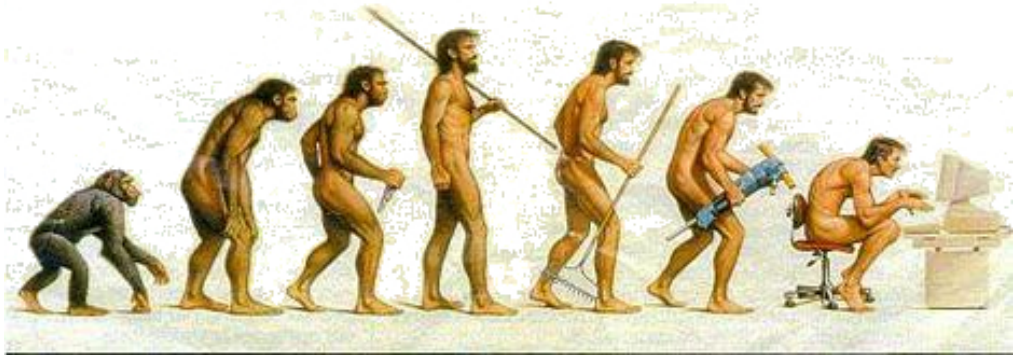
Pioneers of Parametrics



ISPA-SCEA Conference
New Orleans, LA
June 14, 2007

Ricardo Valerdi, PhD

MIT

This work has been funded by the National Science Foundation, the USC Center for Systems & Software Engineering, and the MIT Lean Aerospace Initiative

Outline

- Motivation
- Meet the Pioneers
- Major achievements in parametrics
- Industry/Government/Academia synergy
- Impact of research on practice and practice on research
- Challenges for the future

Motivation

- Much of the “how” is known; but not much of the “why”
- Little information is available about the evolution of the field
- People who pioneered the field are retired or approaching retirement

“History, by apprising [people] of the past, will enable them to judge the future.”

-Thomas Jefferson

Approach

Literature review encompassing:

- Current practice and go backwards
- Commercial best practices that surround cost estimation
- Where “research” was done in gov’t, industry, and academia
- Links to other developments in cost estimation efforts in software, services, and systems
- Evidence of impact
- Influences from outside of the US

But papers and books don’t tell the whole story…

Pioneers of Parametrics

Name (Affiliation)	User	Evaluator	Developer
Vic Basili (University of Maryland)		X	X
Barry Boehm (TRW/USC)			X
Stephen Book (MCR Federal)		X	
Dan Galorath (Galorath, Inc.)			X
Jairus Hihn (NASA JPL)	X	X	
Randy Jensen (Hughes/Jensen Consulting/US Air Force)			X
Capers Jones (Software Productivity Research)			X
Larry Putnam (Quantitative Software Measurement, Inc.)			X
Don Reifer (RCI Consulting)			X
Dieter Rombach (Fraunhofer IESE/TU Kaiserslautern)		X	
Walker Royce (TRW/Rational/IBM)	X		X
Marilee Wheaton (The Aerospace Corporation/TRW)	X		

Sample Interview Questions

- What got you interested in software resource estimation?
- How had you previously been doing SRE?
- What problems with you current approaches were you trying to address?
- How did you determine the functional forms for your resource estimating relationships (RERs)?
- To what extent did you draw on previous SRE models or studies of software cost and schedule drivers?
- Where did you get data to calibrate/validate the RERs?
- Did the data and calibration cause any changes in your RERs?
- Were there any environments that were difficult to model with your RERs?
- What kinds of organizations initially and eventually were major users of your RERs?
- What kinds of impacts did your RERs have on their practice?

Major Achievements in Parametrics

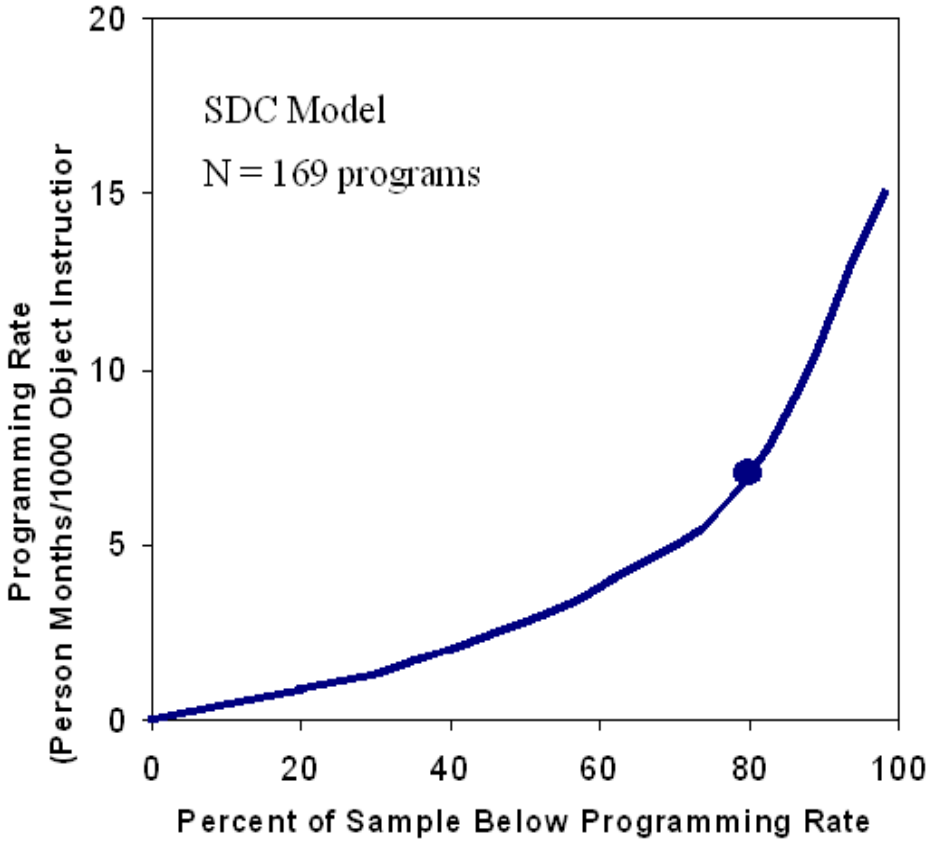
Early achievements: 1965-1985

Model	Year
System Development Corporation	1965
Martin Marietta PRICE H	1973
TRW Wolverton	1974
Putnam	1976
Doty	1977
RCA PRICE S	1977
IBM-FSD	1977
Boeing-Black	1977

Model	Year
IBM Function points	1979
SLIM	1980
Bailey-Basili Meta-Model	1981
SoftCost, -R	1981
COCOMO	1981
Jensen/SEER-SEM	1983
ESTIMACS	1983
SPQR/Checkpoint	1985

Major Achievements in Parametrics

Search for good model forms



$$PM = A * (Size)^{B} * (EM)$$

ADDITIVE

EXPONENTIAL

MULTIPLICATIVE

Major Achievements in Parametrics

Development of Model Evaluation Criteria

- Scope: Covers desired range of situations?
- Granularity: Level of detail sufficient for needs?
- Accuracy: Estimates close to actuals?
- Objectivity: Inputs repeatable across estimators?
- Calibratability: Sufficient calibration data available?
- Contructiveness: Helps to understand job to be done?
- Ease of use: Parameters easy to understand, specify?
- Prospectiveness: Parameters values knowable early?
- Parsimony: Avoids unnecessary parameters, features?
- Stability: Small input changes mean small output changes?
- Interoperability: Easy to compare with related models?

Major Achievements in Parametrics

Emergence of a Model Marketplace and Community of Interest

SEER®

PRICE

QSM The Intelligence behind Successful Software Projects

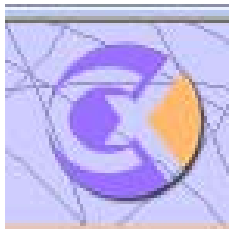
SOFTSTAR

ESTIMACS

r2ESTIMATING

COCOMO II

Softcost-R



SPR KnowledgePLAN®

Major Achievements in Parametrics

Intermediate achievements: 1985-2005

- *Mainstream Refinements: 1985-1995*
  - *Function points*
  - *Risk analyzers*
  - *Monte Carlo approaches*
  - *Breakdown of estimates by phase, activity, or increment*
- *Proliferation of Software Development Styles: 1995-2005*
  - *Need for exponential scale factor*
  - *Reuse*
  - *Early sizing methods in SW and HW*
- *Alternative Model Forms*
  - *Analogy-based estimation*
  - *Neural networks*
  - *Systems dynamics*

Systematic review of SW Development Cost Estimation Studies

(Jorgensen & Shepperd, TSE, Jan 2007)

- Classification of 304 papers in 76 journals
- Found TSE to be the most highly cited journal for cost estimation
- Found that only 13 researchers with more than five journal papers
- Regression-based estimation approaches dominate
- More estimation by analogy and expert judgment studies are taking place
- First wave of theories is behind us

TABLE 5
Estimation Approaches

Estimation approach	-1989	1990-1999	2000-2004	Total
Rg	21 (51%)	76 (47%)	51 (51%)	148 (49%)
An	1 (2%)	15 (9%)	15 (15%)	31 (10%)
Ej	3 (7%)	22 (13%)	21 (21%)	46 (15%)
Wb	3 (7%)	5 (3%)	4 (4%)	12 (4%)
Fp	7 (17%)	47 (29%)	14 (14%)	68 (22%)
Ct	0 (0%)	5 (3%)	9 (9%)	14 (5%)
Si	2 (5%)	4 (2%)	4 (4%)	10 (3%)
Nn	0 (0%)	11 (7%)	11 (11%)	22 (7%)
Th	20 (49%)	14 (9%)	5 (5%)	39 (13%)
By	0 (0%)	1 (1%)	6 (6%)	7 (2%)
Cb	0 (0%)	3 (2%)	2 (2%)	5 (2%)
Ot	2 (5%)	7 (4%)	16 (16%)	25 (8%)

The abbreviations used are: Rg = Regression, An = Analogy, Ej = Expert judgment, Wb = Work break-down, Fp = Function Point, Ct = Classification and regression trees, Si = Simulation, Nn = Neural network, Th = Theory, By = Bayesian, Cb = Combination of estimates, Ot = Other

Amendments to Jorgensen & Shepperd 2007

- Books considered to be milestones
  - Boehm, B., Software Engineering Economics, Prentice Hall, 1981.
  - Boehm, B. W., Abts, C., Brown, A. W., Chulani, S., Clark, B., Horowitz, E., Madachy, R., Reifer, D., & Steece, B., *Software Cost Estimation with COCOMO II*, Prentice Hall, 2000.
  - Conte, S. D., Dunsmore, H. E., & Shen, V.Y., *Software Engineering Metrics and Models*, Benjamin/Cummings Publishing Company, Menlo Park, 1986.
  - Galorath, D., and Evans, M., *Software Sizing, Estimation, and Risk Management*, Auerbach Publications, Boston, MA, 2006.
  - Gilb, T., *Software Metrics*, Winthrop Publishers, Cambridge, MA, 1977.
  - Jones, C., *Applied Software Measurement: Assuring Productivity and Quality*, McGraw Hill, New York, NY, 1991.
  - Jones, C., *Estimating Software Costs*, 2nd Ed., McGraw-Hill Osborne Media, 2007.
  - Kitchenham, B., & Littlewood, B., *Measurement for Software Control and Assurance*, Elsevier, 1989.
  - Kruchten, P., *The Rational Unified Process*, 2nd ed: Addison Wesley, 2001.
  - Nelson, E. A., *Management Handbook for the Estimation of Computer Programming Costs*, AD-A648750, Systems Development Corp., 1966.
  - Putnam, L. H., Myers, W., *Five Core Metrics: The Intelligence Behind Successful Software Management*, Dorset House Publishing Company, 2003.
  - Reifer, D., *Practical Software Reuse,* Wiley, 1997.
  - Royce, W., *Software Project Management,* Addison Wesley, 1998.
  - Stutzke, R., *Estimating Software-Intensive Systems*, Addison Wesley, 2005.

Industry/Government/Academia synergy

- Industry: best data, but limited scope, proprietary
  - Motivations: planning, proposals, overrun avoidance
  - Seedbed for open models
    - IBM Function points, TRW COCOMO
  - Special role of commercial cost model companies
- Government: strong funding, usage, standards leverage
  - Motivations: planning, source selection, overrun avoidance
  - Challenges: cross-organization coordination, sustainment
- Academia: model exploration, extension, evaluation
  - Motivation: scientific understanding of SW phenomena
  - Challenges: data access, context understanding, sustainment
- Key roles of professional societies, databases

Evolution of Industry Roles

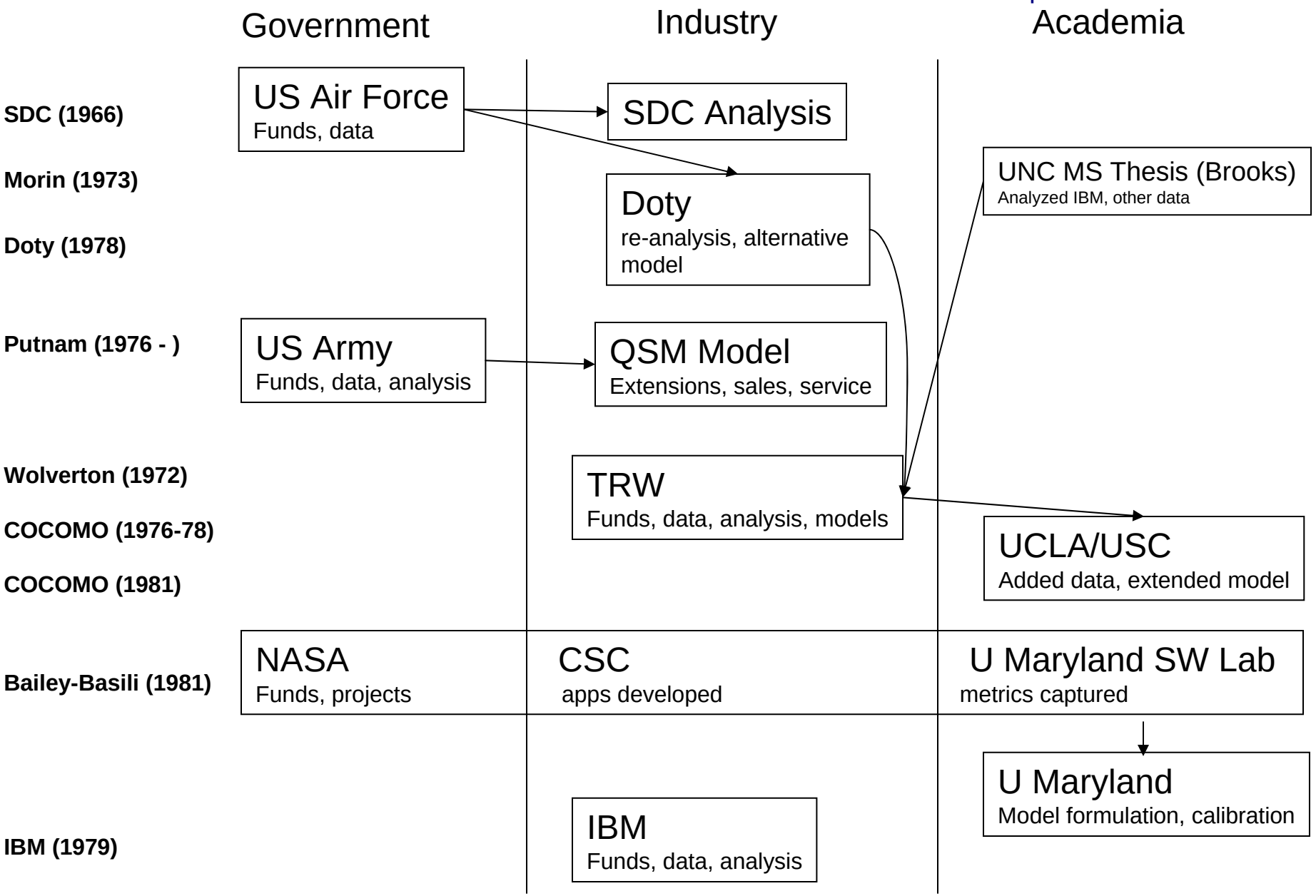
- Motivations: planning, proposals, overrun avoidance
  - 1970's: large projects (IBM, AT&T, aerospace companies)
  - 1980's: business software (Function points, analogy)
  - 1990's: product time to market (HP, telecoms, services)
  - 2000's: rapid change (agile methods, dynamic estimation)
- Increasing concern with total cost/value of ownership
  - Maintenance costs, effects on business cost and value
  - Concerns on what/how to outsource
- Special role of commercial cost model companies
  - Cross-company/area data opportunities, challenges
  - PRICE '76, QSM '78, Jensen/SEER '81, SPR '84
  - Over a dozen COCOMO-family vendors

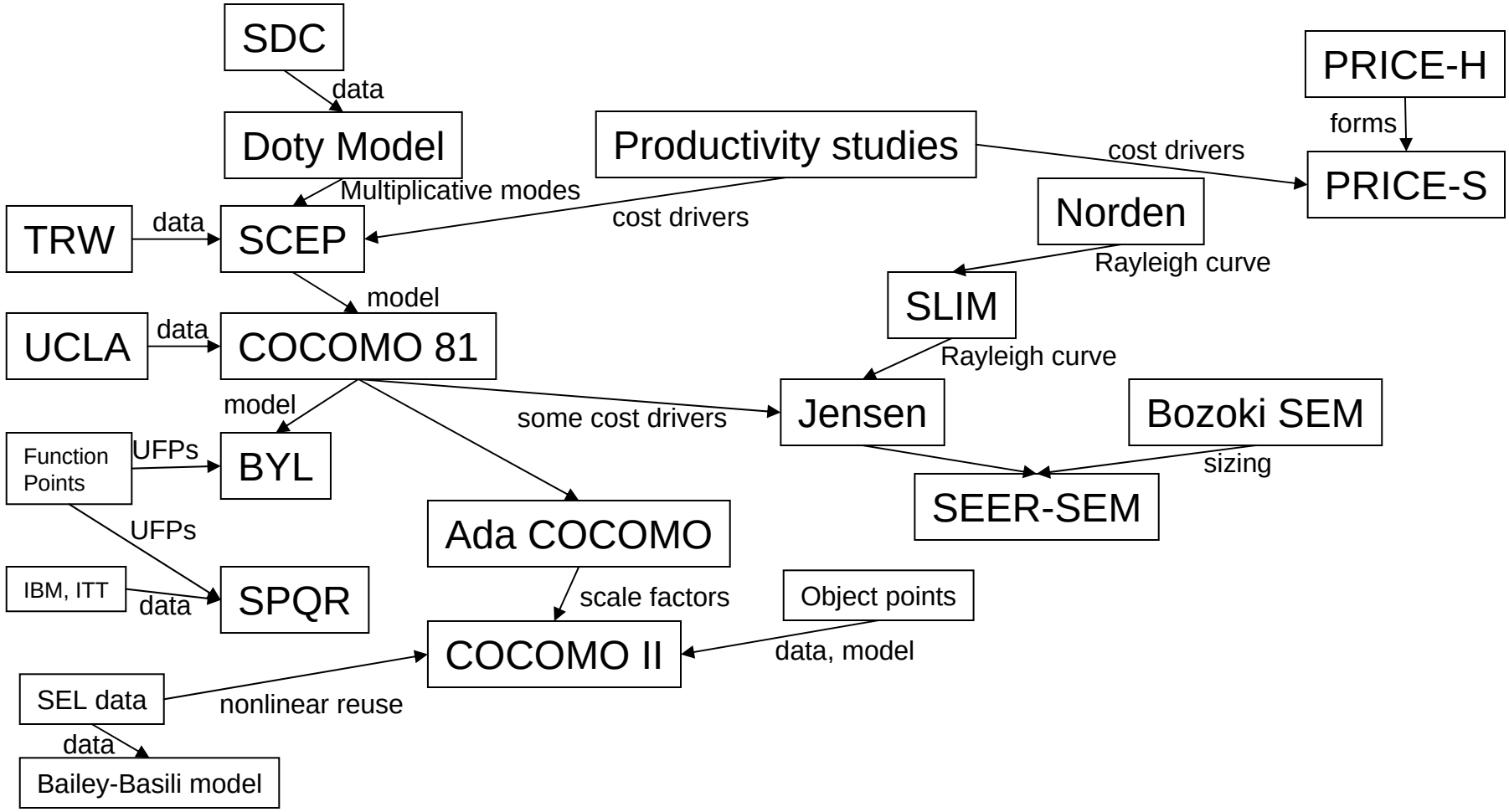
Evolution of Government Roles

- Motivations: planning, source selection, overrun avoidance
  - 1965: Air Force/SDC study: 169 data points
  - 1970's: Space software initiatives: NASA SEL, SSCAG
  - 1980's: Software cost analysis offices: DoD, NASA, FAA, UK MoD
  - 1990's: Usage/improvement stimulation via maturity models
- Government SW cost/productivity initiatives
  - European Community, UK, Norway, Japan, Australia, China
- Sponsorship of new model research, evaluations
  - Academia, small business grants, nonprofits
- Challenges: cross-organization coordination, sustainment

Evolution of Academia Roles

- Motivation: scientific understanding of SW phenomena
  - 1970's: simple models (Purdue, Maryland)
  - 1980's: industry model extensions (COCOMO, Function Points)
  - 1990's: new models: COTS, RAD, SysEngr, analogy, neural nets
- Model coordination
  - COSMIC function points, COCOMO/PRICE/SEER Rosetta Stones
- Model evaluation
  - 1980's: Kitchenham, Jeffery, Kemerer, Rubin
  - 1990's-2000's: Ferens, Jorgensen-Shepperd
- Challenges: data access, context understanding, sustainment





Professional Society Roles

- Society of Cost Estimation & Analysis (formed in 1987 from NES & ICA)
  - National Estimating Society (since 1968)
  - Institute of Cost Analysis (since 1981)
- ACM SIGMETRICS, IEEE Trans. SW Engineering (1970s)
- International Society of Parametric Analysts
  - Formed in 1980 from PRICE users' group
- International Software Engineering Research Group (1993)
- Users' groups: Function Points, COCOMO, SEER
- Conferences: ISPA/SCEA, IFPUG, EuroSCE, COCOMO/SSCM, ACM/IEEE Metrics+ISESE -> ESEM (2007)
- Journals: ISPA, Empirical Software Engineering

Key Role of Databases

- Hard to do good research without sound, comparable data
- Databases mostly proprietary
  - Companies: best data; limited scope
  - Cost model vendors: model-centered, comparability challenges
  - Academic use requires data protection, nondisclosure agreements
- Challenges with open databases
  - Hard to extend: COCOMO 81
  - Limited detail: SSCAG, ISBSG
  - Limited quality control: SSCAG, DACS
  - Limited scope: NASA/Maryland SEL
- General challenge: evolution
  - New processes, methods, artifacts
  - Schema migration, standards coordination

Enhancing Impact: Key Role of Databases

- Through data definitions
  - SEI, IFPUG counting rules
- Sustained Gov't leadership
  - Industry improvement incentives
  - Data collection: contextualized, evolving
  - Research sponsorship
- Sustained partnerships
  - User groups: IFPUG, ISBSG, cost models
  - Government/industry/academic partnerships
    - SEL, NICTA (Australia), EASE (Japan)

Pioneers' Perceived Impacts

- Data & validation had an impact on research
  - Multiple modes of estimation
  - Bayesian approximation
  - Reuse models
  - Flexibility across > 650 programming languages
  - Need for handling new methods (i.e., RAD)
  - Compatibility with standards
  - Confirmed that model was off by a factor of 2.5
  - Introduced activity allocations (i.e., MBASE/RUP)
  - Calibration factors
  - Effective technology constant
  - Addition of management factors
- Research had an impact on practice
  - More objective scoping
  - Closed-loop planning & control
  - Reduced resistance to measurement
  - Improvement culture needed to coexist with models
  - CMMI Level 5 organizational attributes
  - Orgs began to plan for schedule slips
  - Review boards at NASA demanded more detail, sophistication
  - Made users more realistic
  - Provided new ways to do tradeoffs

Challenges for the Future

- Evolution
  - Integration of software and systems engineering cost estimation
  - Sizing for new product forms
  - Accommodating future trends
- Sustainment
  - Sustained partnerships
  - Sustained professional society membership & involvement

Questions?