Gartner TeraQuest

Benchmarking the ROI for SPI:

Some New Thoughts

On an Old Problem

Joyce Statz, TeraQuest
Bob Solon, Gartner

Agenda

Business Cases and ROI

Challenges with ROI for SPI

Available Data on ROI for SPI

Benchmarking Approaches

Business Benefits from PI

PI business cases identify benefits such as

- **new revenue from new capabilities**
- **more business (or products) from faster cycle time**
- **reduced costs from reductions in rework**
- **reduced costs of operations**
- **new revenue from improvements in organizational capability**
- **more revenue or reduced cost from productivity improvement**
- **more revenue from additional business from improved customer satisfaction**

Are these measurable? In use?

Are they comparable across organizations?

Costs of PI

PI business cases identify costs such as

- **labor for defining or improving the processes**
- **travel costs for the PI work**
- **administrative costs for PI**
- **fees for training and other start-up activities**
- **tools, repositories, systems for process management**
- **cost of specialty services, such as assessments**

Which of these are measurable? In use?

Are they comparable across organizations?

What is ROI?

Traditional accounting definition

$$\text{Return on Invested Capital} = \frac{\text{Earnings Before Interest and Taxes}}{\text{Average Invested Capital}}$$

Definition often used by process improvement program

$$\text{Return on PI Program} = \frac{\text{Program Benefits} - \text{Program Cost}}{\text{Program Cost}}$$

Which of the PI benefits fit the traditional accounting definition?

Challenges with ROI on PI

Balancing long-term and short-term focus

- long-term requires much patience
- long-term view gets systemic change
- short term view leverages project-local improvement
- short-term invites cost savings at the wrong times

ROI on process improvement is challenging for

- improvements that rely on significant organizational change
- changes that affect multiple organizations

Greatest challenge: initial PI justification when you don't yet have your own data….

Using Reported Results

Without local data, organizations look to industry data

Much anecdotal data on software process improvement

- ROI values reported range from 4 to 70 x cost
- Median reported ROI is around 5:1
- Benefits vary, based on organization goals: productivity, defect levels, cost, schedule attainment, effort spent, customer satisfaction, staff attitude
- Costs per individual in the organization vary from $200 to $2500 on successful programs

Systems process improvement data is much in demand

SEI Industry Analysis – a Classic

Improvement benefit	Orgs	Annual median	Annual range
Productivity growth	4	↑ 35%	9% - 67%
Pre-test defect detection	3	↑ 22%	6% - 25%
Time to market	2	↓ 19%	15% - 23%
Field error reports	5	↓ 39%	10% - 94%
Return on investment	5	5.0:1	4:1 - 8.8:1

Source: Herbsleb et al. (1994)

A New COCOMO II Factor

Scale factor in new COCOMO II model - Process Maturity (PMAT)

- analysis of 161 data points in COCOMO II database shows statistically significant correlation between improvements in PMAT and reductions in software project effort
- PMAT values of 7.8, 6.24, 4.68, 3.12, 1.56 used in the exponent hat relates size to effort
- example impact going from Level 2 to Level 3

Project Type	Typical Size	Productivity Improvement
Small	10 KSLOC	4%
Medium	100 KSLOC	7%
Large	2000 KSLOC	11%

Source: Boehm, et.al. Software Cost Estimation with COCOMO II, 2000, p. 67

The Holy Grail of ROI for PI

The Challenge: A common ROI benchmark

Costs can be normalized by staff size – not much of an issue

What elements should be included as benefits?

- **revenue growth?**
- **market share?**
- **cycle time?**
- **quality level?**
- **productivity?**

Many look for productivity improvements in some form

Productivity Comparison

Productivity can be defined several ways, too

- classic approach – units of output/unit time
- an alternative – revenue/person/year

Software productivity analysis often uses SLOC as the basis for size, making benchmark comparisons difficult

- SLOC cannot be directly aggregated across languages
- even different versions of the same language cannot always be aggregated
- newer development environments can produce much functionality with few if any lines of code produced

Functional Measures of Size

Function Point Analysis (FPA) provides a measure of automated functionality delivered to the end user

Function point (FP) counts are independent of the technologies, processes or platforms used

FPA-based sizes are linear, scalable and comparable

- **1000 FPs is twice as large as 500 FPs**
- **3000 FPs written in COBOL is functionally equivalent to 3000 FP written in Java**

Formal definitions are controlled by International Function Point Users Group; other variants exist

Using FP Counts for ROI

Software size in FPs can be used as the normalized unit in key ROI calculations of productivity...

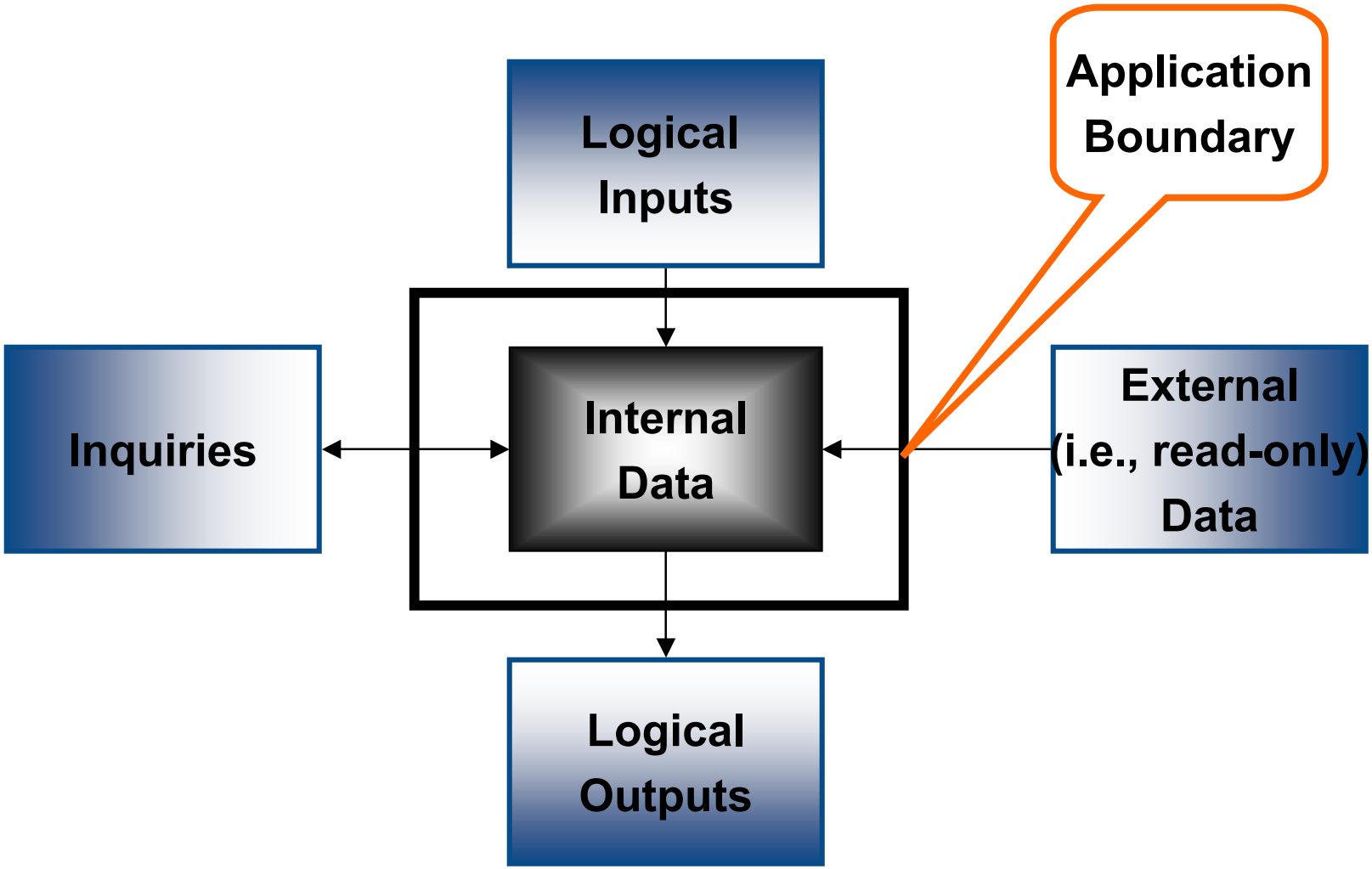
- **productivity = units of output/unit time**
- **productivity = xx function points/developer year (FTE)**

...and of efficiency

- **development efficiency = units of output/financial investment**
- **development efficiency = yy function points/$$$ spend**

Use of FPA avoids issues of normalization

FP Counting Logical Model



Some Existing Productivity Data

Gartner's Application Development (AD) performance benchmark data provides FP-based productivity information

- 43,700 development projects, 44,616,000 FPs
- 55,700 supported applications, 124,588,000 FPs
- all major technologies, languages, databases
- Includes project data from ~1991

The AD benchmark gathers size and labor data at the application/project level

- allows discrete analysis of technical and performance data at low level
- productivity can be calculated at application/project level or aggregated

Existing Correlations with Process

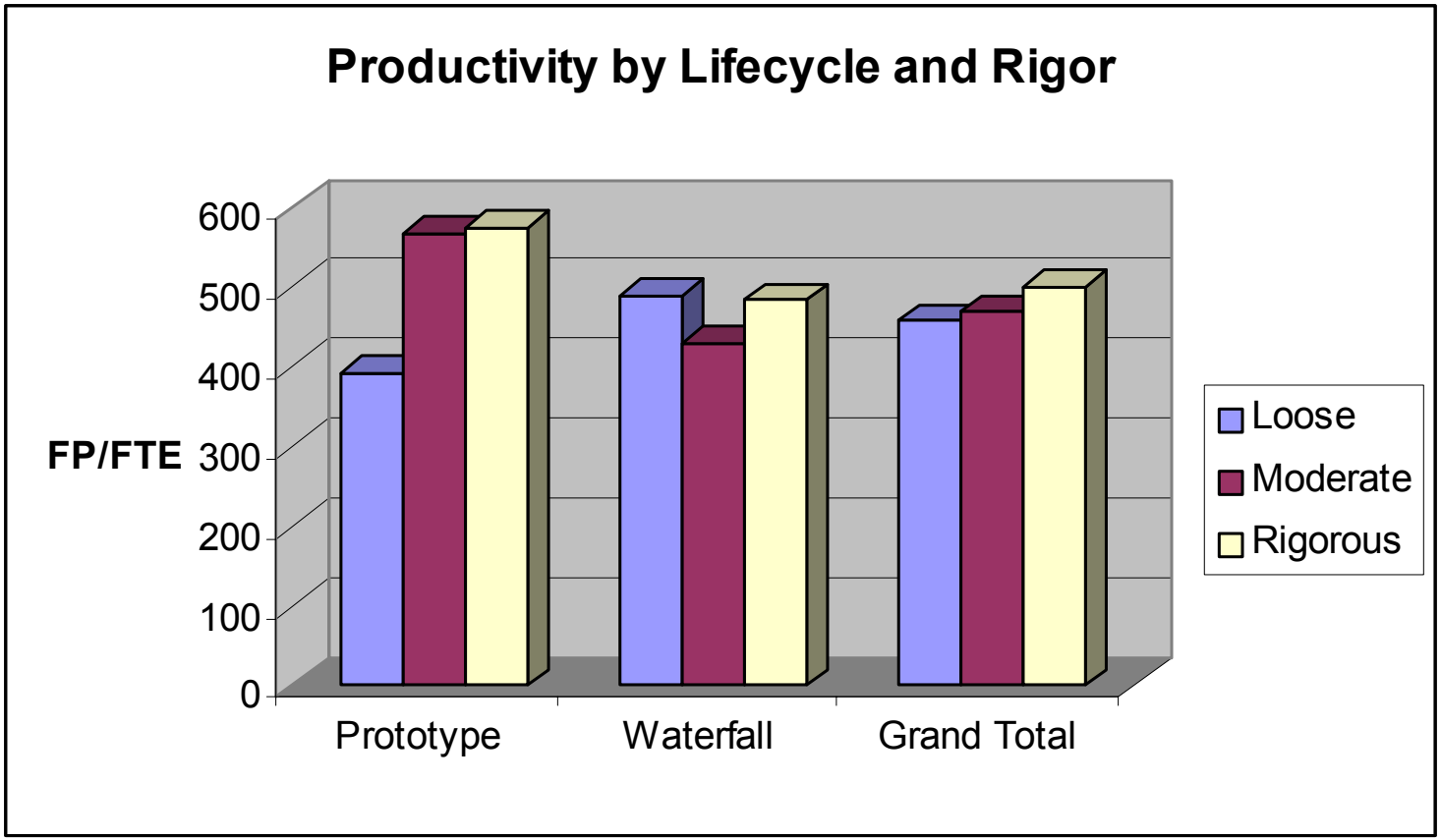
AD benchmark currently asks respondents to rate their development process by generic life cycle and level of rigor:

- **Waterfall/Prototyping**
- **Loose/Moderate/Rigorous**

Definitions of levels of rigor:

- **Loose: informally followed; little or no documentation**
- **Moderate: checkpoints at major phase boundaries; responsibility rests with project managers; little or no external oversight**
- **Rigorous: extensive documentation; independent oversight/quality assurance/process management tools often used**

Initial Analysis of Process Data



Source: Gartner Measurement, data from 2000-2001

Basis for Initial Analysis

Analysis uses project data from the last two years

Examining Grand Total shows that productivity rises as process rigor increases

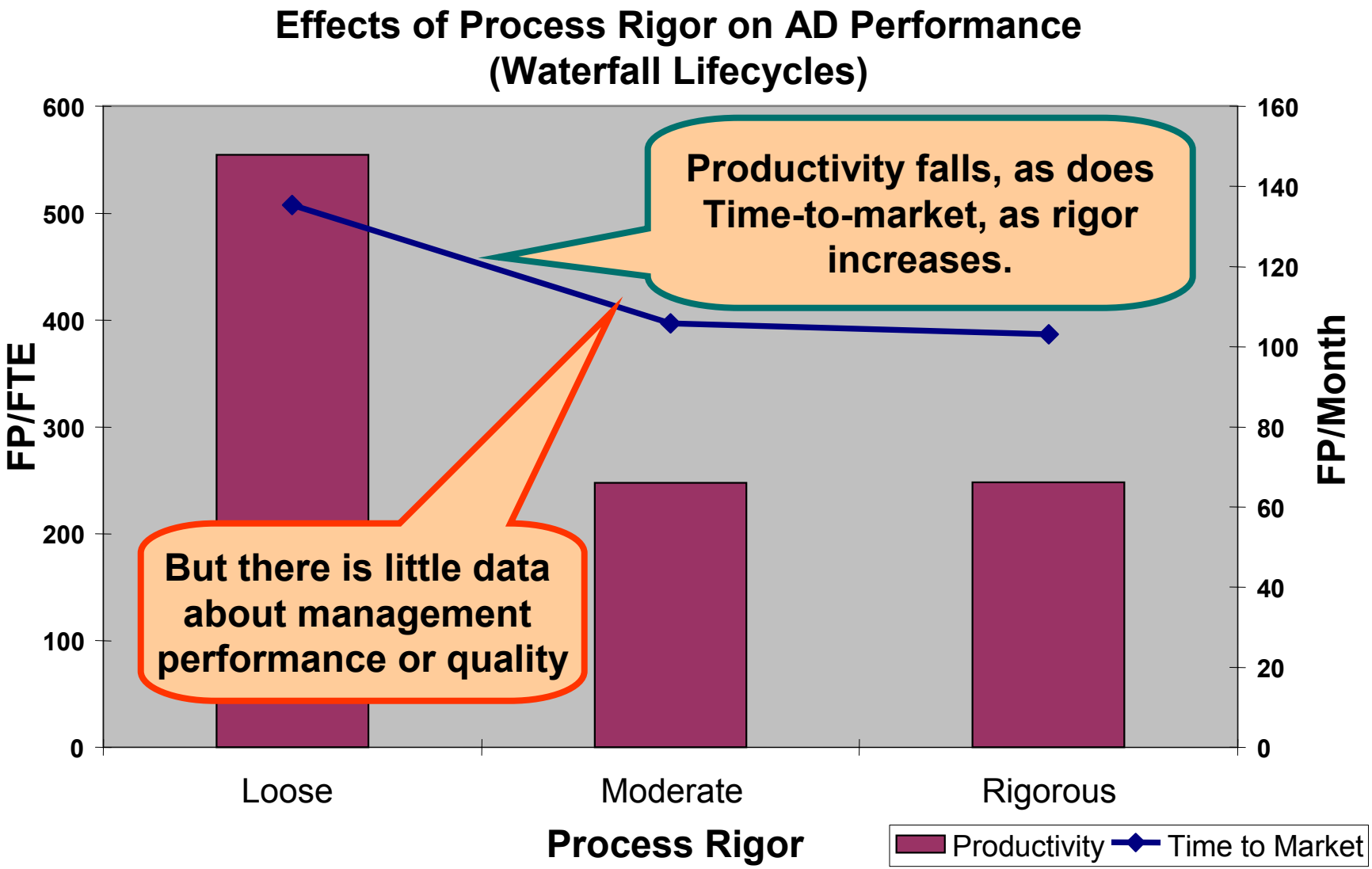
Prototyping life cycle shows similar trend, with greatest increase between Loose and Moderate rigor

However, the traditional Waterfall life cycle shows lowest productivity at the Moderate level of rigor

Note that the rigor levels do not correspond with any particular CMM Level or other external process model

After excluding outliers, and looking at Waterfall data only….

Further Analysis



Many Questions Remain

New studies will gather performance data to CMM Level as well as existing benchmark data

Our Hypotheses

CMM Level 2:

- Somewhat improved productivity and cost over Level 1
- Somewhat longer time to market than Level 1

CMM Level 3:

- Notable improvement in productivity and cost
- Improvement in time to market
- Improvement over time in defect densities

CMM Level 4 and 5:

- Further improvement in all performance criteria
- Expansion of analysis into other areas of performance

Similar Results?

Study of 30 software products of $1B/year IT firm

- **3.3 million lines of COBOL MRP system**
- **created over period of 12 years, 1984-96**
- **simultaneously looked at product quality, cycle time, and cost**

Increasing CMM process maturity associated with

- **higher product quality**
- **increases in development effort**

***<u>However</u>*, reductions in cycle time and effort because of improved quality outweigh increases in development effort, so the net effect is reduced cycle time and effort**

Source: Harter, et.al. Management Science, 2000.

You Can Help!

Once sufficient data is available for analysis, we will test the performance hypotheses against normalized performance data to investigate how performance varies by CMM Level.

We always need more data. Consider participating in the current studies by the SEI, Gartner, and TeraQuest on the relationships between CMM Levels and other performance data. Contact the speakers for information on how you can get involved.

Contact Us

Joyce Statz, Vice President

TeraQuest Metrics, Inc.
12885 Research Blvd, Suite 207
Austin, TX 78750
(512) 219-9152
joyce.statz@teraquest.com

Bob Solon, AD Measurement Practice Manager

Gartner, Inc.
3836 North Drexel Avenue
Indianapolis, IN 46226
(317) 237-4039
robert.solon@gartner.com

References - 1

Boehm, Barry W., et.al. Software Estimation with COCOMO II. Upper Saddle River, NJ: Prentice-Hall, 2000.

Clark, Bradford. "The Effects of Software Process Maturity on Software Development Effort." Ph.D. Dissertation. August 1997, University of Southern California

Harter, Donald E. , Mayuram S. Krishnan, and Sandra A. Slaugher. "Effects of Process Maturity on Quality, Cycle Time, and Effort in Software Product Development. Management Science, vol. 46, No. 4, April, 2000, p. 451-466.

References - 2

Herbsleb, J, Carleton, A., Rozum, J., Siegel, J., and Zubrow, D. (1994). Benefits of CMM-Based Software Process Improvement: Initial Results (Tech. Rep. CMU/SEI-94-TR-13). Pittsburgh: Software Engineering Institute, Carnegie Mellon University.

McGibbon, Thomas. "A Business Case for Software Process Improvement Revisited." Updated DACS State-of-the-Art Report, September 30, 1999.

van Solingen, Rini. "The Cost and Benefits of Software Process Improvement." unpublished white paper, 2001.

Excellent focus issue on benchmarking – *IEEE Software*, September-October, 2001.