Simulation modeling for cost estimation

BY RICHARD WATERMAN

Department of Statistics, University of Pennsylvania,

and

DONALD RUBIN

Department of Statistics, Harvard University,

and

NEAL THOMAS

Datametrics Research Inc.

and

ANDREW GELMAN

Department of Statistics, Columbia University.

June 23th 1999

Outline

- Context for the simulation model.

- Objectives.

- The types questions it answers.

- How it works, it's scope and requirements.

- Insights and benefits.

Context I, ABC.

- Ch.5, Cooper & Kaplan. The Design of Cost Management Systems.

- Introduce the idea of an *optimal cost system*.

- One that trades off between

(A) The cost of errors.	(B) The cost of measurement.

- Error consequences include:

```
* Poor product related decisions.
* Poor product design decisions.
* Poor capital investment decisions.
* Inaccurate budgeting decisions.
```

Context II

- US Postal Service Data Quality Study (June 1997 – April 1999).

- Part of the study mandate:

   ... provided sufficiently (1) complete and (2) accurate data for rate making, ...

- First level question: *what are costs?*

- Second level question: *how well are costs estimated?*

- Simulation model addresses the second question.

Simulation model objectives

- Examine how well costs are estimated.

- Establish the impact of measurement errors.

- Explore consequences of competing viewpoints.

- Prioritize information sources.

- Take a holistic view of the cost estimation process.

A flavor of the questions addressed

- What is the uncertainty in overall Marginal Cost estimates?

- How do cost elasticity assumptions impact MC estimates of products?

- What impact would halving data collection resources have on cost estimates?

- Which costpool contributes more **uncertainty** to overall MC estimates? Delivery or Transportation?

- Where are the strongest/weakest links in the chain of cost estimate components?

Why it's a hard problem

- **Magnitude**: multiple products (i), multiple cost drivers (j).

- **Complexity** of estimation equation.

- **Multiple** information sources.

- **Inter-related** data inputs.

- **Spheres of influence**. Accountants, Economists and Statisticians.

Schematic of the Marginal Cost estimation equation.

$$UVVC(i)\ \$/Piece = \sum_{costpools(j)} \frac{Cost(j)\ \$ \cdot Elasticity(j) \cdot DKS(i,j)}{Volume(i)\ Pieces}$$
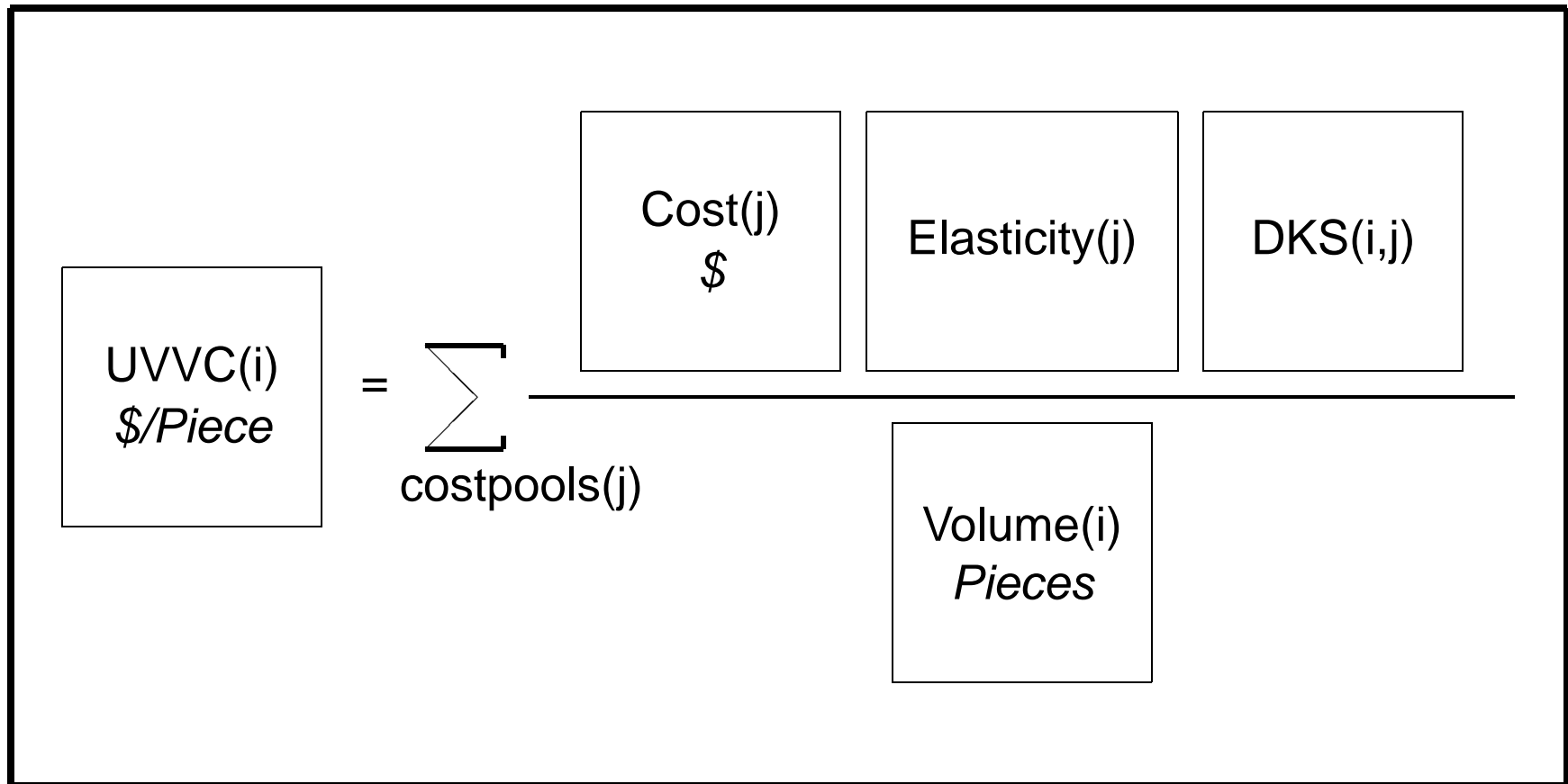
Simulation model description

- Take an established theoretical basis and combine this with best available data.

- Trade *analytical complexity* for *computational intensity*.

- Incorporate **uncertainty**, potential bias and dependencies in all input elements.

- Measure **uncertainty** in outputs (Marginal Cost estimates).

- Reflect system wide uncertainty, not just component level uncertainty.

Model Scope

- 8 mail subclasses.

- 29 Cost pools.

- Estimates Unit Volume Variable Cost/ Marginal Cost.

- Incorporates 4000 separate data inputs.

- Combines inputs from econometric studies, sample surveys and expert judgment – **multiple data sources**.

- Requires up to date data inputs to remain useful.

Cost estimation system evaluation

- How good a measuring stick do we have?

- Need to know "truth", to judge how far we are from truth.

- Generate *Hypothetical Worlds*, use these as testbeds.

- Benchmark *Hypothetical Worlds* against FY96 data to enhance credibility.

- Test measuring procedures (SEP's) against these *Hypothetical Worlds*.

How it works: an individual volume element
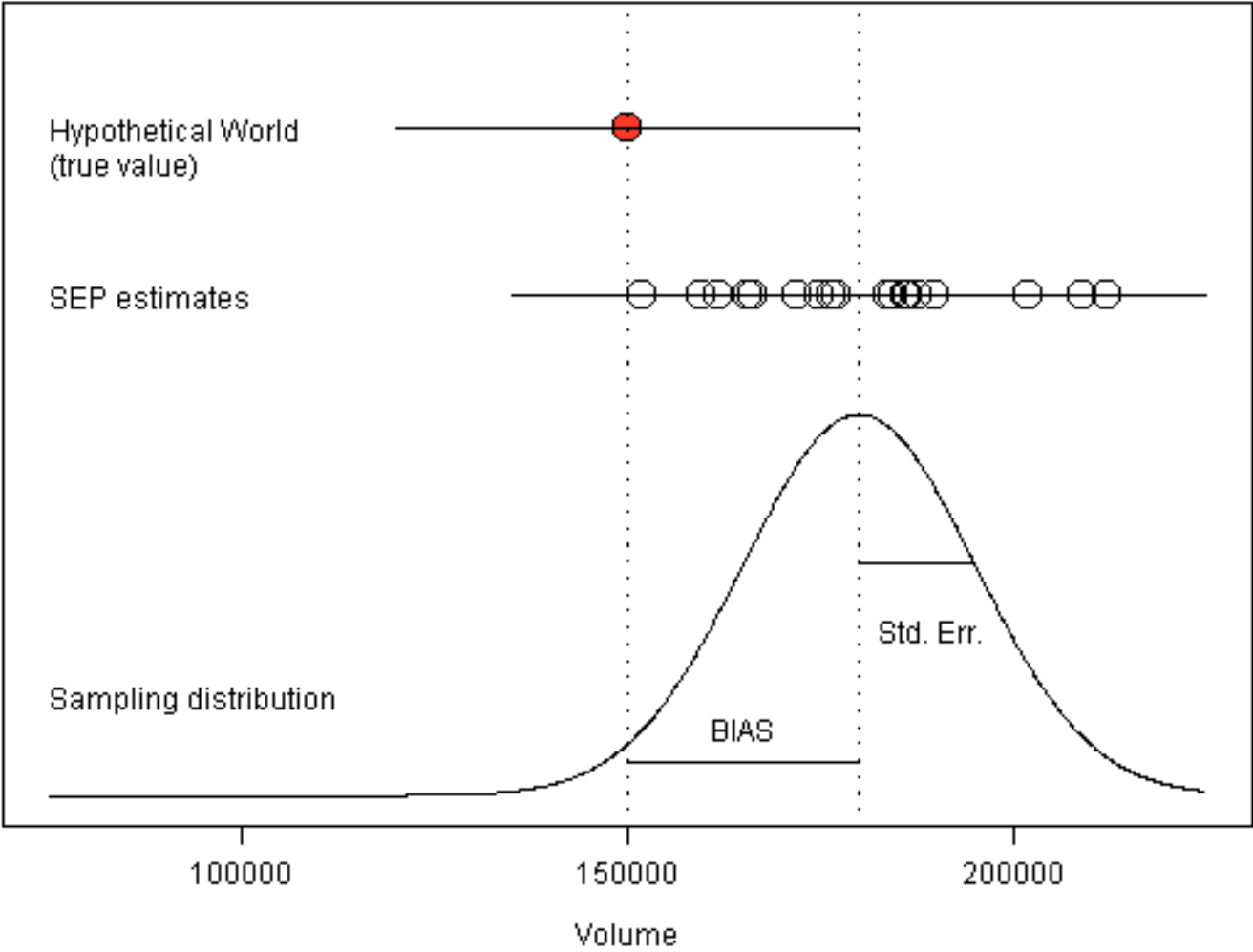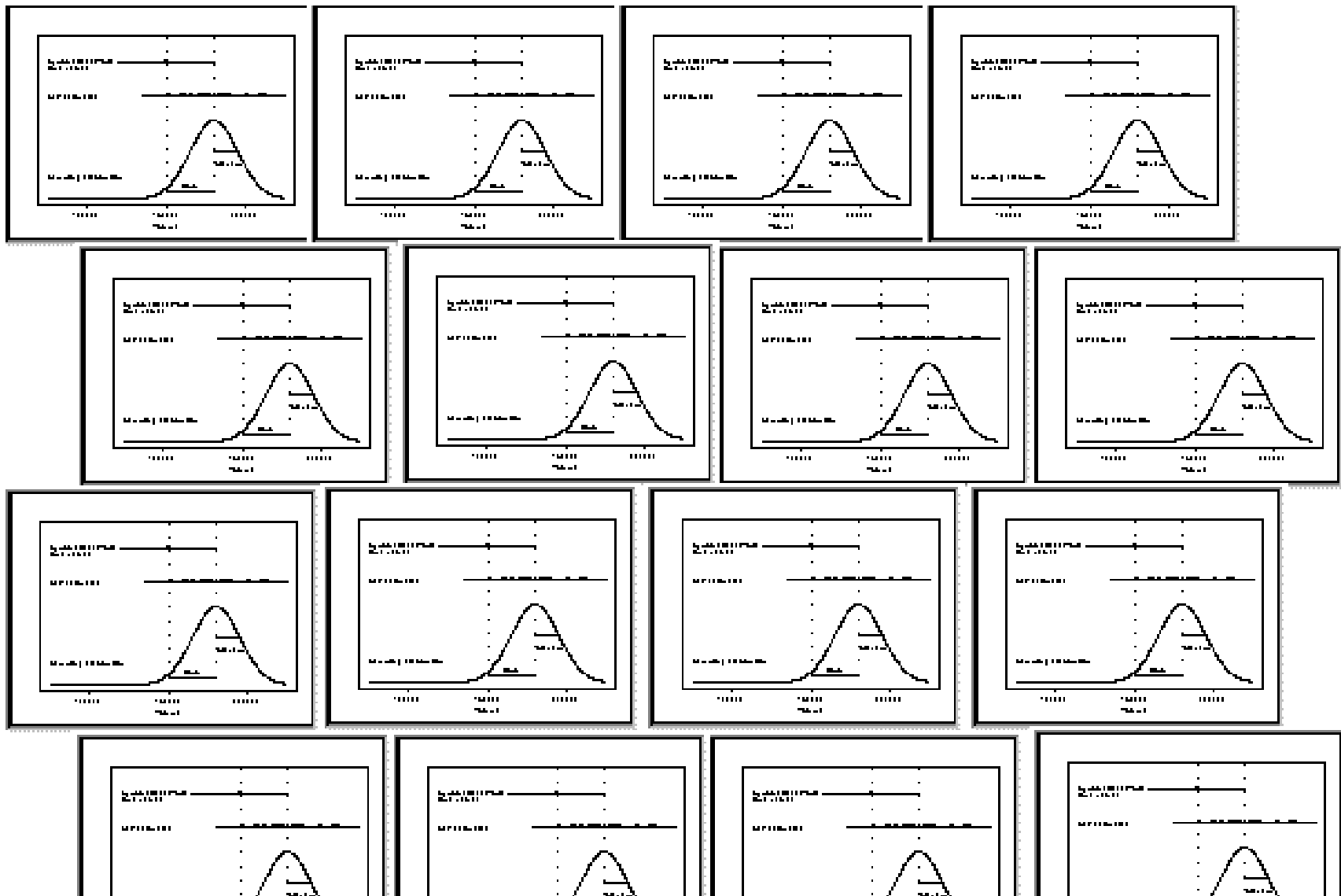
Uncertainty components

Summary of the sampling variability in the UVVC estimates for the 8 products used in the simulation model.

Subclass	UVVC mean	SEP cv	Subclass cv when a specific subsystem is "turned off"					
			RPW	IOCS	TRACS	CCS	ELAS	SS
First Class Letters Flats & IPPs	11.60	1.26	1.24	1.19	1.14	*1.30	1.24	0.81
Periodicals within County	3.38	10.09	*10.23	1.30	*10.14	10.06	10.07	10.02
Periodicals Regular Rate	10.04	2.51	2.51	2.05	1.53	2.51	2.47	*2.52
Standard A Enhanced Carrier Route	3.07	8.00	7.90	*8.14	*8.04	8.00	8.06	1.32
Standard A Regular Rate	8.52	2.11	2.00	1.88	2.03	2.06	2.05	1.04
Standard A Non-profit	6.65	2.70	2.64	2.22	2.57	*2.73	2.67	1.76
Standard B Parcel Post	146.17	4.59	4.25	4.14	3.75	*4.71	*4.64	3.85
Standard B Library	96.33	13.48	8.79	11.46	12.49	13.26	*13.49	13.24

Summary and benefits

- It is feasible to assess cost uncertainty.

- A first step toward understanding the *cost of errors*.

- Model forces assumptions to be explicit.

- Provides a framework/language for discussion.

- Platform for investigating impacts of diverse viewpoints.

- By looking at the whole picture, it can pinpoint which errors have the most serious impact.