# Bridge the Gap between Software Test Process and Business Value: a Case Study

**Qi Li, Barry Boehm**

**University of Southern California-CSSE**

**Ye Yang, Qing Wang**

**Institute of Software, Chinese Academy of Sciences**

**ICSP, Vancouver, May 17th 2009**

# Motivation

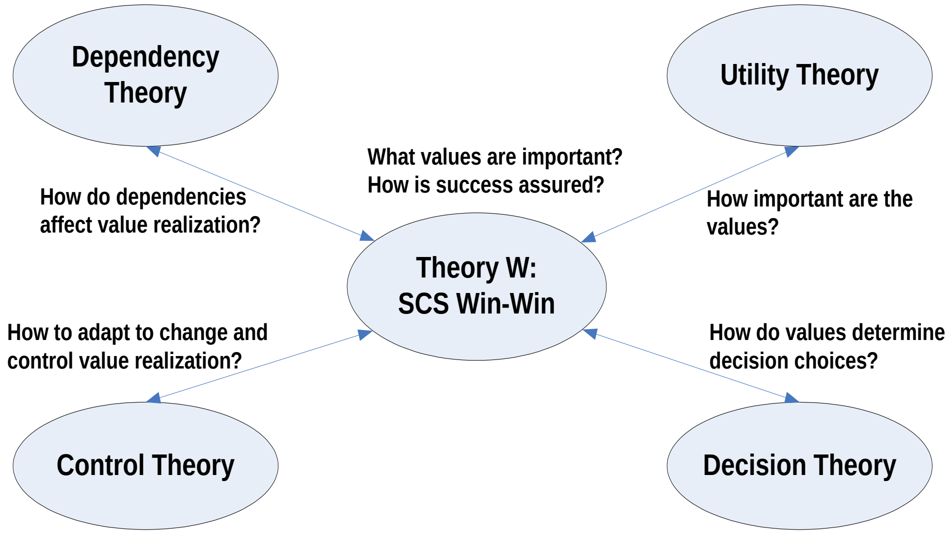
- Software testing is one of the most labor-intensive activities in software development life cycle and consumes 30% - 50% of total development effort
- Traditional testing methodologies are often value neutral
- In practice, 80% of the value often comes from 20% of the software
- Under the market pressure, the testing effort are usually compressed without satisfying the required quality requirement, it is like that we could only "Pick Two" from cost, quality, and schedule

# Motivation (Cont.)

- How to improve the testing efficiency
- Maximize the return on investment (ROI) of testing activities, and improve client's satisfaction
- We propose a value-based software testing method to better link the internal test process with business objectives from customers and market at the system and acceptance testing level
- Prioritize and re-scope which features to test first and to what degree, and in this way we can "Pick Three" from cost, schedule and quality

# Related Work

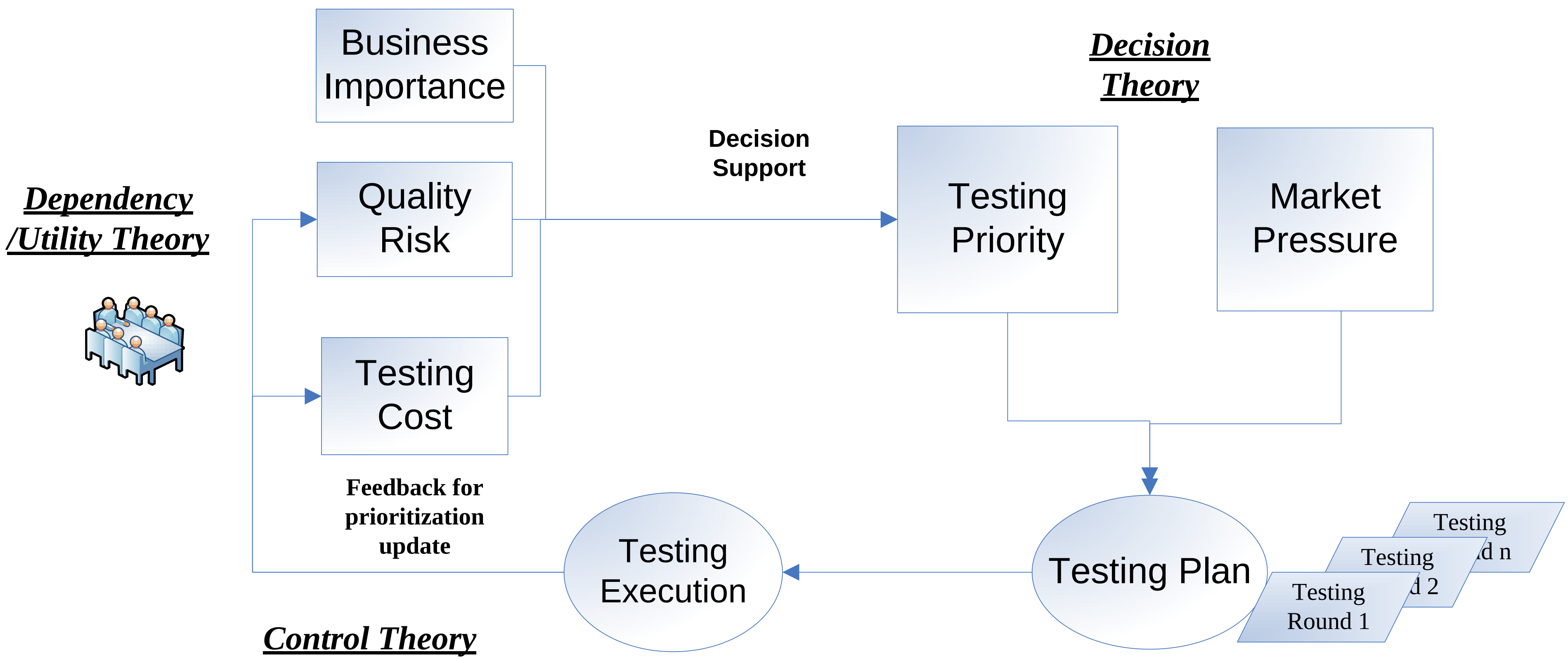
- **The "4+1" Theory of VBSE:**




- **Stale Amland, a risk-based testing approach**
- **Rudolf Ramler, a framework for value-based test management**
- **Liguo Huang, a quantitative risk analysis which helps determine when to stop testing based on COCOMO II and COQUALMO**

# Method Overview

$$ValuePriority = \frac{BI * Probability}{Cost}$$



# Answer Three Questions

- How to calculate priorities for Features?
  - Business Importance (BI)
  - Quality Risk Probabilities
  - Testing Cost
- How to deal with priorities?
- What is the benefit from this method?

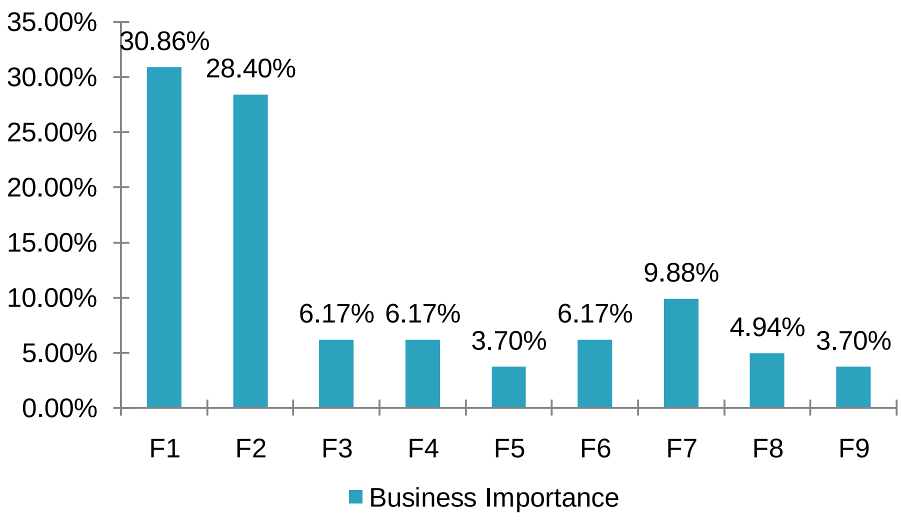
# Case Study Background

- **Organization**:
  - An R&D organization at ISCAS in China
  - Appraised and rated at CMMI maturity level 4 in 2005
  - A research group focused on software process improvement
- **Project**:
  - SoftPM: a tool used to manage software project and has been deployed to many software organizations in China
  - Iterative development method and CMMI process management.
  - Whole size until the case study (the year 2007): 553 KLOC Java codes
- **Case Study:**
  - The case study is based on Version 3.1 that covers 9 new features with an overall size of 32.6 KLOC Java codes
  - Apply this method in this project's system testing process, manage and control its testing process with the help of defect tracking system of SoftPM

# Business Importance (BI)

- Use Karl Wieger's Method to get the relative BI for each feature
- $BI_i = W_{Benefit} * Benifit_i + W_{Penalty} * Penalty_i$

| | Business Importance | | | |
|---|---|---|---|---|
| | Benefit | Penalty | Total BI | BI % |
| Weights | 2 | 1 | | |
| F1 | 9 | 7 | 25 | 30.9% |
| F2 | 8 | 7 | 23 | 28.4% |
| F3 | 1 | 3 | 5 | 6.2% |
| F4 | 2 | 1 | 5 | 6.2% |
| F5 | 1 | 1 | 3 | 3.7% |
| F6 | 2 | 1 | 5 | 6.2% |
| F7 | 3 | 2 | 8 | 9.9% |
| F8 | 1 | 2 | 4 | 4.9% |
| F9 | 1 | 1 | 3 | 3.7% |
| SUM | 28 | 25 | 81 | 1 |



# Quality Risks & Weights Allocation

| | Personnel Proficiency | Size | Complexity | Design Quality | Defects Proportion | Defects Density | Weights |
|---|---|---|---|---|---|---|---|
| Personnel Proficiency | 1 | 1/3 | 3 | 3 | 1/3 | 1/5 | 0.09 |
| Size | 3 | 1 | 3 | 3 | 1 | 1 | 0.19 |
| Complexity | 1/3 | 1/9 | 1 | 1 | 1/7 | 1/9 | 0.03 |
| Design Quality | 1/3 | 1/7 | 1 | 1 | 1/7 | 1/9 | 0.04 |
| Defects Proportion | 3 | 1 | 7 | 7 | 1 | 1 | 0.27 |
| Defects Density | 5 | 3 | 9 | 9 | 1 | 1 | 0.38 |

**Analytical Hierarchy Process (AHP)**

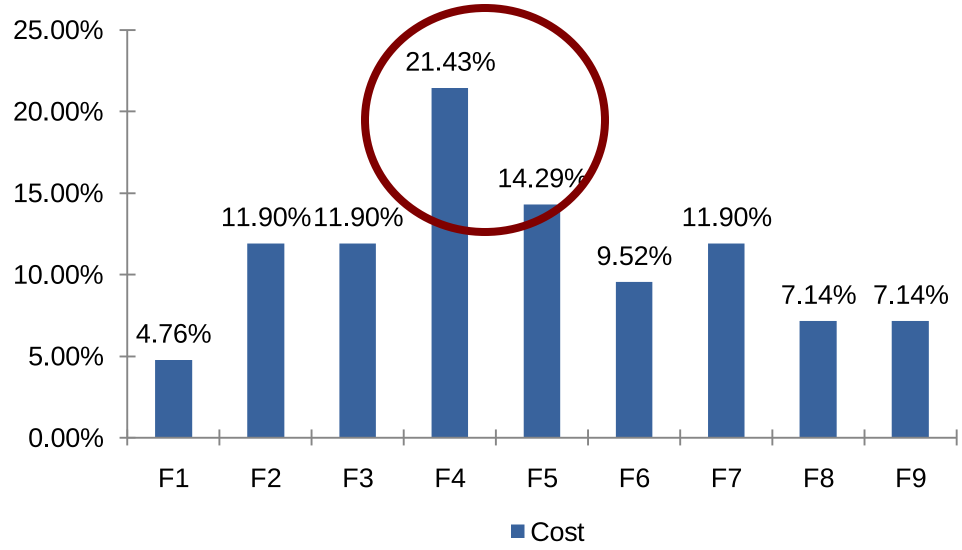
# Quality Risks Probability

$$P_i = \frac{\sum_{j=1}^{n} R_{i,j} * W_j}{9}$$

| | Initial Risks | | | | Feedback Risks | | Probability |
|---|---|---|---|---|---|---|---|
| | Personnel Proficiency | Size | Complexity | Design Quality | Defects Proportion | Defects Density | |
| weights | 0.09 | 0.19 | 0.03 | 0.04 | 0.27 | 0.38 | |
| F1 | 5 | 3 | 1 | 1 | | | 0.13 |
| F2 | 4 | 9 | 5 | 2 | | | 0.26 |
| F3 | 3 | 3 | 5 | 5 | | | 0.14 |
| F4 | 5 | 4 | 7 | 5 | | | 0.19 |
| F5 | 5 | 2 | 3 | 3 | | | 0.12 |
| F6 | 5 | 2 | 5 | 6 | | | 0.14 |
| F7 | 5 | 4 | 5 | 2 | | | 0.17 |
| F8 | 1 | 2 | 1 | 1 | | | 0.06 |
| F9 | 1 | 1 | 1 | 1 | | | 0.04 |

# Testing Cost

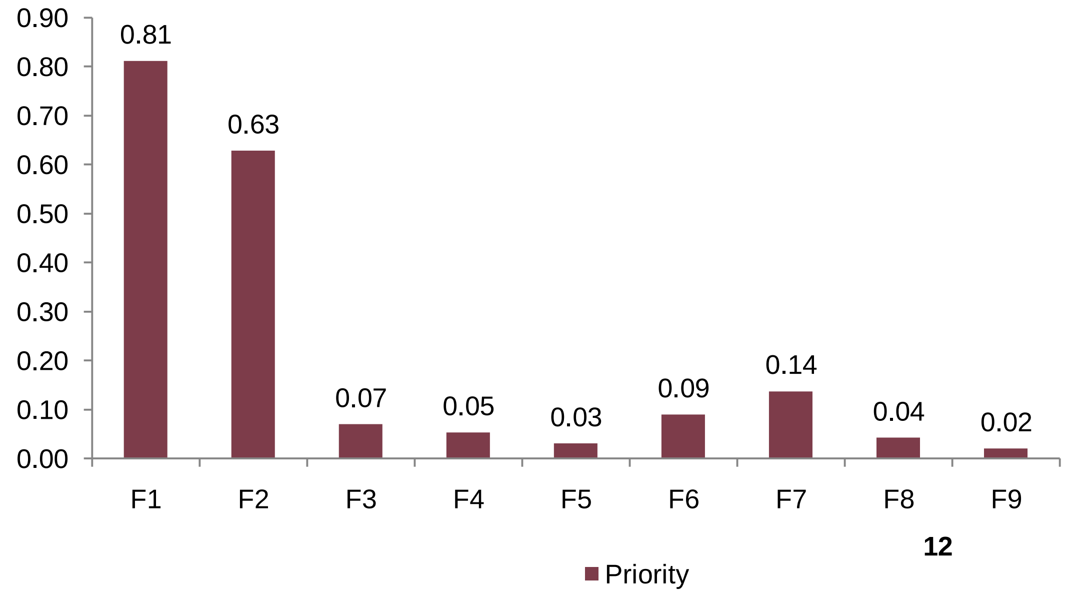
| | Cost | Cost% |
|---|---|---|
| F1 | 2 | 4.8% |
| F2 | 5 | 11.9% |
| F3 | 5 | 11.9% |
| F4 | 9 | 21.4% |
| F5 | 6 | 14.3% |
| F6 | 4 | 9.5% |
| F7 | 5 | 11.9% |
| F8 | 3 | 7.1% |
| F9 | 3 | 7.1% |
| SUM | 42 | 1 |



# Put the three together: Value Priority

| | BI % | Probability | Cost% | Value Priority |
|---|---|---|---|---|
| F1 | 31 | 0.13 | 5 | 0.81 |
| F2 | 28 | 0.26 | 12 | 0.63 |
| F7 | 10 | 0.17 | 12 | 0.14 |
| F6 | 6 | 0.14 | 10 | 0.09 |
| F3 | 6 | 0.14 | 12 | 0.07 |
| F4 | 6 | 0.19 | 21 | 0.05 |
| F8 | 5 | 0.06 | 7 | 0.04 |
| F5 | 4 | 0.12 | 14 | 0.03 |
| F9 | 4 | 0.04 | 7 | 0.02 |

$$Value\ Priority = \frac{BI * Probability}{Cost}$$



# How to Deal with Priorities?

- For high-priority features:
  - Assign the most skilled testers; Put more effort and more thoroughly testing at the early stage of testing;
  - Make high-priority features satisfy the stop-test criteria first, in our case study, the stop-test criteria are:
    - Testing cases coverage rate is 100%
    - All planned test cases executed and passed
    - No non-trivial defects are detected during at least one day of continuous testing

# Adjust Priority According to Feedback during Testing Process

| | Initial Risks | | | | Feedback Risks | | Probability |
|---|---|---|---|---|---|---|---|
| | Personnel Proficiency | Size | Complexity | Design Quality | Defects Proportion | Defects Density | |
| weights | 0.09 | 0.19 | 0.03 | 0.04 | 0.27 | 0.38 | |
| F3 | 3 | 3 | 5 | 5 | 2 | 1 | 0.24 |
| F4 | 5 | 4 | 7 | 5 | 7 | 4 | 0.55 |
| F5 | 5 | 2 | 3 | 3 | 8 | 9 | 0.74 |
| F6 | 5 | 2 | 5 | 6 | 3 | 3 | 0.35 |
| F7 | 5 | 4 | 5 | 2 | 3 | 2 | 0.33 |
| F8 | 1 | 2 | 1 | 1 | 1 | 1 | 0.13 |
| F9 | 1 | 1 | 1 | 1 | 1 | 1 | 0.09 |

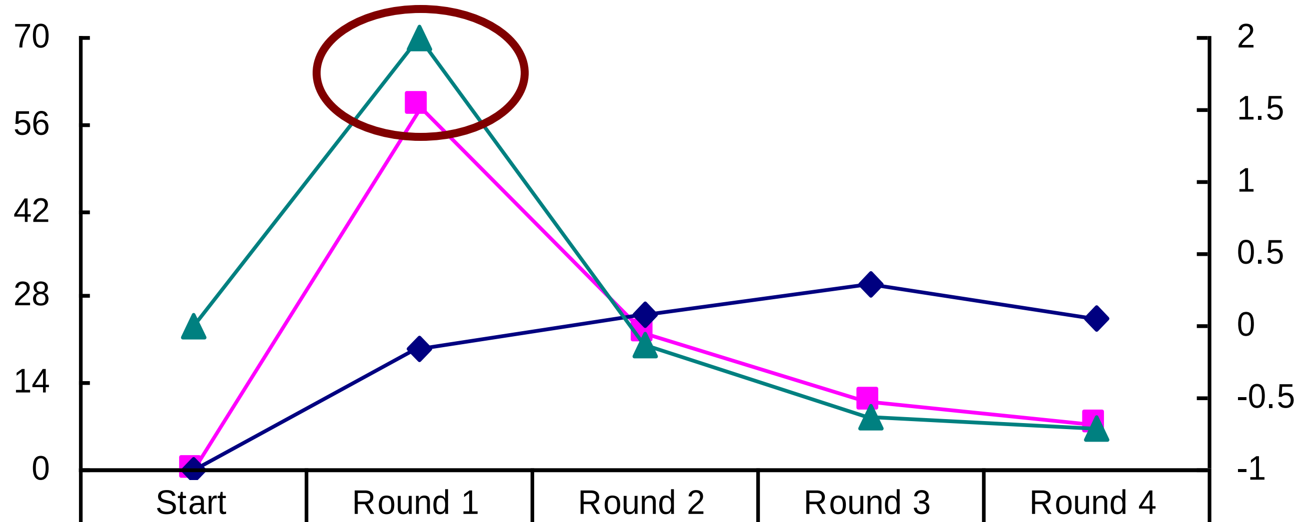
**Re-calculate priority according to updated quality risk probability and testing cost estimation at the end of each test round**

# Controlled Process Results

- 4 Testing rounds, 7-10 days for each round

| Testing Round | Features satisfied with stop-test criteria |
| --- | --- |
| 1 | F1, F2 |
| 2 | F3, F6, F7 |
| 3 | F4, F8 |
| 4 | F5, F9 |

# Performance Evaluation

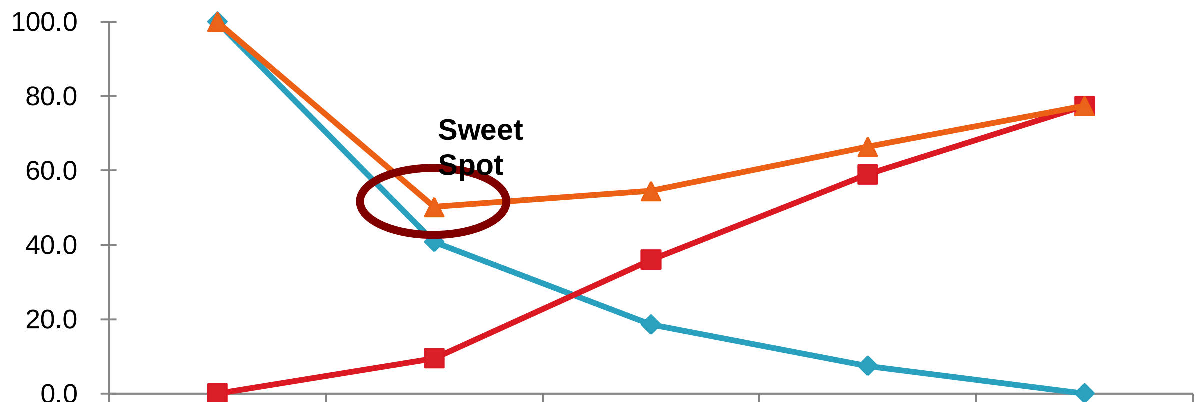



| | Start | Round 1 | Round 2 | Round 3 | Round 4 |
|---|---|---|---|---|---|
| BI Earned | 0 | 59.2 | 22.2 | 11.1 | 7.4 |
| Cost | 0 | 19.8 | 25.3 | 30.3 | 24.6 |
| ROI | 0 | 1.99 | -0.12 | -0.63 | -0.70 |

- **Testing value realized when features satisfy the stop-test criteria**
- **Achieving testing ROI peak at the early stage of testing**
- **Earn most business importance at the early stage of testing**
- **Especially useful when the market pressure is high, which mean time for testing is very limited**

# How Much Testing is Enough

- **Market erosion**
  - **Present_BI= Initial_BI / (1+ Pressure_Rate)$^{Time}$**
  - **Loss BI by market erosion=100-100/(1+Pressure_Rate)$^{n}$**
- **Trade-offs between quality and time to market**

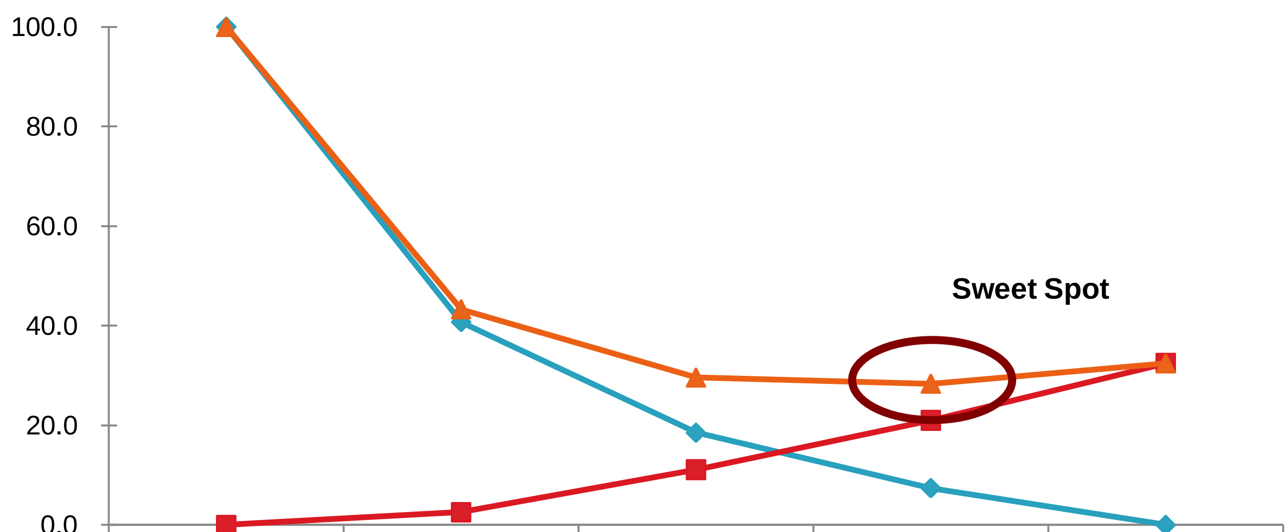



| | Start | Round 1 | Round 2 | Round 3 | Round 4 |
|---|---|---|---|---|---|
| loss by quality | 100.0 | 40.7 | 18.5 | 7.4 | 0.0 |
| loss by market | 0.0 | 9.4 | 35.9 | 59.0 | 77.3 |
| Total loss | 100.0 | 50.2 | 54.4 | 66.4 | 77.3 |

**BI Loss (High Pressure Rate=16%)**

- **In some cases, continuing to test may be worthwhile in terms of customer-perceived quality**

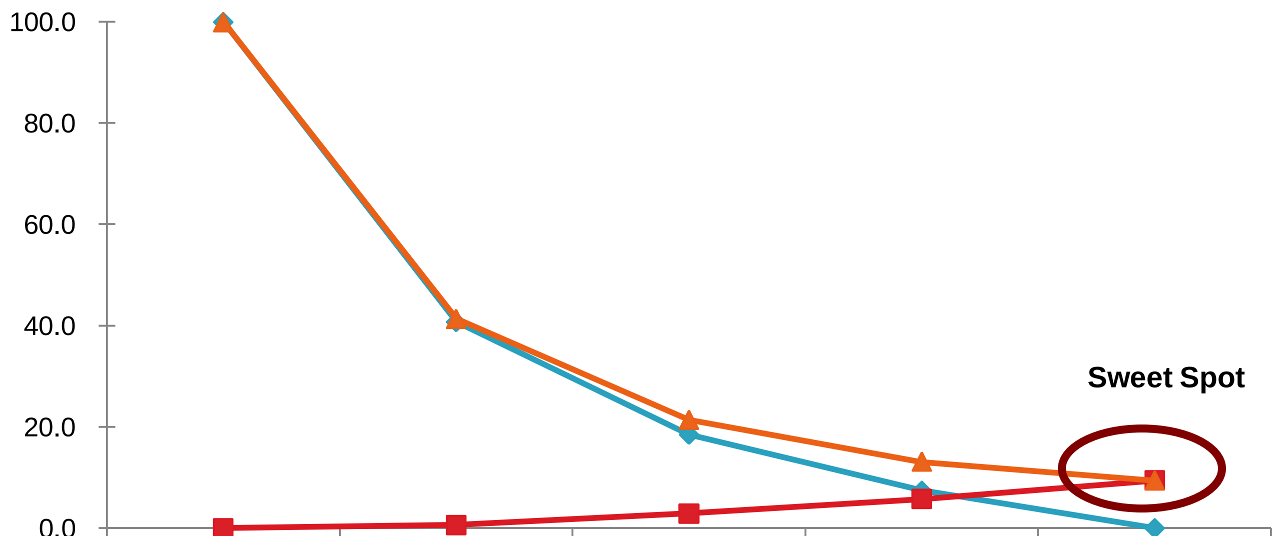
# How Much Testing is Enough (Cont.)




| | Start | Round 1 | Round 2 | Round 3 | Round 4 |
|---|---|---|---|---|---|
| loss by quality | 100.0 | 40.7 | 18.5 | 7.4 | 0.0 |
| loss by market | 0.0 | 2.6 | 11.1 | 21.0 | 32.4 |
| Total loss | 100.0 | 43.3 | 29.6 | 28.4 | 32.4 |

**BI Loss (Moderate Pressure Rate=4%)**

# How Much Testing is Enough (Cont.)




| | Start | Round 1 | Round 2 | Round 3 | Round 4 |
|---|---|---|---|---|---|
| loss by quality | 100.0 | 40.7 | 18.5 | 7.4 | 0.0 |
| loss by market | 0.0 | 0.7 | 2.9 | 5.8 | 9.5 |
| Total loss | 100.0 | 41.4 | 21.5 | 13.2 | 9.5 |

**BI Loss (Low Pressure Rate=1%)**

# Limitation

- Based on the assumption that the features for prioritization are independent
- Applicable for business critical projects, not suitable for safety critical domains
- Most of numbers are expert estimation (ratings ranged from 1 to 9), including the estimation of BI, testing cost estimation and risk quality probability, although we try to reduce the bias from experts, but it is hard to avoid it
- Method validation: a simple function to display the market pressure’s influence to BI; however, may not fit the real life situation

# Future Work

- Deal with prioritization for dependent feature interaction
- Apply this method to a set of e-services projects at USC to get more empirical validation

# Conclusion

- **A value-based approach for prioritizing features for testing which aligns the internal test process with the value objectives coming from the customers and the market**
  - Involve prioritizing features based on their business importance, quality risk, and testing cost
  - Adjust feature's value priority during the testing process
  - Provide stop-testing decision criteria based on the market pressure
- **Benefits:**
  - Decision Support for Testing Plan: helps to identify features with high business importance, high quality risk and low cost
  - Improve testing efficiency and client satisfaction
  - Easy to implement and won't add significant burden to the testing team

# Thank you

# Questions or Comments?