. Combination
6. Example
7. Conclusions

Motivation

- Defect-detection techniques (reviews, tests, . . .) are still the most important methods for the improvement of software quality
- Costs for such techniques are significant. According to Myers (1979) 50% of the total development costs are caused by testing. Jones (1987) assigns 30 – 40% to quality assurance and defect removal.
- Optimisation in this area can save money!
- Two possibilities:
  1. Develop new and *better* defect-detection techniques
  2. Use existing techniques in an optimal way

Outline

1. *Motivation*
2. Types of Costs
3. Economics Model
4. Characteristics
5. Combination
6. Example
7. Conclusions

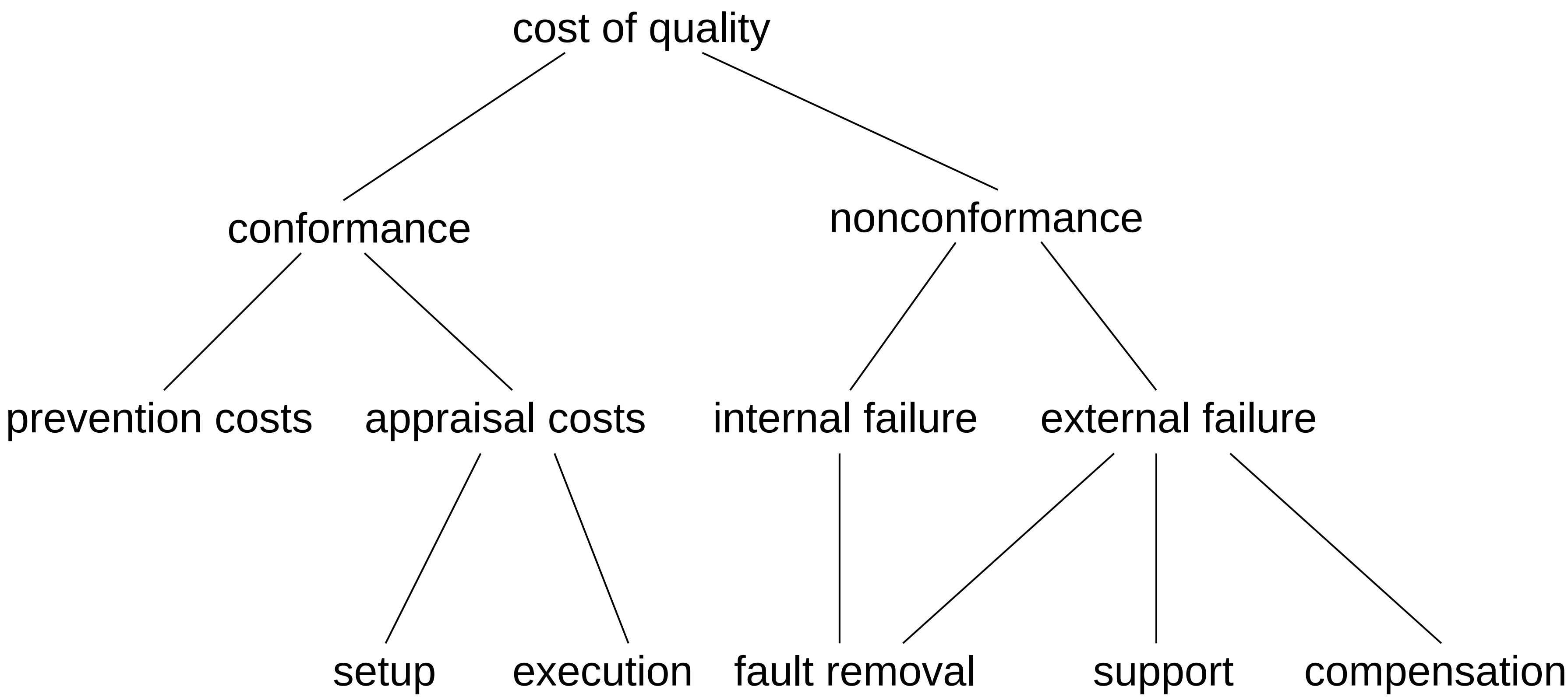
Outline

TECHNISCHE UNIVERSITÄT MÜNCHEN

Quality Costs

- *Quality costs* are all costs associated with
  - preventing,
  - finding,
  - and correcting defective work
- Original models are derived from the manufacturing area
- Division into *conformance* and *nonconformance costs*
- There are derived metrics:
  - Return on Software Quality (ROSQ)
  - Net Present Value of the Software Quality Cash Flows (NPVCF)
  - . . .

Quality Cost Types



Outline

1. Motivation
2. *Types of Costs*
3. Economics Model
4. Characteristics
5. Example
6. Conclusions

Outline

1. Motivation
2. Types of Costs
3. *Economics Model*
4. Characteristics
5. Example
6. Conclusions

TECHNISCHE UNIVERSITÄT MÜNCHEN

Overview

- How can we describe the relationships between the cost types?
- How do we measure them?
- Suitable to compare different defect-detection techniques
- Direct measurements as far as possible
- Estimations on the basis of the MTTF of faults and the severity of its consequences
- Predictions using software reliability models

Directly Measurable Costs

- Costs that can be directly measured or estimated accurately during the usage of a defect-detection technique
- Fixed initial investments: *setup costs* ($c_{\textit{setup}}$)
  - tools,
  - workstations, . . .
- Dynamic part of the appraisal costs: *execution costs* ($c_{\textit{exec}}(p)$)
  - mainly personnel costs
  - hardware usage
- *Fault removal costs* for found faults ($c_{\textit{remv}}(p)$)
- *Direct costs* ($c_{\textit{direct}}$) are the sum for all periods where the defect-detection technique was used.

Prediction Using a Reliability Model

- External failure costs can be predicted
- Reliability models can predict the mean number of experienced failures
- Fraction of a severity class s: $P(s)$
- Average cost of an external failure of severity class s: $c_{\textit{ext}}(s)$
- The *future costs* are then

$$c_{\textit{fut}} = \sum_{i=n}^{u} \frac{\sum_{s=1}^{S} f(i) P(s) c_{\textit{ext}}(s)}{(1+D)^i},$$

where n is the period in which we start the prediction, u is the upper limit of the prediction periods, $f(i)$ is the number of failures in that period, and S is the highest severity class.

Estimation Using Expert Opinion

- Defect-detection Techniques also generate revenues by avoiding failure costs
- They can only be estimated by expert opinion
- Those costs never actually occur
- Saved external failure costs that are saved by finding and removing faults before releasing the software to the customer
- Basis are estimates for each fault
  - mean time to failure (MTTF)
  - severity of failure
- and old cost data for similar failures
- Results in the revenues r

Quality Economics

- Using the metrics established so far, the calculation of the quality economics is straight-forward
- We have
  - the directly measurable costs from the usage,
  - the predicted future costs during operation,
  - and the estimated revenues from saved costs
- Therefore, the *net present value of the cash flows (NPVCF)* is:

$$\mathit{NPVCF} = r - c_{\mathit{direct}} - c_{\mathit{fut}}.$$

Outline

1. Motivation
2. Types of Costs
3. *Economics Model*
4. Characteristics
5. Combination
6. Example
7. Conclusions

Outline

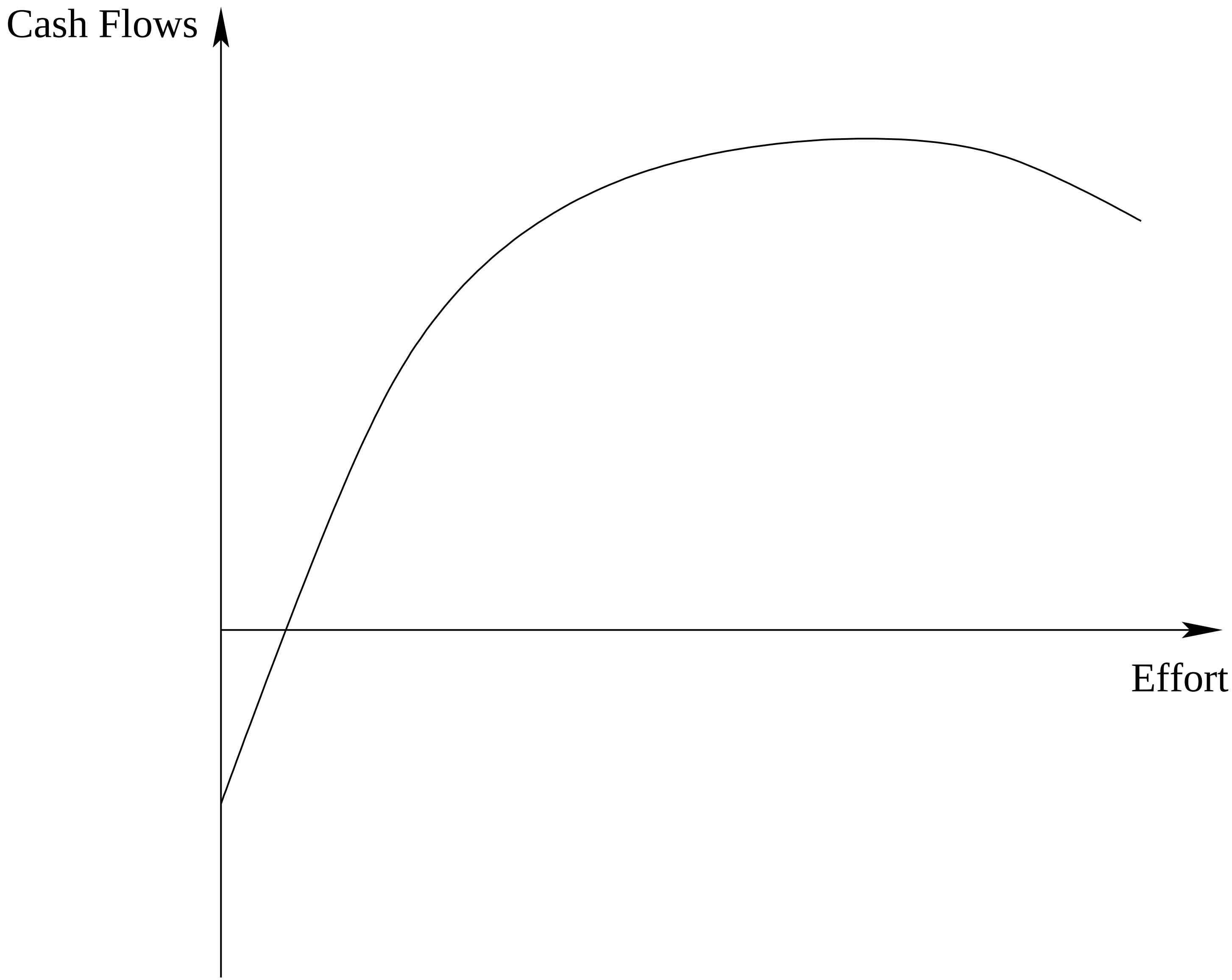
1. Motivation
2. Types of Costs
3. Economics Model
4. *Characteristics*
5. Combination
6. Example
7. Conclusions

TECHNISCHE UNIVERSITÄT MÜNCHEN

Characteristics for a Technique

- How can we characterise a defect-detection technique in general?
- Future costs contain costs for different techniques
- Revenues and direct costs are characteristic cash flows
- Those cash flows in relation to the effort spent for the technique result in a *characteristic curve*
- Starting with negative cash flows from initial investments
- With further effort it becomes positive as faults are found
- Finally, it reaches an area of satisfaction where it becomes more and more difficult to find faults
- Similar to the so-called S-curve of software testing

Characteristic Cash Flows



Problems

- Establishment of such curves for different techniques requires extensive measurement experiments
- Curves from different projects must be normalised because of
  - different software sizes
  - different programming languages
  - different tester experiences
  - . . .
- Investigation on which factors the curves depend are necessary

TECHNISCHE UNIVERSITÄT MÜNCHEN

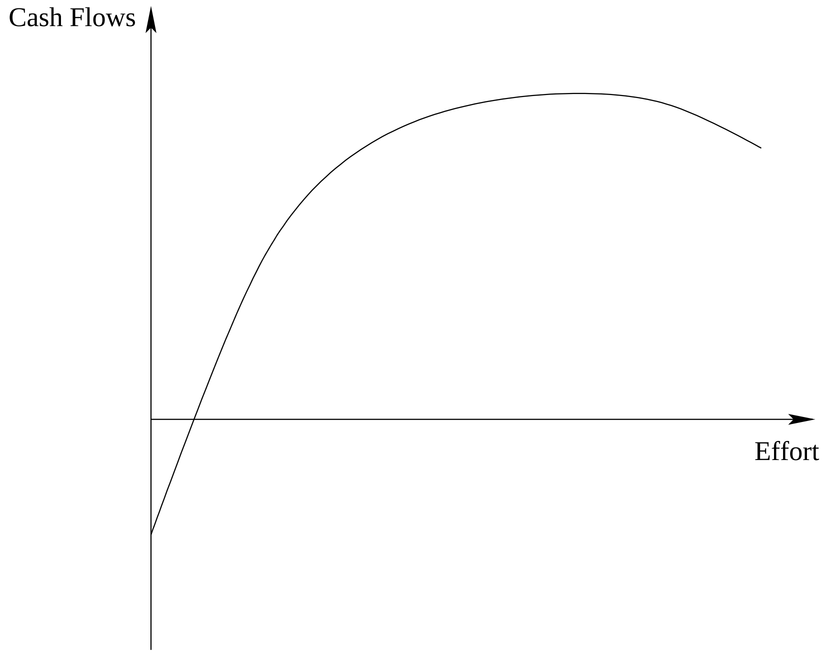
Outline

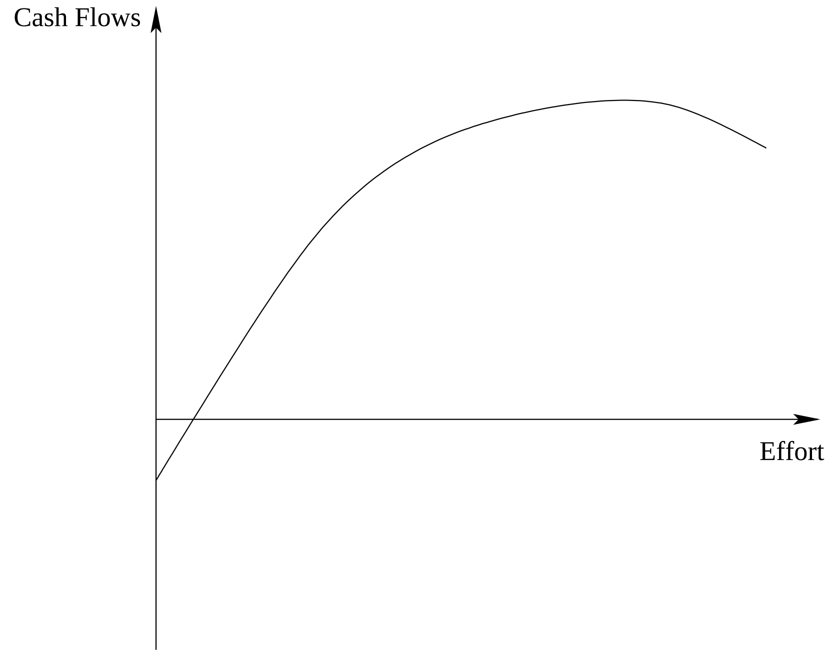
1. Motivation
2. Types of Costs
3. Economics Model
4. *Characteristics*
5. Combination
6. Example
7. Conclusions

Outline

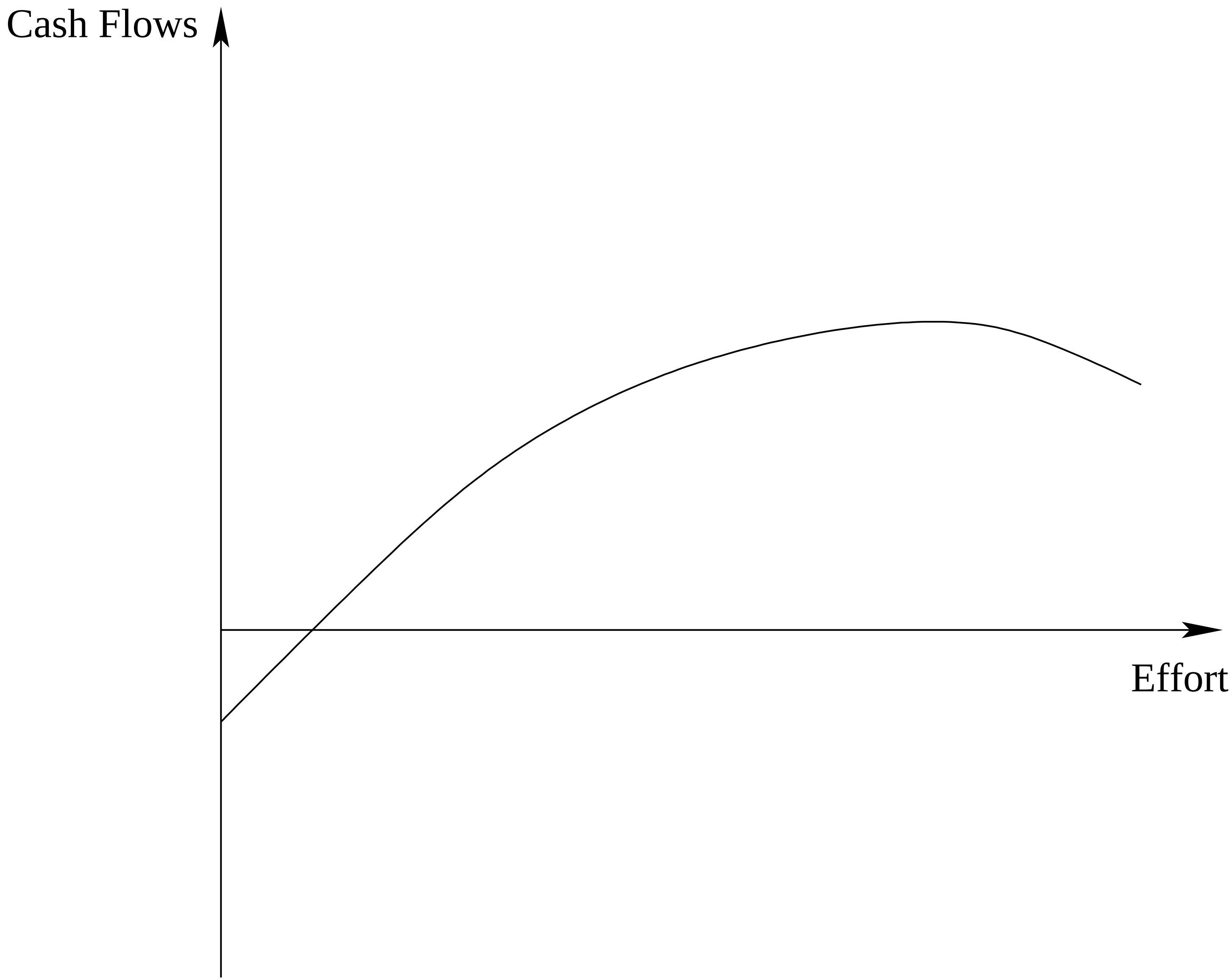
1. Motivation
2. Types of Costs
3. Economics Model
4. Characteristics
5. *Combination*
6. Example
7. Conclusions

Technique A and B

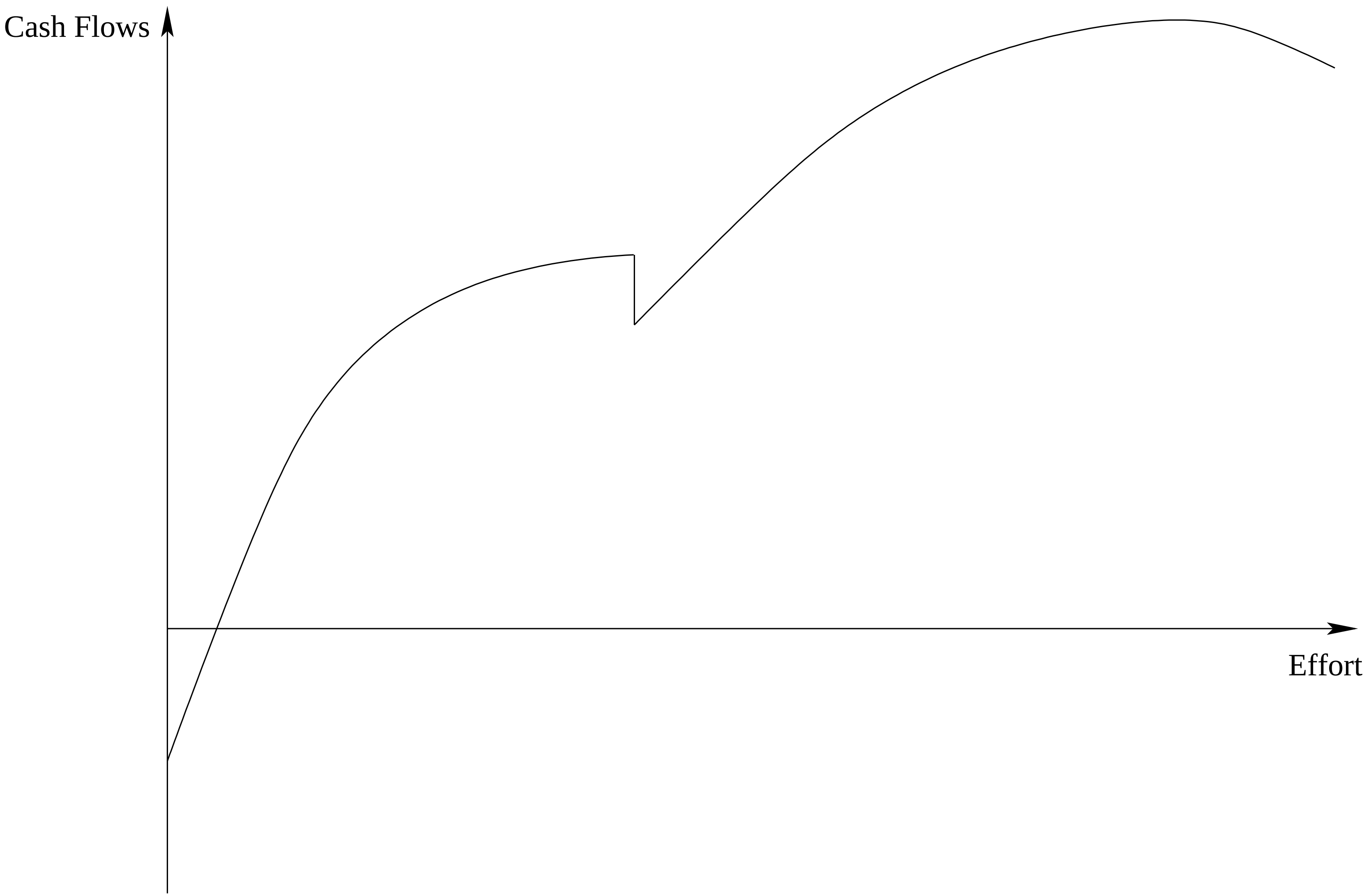




Adjusted B



Combined A and B



Model of Diversity

- Littlewood et al. proposed a model of diversity of defect-detection techniques
- Difficulty function $\theta_A(i)$ that describes difficulty of technique A to find the fault i.
- It varies from one fault to another
- The ineffectiveness of a techique A is then:

$$P(A \text{ fails to detect a randomly chosen fault}) = \sum_i p_i^* \cdot \theta_A(i) = E_{p^*}(\theta_A(i)),$$

- p_i^*: probability distribution of the faults to be present
- E_{p^*}: the mean obtained with respect to this probability distribution.

Model of Diversity — Cont.

- Proportion of faults that are found by a technique:

$$E_{p^*}(1 - \theta_A(i)) = \phi_A$$

- Proportion of the faults that could be found by both techniques:

$$E_{p^*}((1 - \theta_A(i)) \cdot (1 - \theta_B(i))) = \phi_{AB}$$

- The result depends on the covariance of the difficulty functions of the techniques.
- A negative covariance means the techniques are strongly diverse and the effect will be small.
- A positive covariance makes the effect stronger.

Combination Function

- The diversity model can be used as rough approximation in the function *comb*

$$\textit{comb}(T, E, m) = \frac{m(\phi_T - \sum_{e \in E} \phi_{Te})}{\phi_T}$$

- m: monetary value from a technique T
- E: set of earlier used techniques

Reverse Function

- So far only the combination of isolated techniques can be estimated
- If we had data about the combination of two techniques, we would be interested in finding out how the techniques would have behaved in isolation.
- For this we define the reverse function:

$$\overline{\textit{comb}}(T, E, m) = \frac{m \cdot \phi_T}{\phi_T - \sum_{e \in E} \phi_{Te}}$$

Economics Model

- Two parts are affected by the combination

$$r' = \mathit{comb}(T, E, r)$$
$$c'_{\mathit{remv}} = \mathit{comb}(T, E, c_{\mathit{remv}})$$

- The new definitions can be substituted for the old ones
- The definitions using the reverse function $\overline{\mathit{comb}}$ are accordingly

Outline

1. Motivation
2. Types of Costs
3. Economics Model
4. Characteristics
5. *Combination*
6. Example
7. Conclusions

TECHNISCHE UNIVERSITÄT MÜNCHEN

Outline

1. Motivation
2. Types of Costs
3. Economics Model
4. Characteristics
5. Combination
6. *Example*
7. Conclusions

TECHNISCHE UNIVERSITÄT MÜNCHEN

Example

- No application to real project data
- Using example from Littlewood et al. and hypothetical cost data
- Railway signaling system
- Code checking (C)
  - $c_{setup} = 1000$
  - $c_{remv} = 200$
  - $c_{exec} = 1000$
  - $r = 3000$

Example — Cont.

- Testing (T)
  - $c_{\mathit{setup}} = 2000$
  - $c_{\mathit{remv}} = 500$
  - $c_{\mathit{exec}} = 800$
  - $r = 5000$
- $c_{\mathit{fut}} = 2000$ for both cases
- $\mathit{NPVCF}_C = -1200$
- $\mathit{NPVCF}_T = -500$
- Single applications are not profitable
- Naive combination would give $\mathit{NPVCF} = 500$

Diversity

Fault id	Checking	Testing
F11	x	-
F12	-	-
F13	-	x
F14	x	-
F15	x	-
F16	-	x
F17	-	x
F18	x	x

- $\phi_T = .52$
- $\phi_{CT} = .28$

Results

- Then the adjusted values are

$$r' = \textit{comb}(T, \{C\}, r) = \frac{5000 \cdot (.52 - .28)}{.52} = 2,293.05$$

$$c'_{\textit{remv}} = \textit{comb}(T, \{C\}, c_{\textit{remv}}) = \frac{200 \cdot (.52 - .28)}{.52} = 91.72$$

- The adjusted NPVCF is than $-1,798.67$
- The unadjusted value gives obviously a value far too high

Outline

1. Motivation
2. Types of Costs
3. Economics Model
4. Characteristics
5. Combination
6. *Example*
7. Conclusions

TECHNISCHE UNIVERSITÄT MÜNCHEN

Outline

1. Motivation
2. Types of Costs
3. Economics Model
4. Characteristics
5. Combination
6. Example
7. *Conclusions*

TECHNISCHE UNIVERSITÄT MÜNCHEN

Conclusions

- Model of quality economics for defect-detection techniques based on
  - direct measurement,
  - prediction using software reliability models,
  - and estimation using expert opinion
- Concept of a characteristic curve for a technique
- Incorporation of the effects of combination
- Future work includes
  - Incorporation of maintenance costs
  - Quantification of lost sales, annoyed customers, . . .
  - Application to concrete defect-detection techniques