estimate

estimate • analyze • plan • control

Why Can't People Estimate: Estimation Bias and Strategic Mis-Estimation

PSM 2015

Joseph P. Dean

jdean@galorath.com

SEER by GALORATH

Key Points

SEER by GALORATH

Without care estimates are usually biased (even with experts)

Tempering with an "outside view" can mitigate some bias

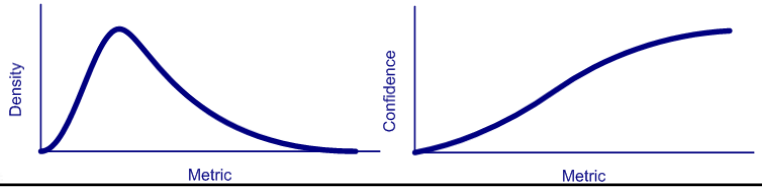
Estimates can be better, squelching bias & strategic mis-estimation… Parametrics help

ESTIMATION & PLANNING:

An Estimate Defined

- An ***estimate*** is the most knowledgeable statement you can make ***at a particular point in time*** regarding:
  - Effort / Cost
  - Schedule
  - Staffing
  - Risk
  - Reliability

- Estimates more precise with progress
- *A WELL FORMED ESTIMATE IS A DISTRIBUTION*

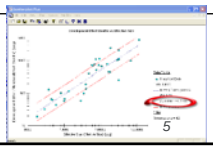


Estimation Methods - 1 of 2

Model Category	Description	Advantages	Limitations
Guessing	Off the cuff estimates	Quick Can obtain any answer desired	No Basis or substantiation No Process Almost always wrong
Analogy	Compare project with past similar projects.	Estimates are based on actual experience.	Truly similar projects must exist
Expert Judgment	Consult with one or more experts.	Little or no historical data is needed; good for new or unique projects.	Experts tend to be biased; knowledge level is sometimes questionable; usually are not consistent.
Top Down Estimation	A hierarchical decomposition of the system into progressively smaller components is used to estimate the size of a software component.	Provides an estimate linked to requirements and allows common libraries to size lower level components.	Need valid requirements. Difficult to track architecture; engineering bias may lead to underestimation.

Estimation Methods - 2 of 2

Model Category	Description	Advantages	Limitations
Bottoms Up Estimation	Divide the problem into the lowest items. Estimate each item… sum the parts and add a factor	Complete WBS can be verified.	The whole is bigger than the sum of the parts. Costs occur in items that are not considered in the WBS.
Design To Cost	Uses expert judgment to determine how much functionality can be provided for given budget.	Easy to get under stakeholder number.	Little or no engineering basis. Always over original cost
Simple CER's	Equation with one or more unknowns that provides cost / schedule estimate.	Some basis in data.	Simple relationships may not tell the whole story. Historical data may not tell the whole story.
Comprehensive Parametric Models	Perform overall estimate using design parameters and mathematical algorithms.	Models are usually fast and easy to use, and useful early in a program; they are also objective and repeatable.	Models can be inaccurate if not properly calibrated and validated; historical data may not be relevant to new programs; optimism in parameters will lead to underestimation.



Bias Types

- Optimism
- Cognitive Bias
- Confirmation Bias
- Negativity Bias
- Loss Aversion Bias
- Affect Heuristic Bias
- Thinking Fast & Thinking Slow
- Illusion of Control

Human Nature: Humans Are Optimists

Harvard Business Review explains this Phenomenon:

- Humans seem hardwired to be optimists
- Routinely exaggerate benefits and discount costs

Delusions of Success: How Optimism Undermines Executives' Decisions (Source: HBR Articles | Dan Lovallo, Daniel Kahneman | Jul 01, 2003)

Solution - Temper with "outside view": Past Measurement Results, traditional forecasting, risk analysis and statistical parametrics can help

Don't remove optimism, but balance optimism with realism

Cognitive Bias: How Fair Are We

(Source BeingHuman.org)

- Cognitive bias: Tendency to make systematic decisions based on cognitive factors rather than evidence
- Human beings exhibit inherent errors in thinking
- Researchers theorize in the past, biases helped survival
  - Our brains using shortcuts (heuristics) that sometimes provide irrational conclusions

"We usually think of ourselves as sitting the driver's seat, with ultimate control over the decisions we made and the direction our life takes; but, alas, **this perception has more to do with our desires—with how we want to view ourselves—than with reality**." Behavioral economist Dan Ariely

- Bias affects everything: from deciding how to handle our money, to relating to other people, to how we form memories

Essence of the problem: Memory is unreliable and we are hard wired to ignore risk & questioning

Confirmation Bias (Source: Beinghuman.org)

- Give more weight to information that confirms what we already believe
  - Automatic unconscious way our brains process information
  - Selectively remember information that confirms what we already think
  - When we approach new information, we interpret it in a biased way
  - Spin news story so it vindicates their own beliefs?
- We subconsciously only pay attention to the information that confirms what is already known

You would think this would help ensure viable estimates but… Its what we believe, not necessarily what is reality

9

Negativity Bias (Being Human.org)

- Unconsciously pay give more weight to negative experiences than positive ones
- Brains react powerfully to negative information than they do to positive information
- Daniel Kahneman explained:
- "The brains of humans and other animals contain a mechanism that is designed to give priority to bad news. By shaving a few hundredths of a second from the time needed to detect a predator, this circuit improves the animal's odds of living"
- More important for our ancestors to be able to avoid a threat quickly than to gain a reward

Again, this should yield viable estimates but is usually overridden

10

Loss Aversion Bias (Source BeingHuman.org)

- Tendency to strongly prefer avoiding a loss to receiving a gain
  - Explains making same irrational decisions over and over
- Kahneman: Experiment giving one third of the participants mugs, one third chocolates, and one third neither
  - Option of trading
    - 86 percent who started with mugs chose mugs
    - 10% who started with chocolate chose mugs
    - 50% who started with nothing chose mugs

- Throwing good money after bad (sunk cost fallacy)is a perfect example of loss aversion
- To avoid feeling the loss we stick with our plan, hoping for a gain, even when that just leads to a bigger loss

Explains why it is so hard to kill a failing program

Affect Heuristic Bias (Source: Beinghuman.org)

- Involuntary response to a stimulus that speeds up the time it takes to process information
  - If we have pleasant feelings, we see benefits high and risks low, and vice versa
  - affect heuristic behaves as a first and fast response mechanism in decision-making
  - Helpful in life or death situations where time was of the absolute essence.
- **System 2** The analytic, rational system of the brain is relatively slow and requires effort
- **System 1** The experiential system is different—speedy, relying on emotional images and narratives that help us to estimate risk and benefit.

Hopefully estimates elicit system 2… But often are off the cuff via system 1

Thinking Fast & Thinking Slow (Source: Kahneman)

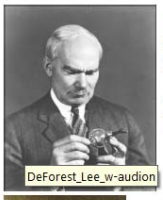
System 1: Thinking Fast	System 2: Thinking Slow
• Operates Automatically • No effort • Quick • No voluntary control	• Allocates attention to mental activities that demand it • Complex computations
• Coherent interpretation of what is going on	• Good at balancing probabilities but often indecisive
• Intuitive answers quickly	• Takes over when System 1 can't process the data • If the person is willing • Can correct or override System 1 if it determines intuition is wrong

Illusion of Control (Source: BeingHuman.org)

- Tendency to overestimate their influence over outcomes that they cannot affect
- Psychologist Ellen Langer Subjects given lottery tickets; either at random or allowed to choose their own
  - Had chance to trade tickets for others that had a higher chance of paying out.
  - Subjects who chose a ticket were less likely to part with it than those who had a random ticket
  - Subjects felt their choice of ticket had some bearing on the outcome—demonstrating the illusion of control.
- Illusion of control especially strong in stressful and competitive situations, like gambling or financial trading or ESTIMATING

Illusion of control can lead bad decisions or irrational risks

Cognitive Bias At Work (Adapted From: Northrop)



"Man will never reach the moon regardless of all future scientific advances."

- Dr. Lee DeForest, Inventor of Television

"There is no reason anyone would want a computer in their home."

- Ken Olson, president and founder of Digital, 1977

"Airplanes are interesting toys but of no military value."

- Marechal Ferdinand Foch, Professor of Strategy, Ecole Superieure de Guerre

"640K ought to be enough for anybody."

- Bill Gates, 1981

"Any general who's worth his salt knows that war is not a Nintendo game, war is not something that's fought by robots."

- Norman Schwarzkoph, 1991

"To throw bombs from an airplane will do as much damage as throwing bags of flour. It will be my pleasure to stand on the bridge of any ship while it is attacked by airplanes."

- Newton Baker, Sec. of War, 1921

Trouble Starts By Ignoring Project / Program Iron Triangle Realities

- Typical Trouble: Mandated features needed within specific time by given resources

Scope (features, functionality)

Resources **Schedule**

- At least one must vary otherwise quality suffers and system may enter impossible zone!

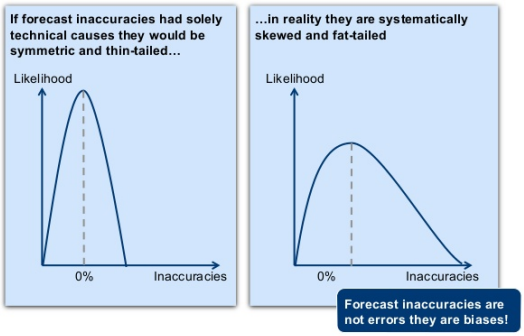
Pick Two

Explanations for Poor Estimating

(Adapted From Source Master Class on Risk, Flybjerg, 2013)

1. Technical: Inadequate data & Models (Vanston)
2. Psychological: Planning Fallacy, Optimism Bias - causes belief that they are less at risks of negative events
3. Political / Economic: Strategic misrepresentation - tendency to underestimate even when experienced with similar tasks overrunning (Flyvberg)

Technical Explanations are Not Enough...



Channel Tunnel Disaster

(Source Master Class on Risk, Flybjerg, 2013)

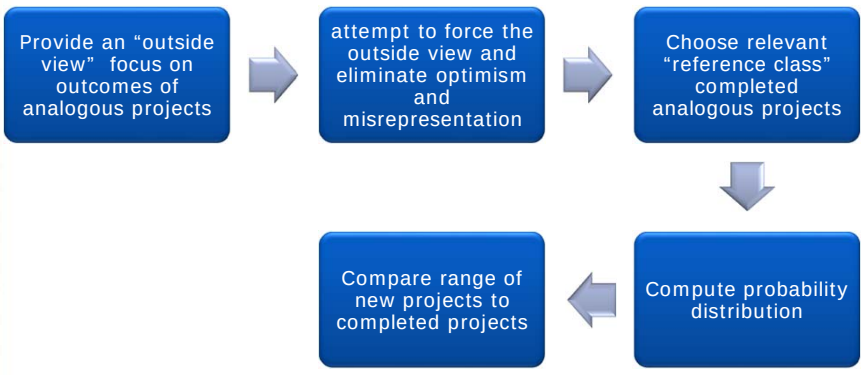
- Actual Costs 200% of Estimates
- Actual Benefits ½ times estimates
- Actual NPV 17.8 Billion Pounds

Business Case results were eliminated because of over-optimism in costs and over-optimism in benefit

Reference Class Forecasting (adapted from http://www.slideshare.net/assocpm/a-masterclass-in-risk)

- Best predictor of performance is actual performance of implemented comparable projects (Nobel Prize Economics 2002)
- Provide an "outside view" focus on outcomes of analogous projects
- "Reference Class Forecasting" attempts to force the outside view and eliminate optimism and misrepresentation
  - Choose relevant "reference class" completed analogous projects
  - Compute probability distribution
  - Compare range of new projects to completed projects

Reference Class Forecasting



Best predictor of performance is actual performance of implemented comparable projects (Nobel Prize Economics 2002

3 point estimates

- Optimistic value (s_{opt})
- Most likely value (s_m)
- Pessimistic value (s_{pess})

- Expected value (*EV*)

$$EV = \frac{(s_{opt} + 4s_m + s_{pess})}{6}$$

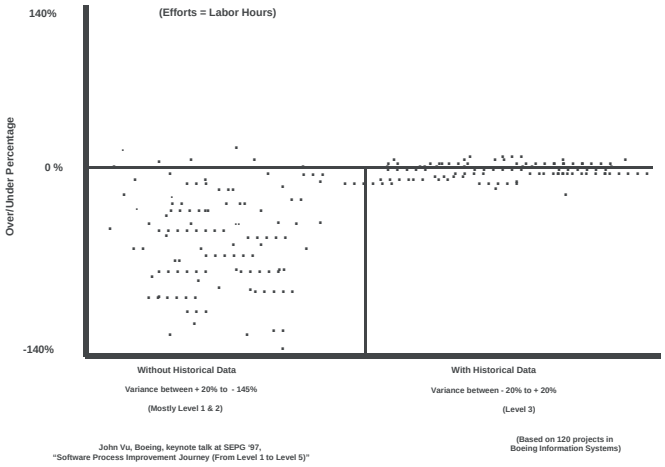
Questions After Developing Likely Estimate

- What might go wrong?
- What are the likely consequences?
- Is the staff involved experienced
- Have problems occurred with this kind of work before?
- Does this activity depend on inputs, resources, or other factors we don't control?
- Are there aspects of this work that we don't understand well
- If betting would your estimate change?
- Capture all potential difficulties as identified risks

Understanding The Data: Pepsi Challenge Example

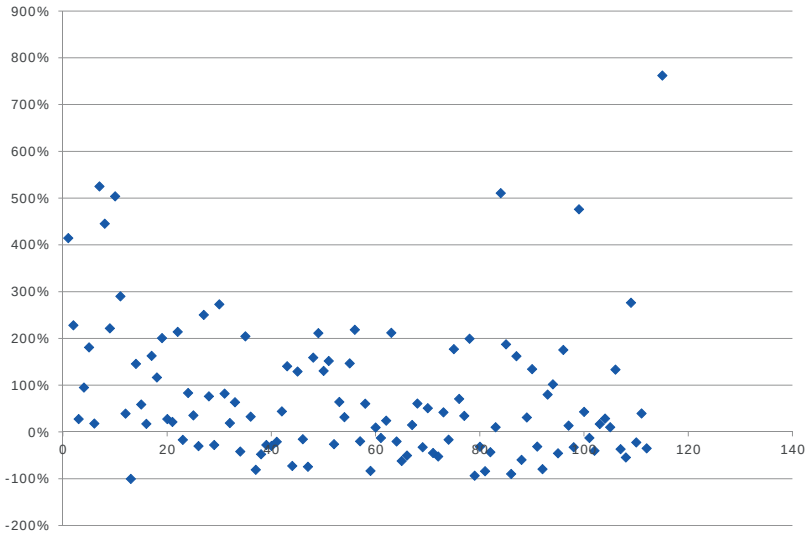
- Coke and diet Coke outsold Pepsi
- Blind sip test showed 57% preferred Pepsi vs Coke
- This was the reason for new Coke
  - "New Coke" beat Pepsi in sip test
  - People didn't like "new Coke" so Coke went back to classic
- Problem: measuring sips versus bottles
  - Consumers like the sweetness of Pepsi in a sip
  - Not so in a whole portion

Data Improves Estimates For New Programs Source: John Vu, Boeing SEPG 1997



SRDR Estimated New SLOC vs Actual

(Note: HUGE outliers removed to make the graph more readable)

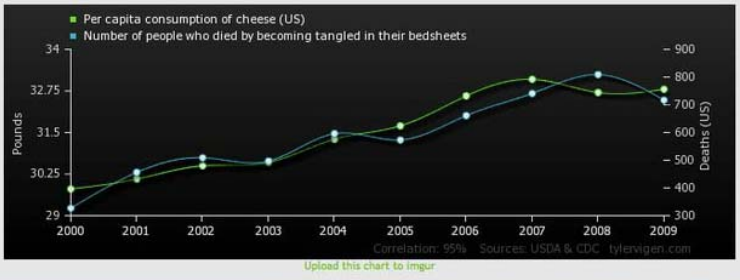


Gross underestimation of software size versus actual

25

Correlation Doesn't Always Mean Causation (Source: www.memolition.com)

Per capita consumption of cheese (US)
correlates with
Number of people who died by becoming tangled in their bedsheets



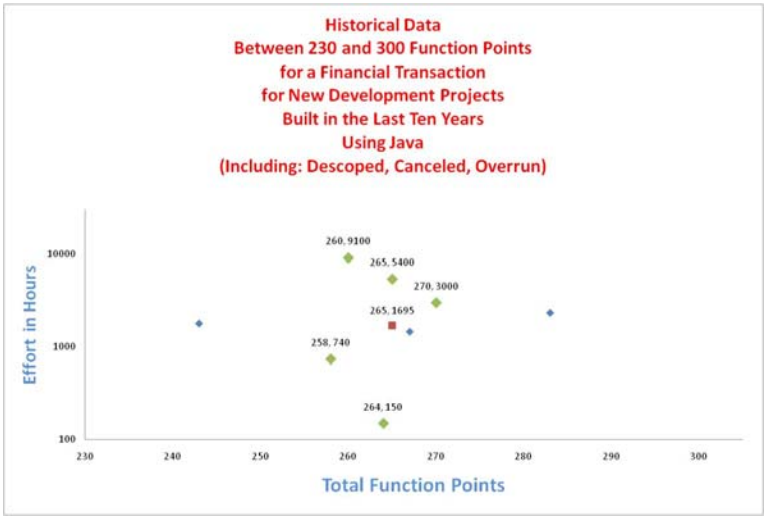
	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009
Per capita consumption of cheese (US) *Pounds (USDA)*	29.8	30.1	30.5	30.6	31.3	31.7	32.6	33.1	32.7	32.8
Number of people who died by becoming tangled in their bedsheets *Deaths (US) (CDC)*	327	456	509	497	596	573	661	741	809	717

Correlation: 0.947091

26

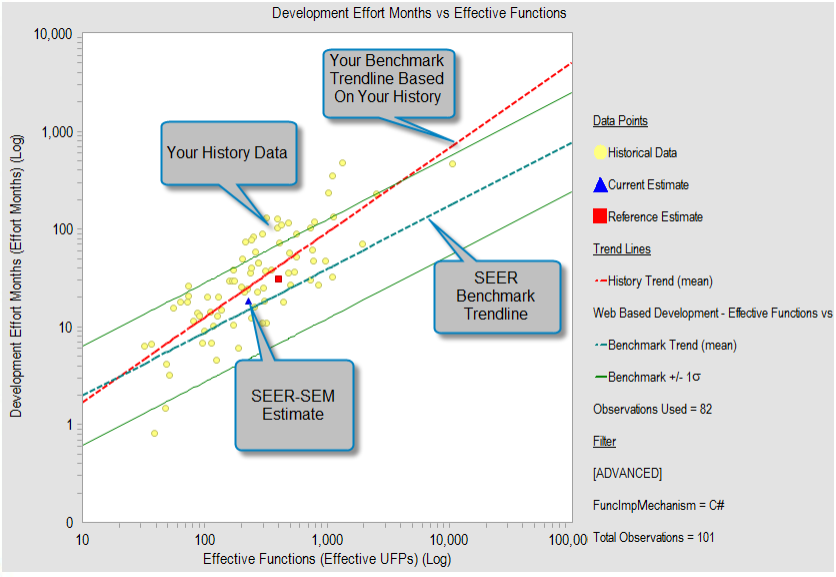
Fallacy of Silent Evidence

What about what we don't know?

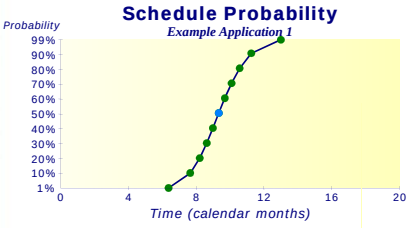


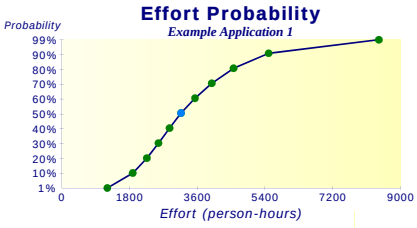
How confident would you feel if the Silent Evidence was visible?

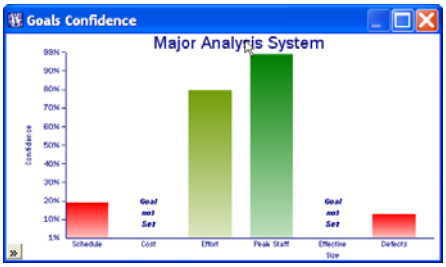
Example: Parametric Estimate Compared With History

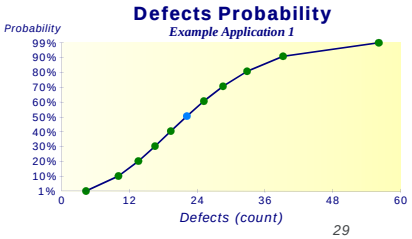


Understand Project Risks Include Them In Planning Decisions (Example SEER-SEM Outputs)









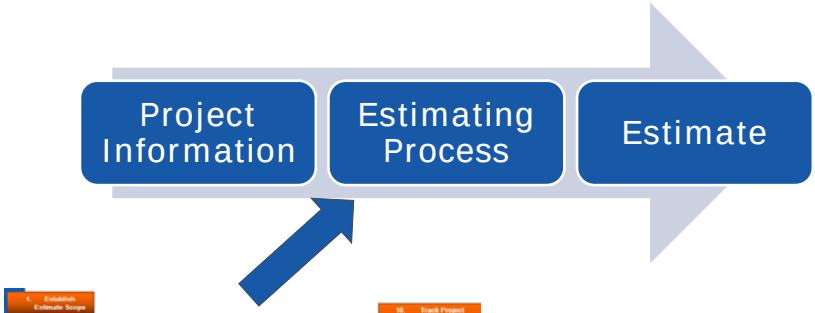
29

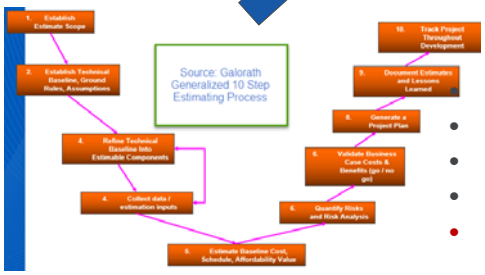
5 Levels of Risk Management (Adapted from Flyvbierg)

1 Opinions	2 Benchmarking	3 Diligence	4 Rigorous Estimating	5 Risk Analysis
As unbiased as possible	Comparing to viable database	Estimate review	Parametric Relative Reference Class Forecasting	Risk management Black Swan mitigation

30

Estimating Process Should Help Mitigate Bias (Adapted from Andy Prince)





Process Provides

- Traceability
- Repeatability
- Best Practices
- Analytical Mindset
- STEPS TO MITIGATE BIAS

Example Bias Mitigation Using Multiple Sources

Evaluate All Sources of Software Size…

New Size | Pre-existing Size (rework) | Generated Code | COTS/GOTS | Glue Code | Integrated Code

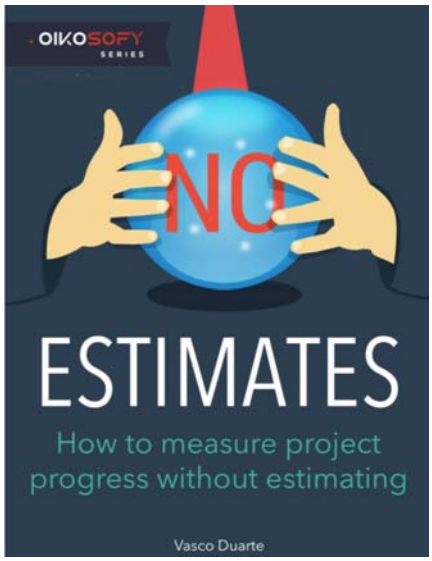
Total Size Estimates	Least	Likely	Most
Expert Judgement	12000	15500	17000
Relevant Range by Analogy	19850	24750	32540
Sizing Database	8000	32000	46000
Functional Analysis	19680	27540	35400
SEER-EstimateByCompare	15450	22650	29850
Delphi Analysis	16788	19750	22713
Estimate Range	**12000**	**22650**	**46000**

…Using Multiple Methods

Root Causes Of Bad Estimates In Agile Projects As An Example

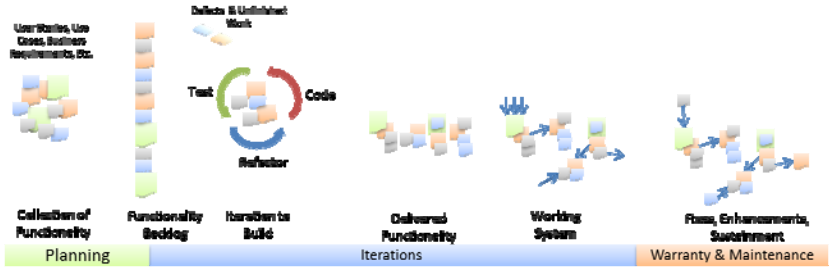
- Team not really doing Agile
  - Everyone seems to have their own “hybrid” which is code for management controls
- Immature process
  - No one with previous experience, i.e.: no Scrum Master
  - No training in the process being used
- Management gets in the way
  - Micromanage the burn down chart
  - Want to use velocity as productivity
  - Assume Ideal Days = Capacity Days
- Bad Story Counting
  - Trying to use counts across teams
  - Using historical story point counts for new work

Add In The Agile Bashing of Estimating For a Full View



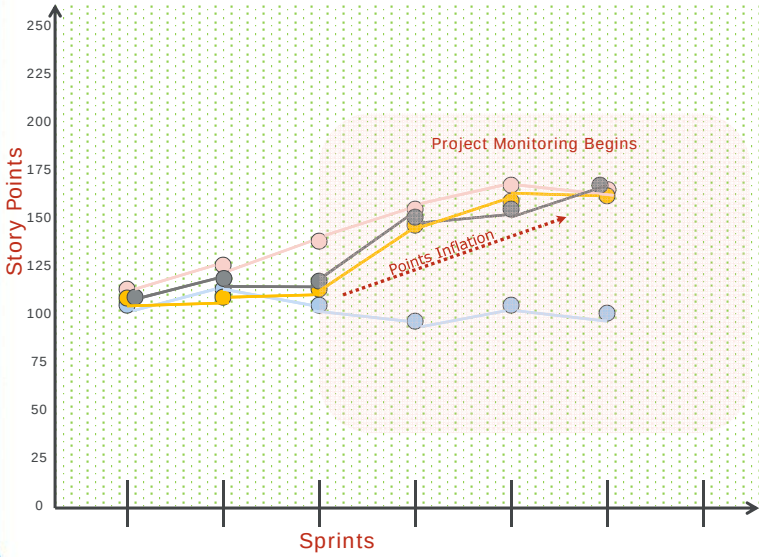
The Agile "Life Cycle"
(Scrum Example)

- Focus is on what features can be delivered per iteration
- Not fully defined what functionality will be delivered at the end?



- Iterations are often called a "Sprint"

Inflation in Story Point Productivity



Comparison of Parametric & Bottoms Up Methods (Source Hamaker)

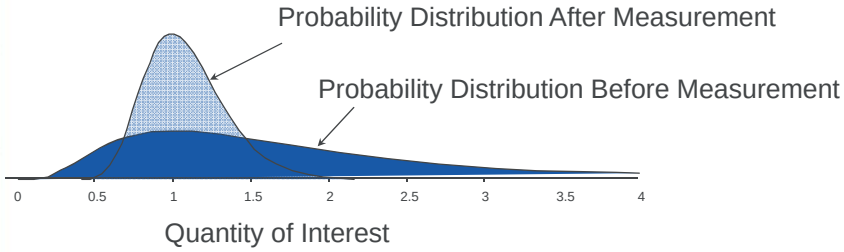
Parametric Estimates	Detailed Build-Up Estimates*
Benefits	**Benefits**
Top down	**Bottoms up**
Less detail	**More detail**
Based on performance metrics	**Accepted method**
Less labor intensive	**Generally understood**
Quicker	**Based on time and material**
Ease of trade-offs analyses	**Issues**
Generally more disciplined	**Labor intensive**
• Standard methodology	**Time consuming**
• Independent	**Trade offs need details**
• Done by trained analysts	**Performance standards**
• Captures totality of past programs	**More susceptible to distortions**
Issues	• Optimism/Pessimism
Parametric database Not always accepted	• Special interest/buy-in
"Black Magic" aura	• Done by managers/engineers
	• Missing
	- "I forgots"
	- Unknowns

*AKA "labor-material build up", "grass roots", "bottoms up" "engineering estimates"

2/24/2016

Hubbard: Measure To Reduce Uncertainty

- Perception that measurement is a point value is a key reason why many things are perceived as "immeasurable"
- Measurement: Quantitatively expressed reduction in uncertainty based on observation



38

Gunning for Models (Adapted from Hubbard)

- Be careful of red herring arguments against models
  - "We cannot model that…it is too complex."
  - "Models will have error and therefore we should not attempt it."
  - "We don't have sufficient data to use for a model."
  - "It works but we cant see all data so we should not use it"
- Build on George E. P. Box: "Essentially, all models are wrong, but some are useful."
  - Some models are more useful than others
  - Everyone uses a model – even if it is intuition or "common sense"
  - So the question is not whether a model is "right" or whether to use a model at all
  - Question is whether one model measurably outperforms another
  - A proposed model (quantitative or otherwise) should be preferred if the error reduction compared to the current model (expert judgment, perhaps) is enough to justify the cost of the new model

39

Key Points

Without care estimates are usually biased (even with experts)

Tempering with an "outside view" can mitigate some bias

Estimates can be better, squelching bias & strategic mis-estimation… Parametrics help.

40