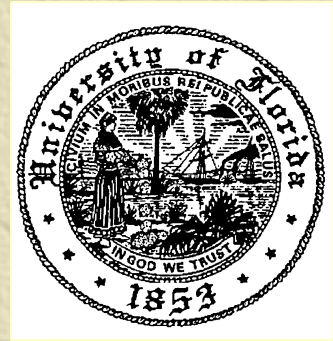
In Search of Data: An Editorial

Steven M. Shugan

bear.cba.ufl.edu/centers/MKS/general/**search**.pdf

Warrington College of Business Administration, University of Florida,
201 Bryan Hall, Campus Box 117155, Gainesville, Florida 32611-7155



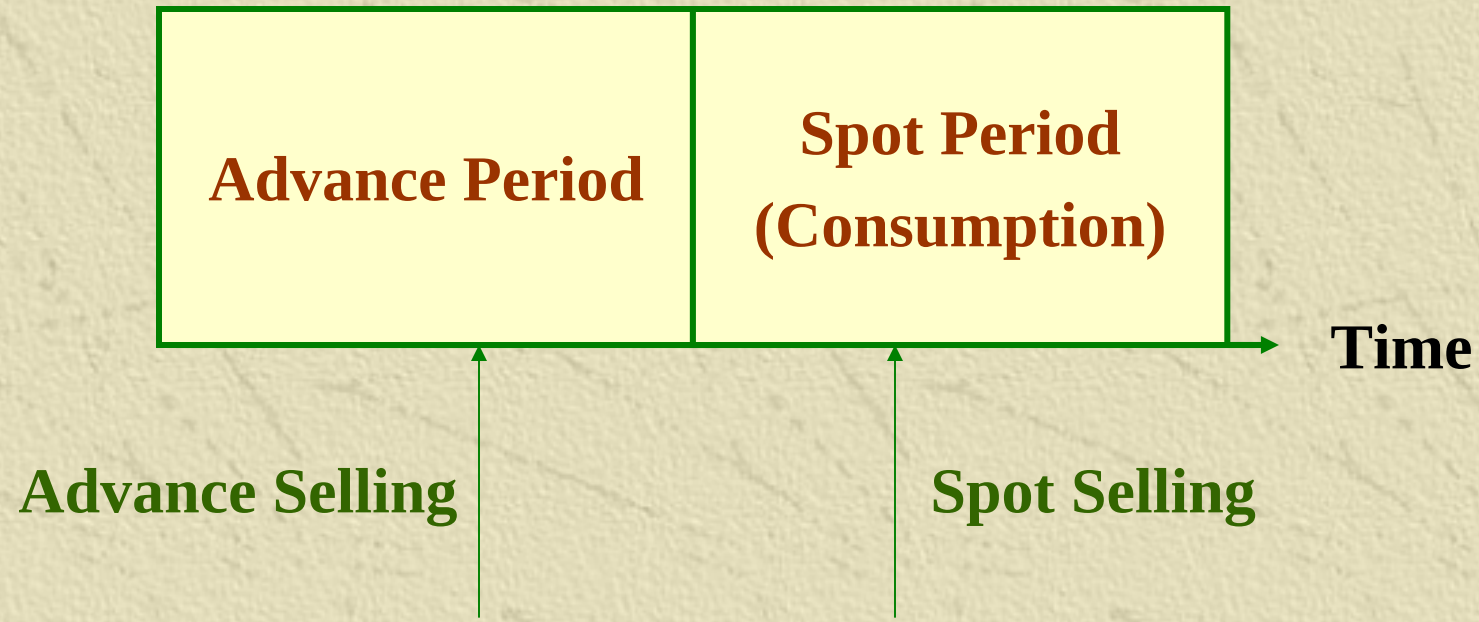

THE MARKETING JOURNAL OF THE INSTITUTE FOR OPERATIONS RESEARCH AND THE MANAGEMENT SCIENCES
informs
MARKETING SCIENCE
Editorial: Endogeneity in Marketing Decision Models
Steven M. Shugan
1
Long-Run Effects of Promotion Depth on New Versus Established Customers: Three Field Studies
Eric T. Anderson and Duncan I. Simester
4
Centralized Pricing Versus Delegating Pricing to the Salesforce Under Information Asymmetry
Birendra K. Mishra and Ashutosh Prasad
21
An Empirical Analysis of Determinants of Retailer Pricing Strategy
Venkatesh Shankar and Ruth N. Bolton
28
A Dynamic Changepoint Model for New Product Sales Forecasting
Peter S. Fader, Bruce G. S. Hardie, and Chun-Yao Huang
50
Multiple Discreteness and Product Differentiation: Demand for Carbonated Soft Drinks
Jean-Pierre Dubé
www.marketingscience.org
Dimitrios Vakratsas, Fred M. Feinberg, Frank M. Bass, and Gurumurthy Kalyanaram
109
Is Having More Channels Really Better? A Model of Competition Among Commercial Television Broadcasters
Yong Liu, Daniel S. Putler, and Charles B. Weinberg
120
Consumer Learning, Brand Loyalty, and Competition
J. Miguel Villas-Boas
134
Contingent Pricing to Reduce Price Risks
Eyal Biyalogorsky and Eitan Gerstner
146
Consumer Learning and Brand Valuation: An Application on Over-the-Counter Drugs
M. Tolga Akçura, Füsun F. Gönül, and Elina Petrova
156
Focus on Authors
170
current editor-in-chief: Steve Shugan

Why Publish in Marketing Science

- INFORMS Sponsor
- Highly Cited (SSCI #3 of 94 business journals)
- On Financial Times List MBA-rankings
- Accessible- libraries throughout world
- Accessible- EBSCO, ProQuest ABI/INFORM, JSTOR, INFORMS (on-line), Web-of-Science (abstracts)…
- Transparent
- FAST! Minimum Revisions
- Directed Revisions
- Link for Own Posting
- Specialized AE System (2 chances)
- Completely Double Blind
- Paperless

Sequence of Events

Advance selling occurs when sellers allow buyers to purchase at a time preceding consumption.



Academic Research (Shugan-Xie)

- Shugan-Xie (Journal of Service Rearch 2000)
  - Basic Idea and demonstration that uncertain valuations can increase profits without price discrimination.
- Xie-Shugan (Marketing Science 2001)
  - Full Model Monopoly with Capacity Constraints
- Shugan-Xie (California Management Review 2004)
  - Implementation and Strategic Issues
- Shugan-Xie (International Research in Marketing 2005)
  - Advance Selling with Competition
- Google SHUGAN for articles

New Technology: The Smart Card

GEMPLUS

State Dependency

Not Price Discrimination

- Complex issue … but general foundation (buyer uncertainty)
- EXAMPLE: 100 potential ticket buyers
- Future valuations unknown
- Suppose $5 or $2 with equal likelihood
- Profits
  - Spot sell at $2; profit = $2 × 100 = $200
  - Spot sell at $5; profit = $5 × 50 = $250

Not Price Discrimination

- Complex issue … but general foundation (buyer uncertainty)
- EXAMPLE: 100 potential ticket buyers
- Future valuations unknown
- Suppose $5 or $2 with equal likelihood
- Profits
  - Spot sell at $2; profit = $2 × 100 = $200
  - Spot sell at $5; profit = $5 × 50 = $250
  - Maximum advance price = $2 × ½ + $5 × ½ = $3.50

Not Price Discrimination

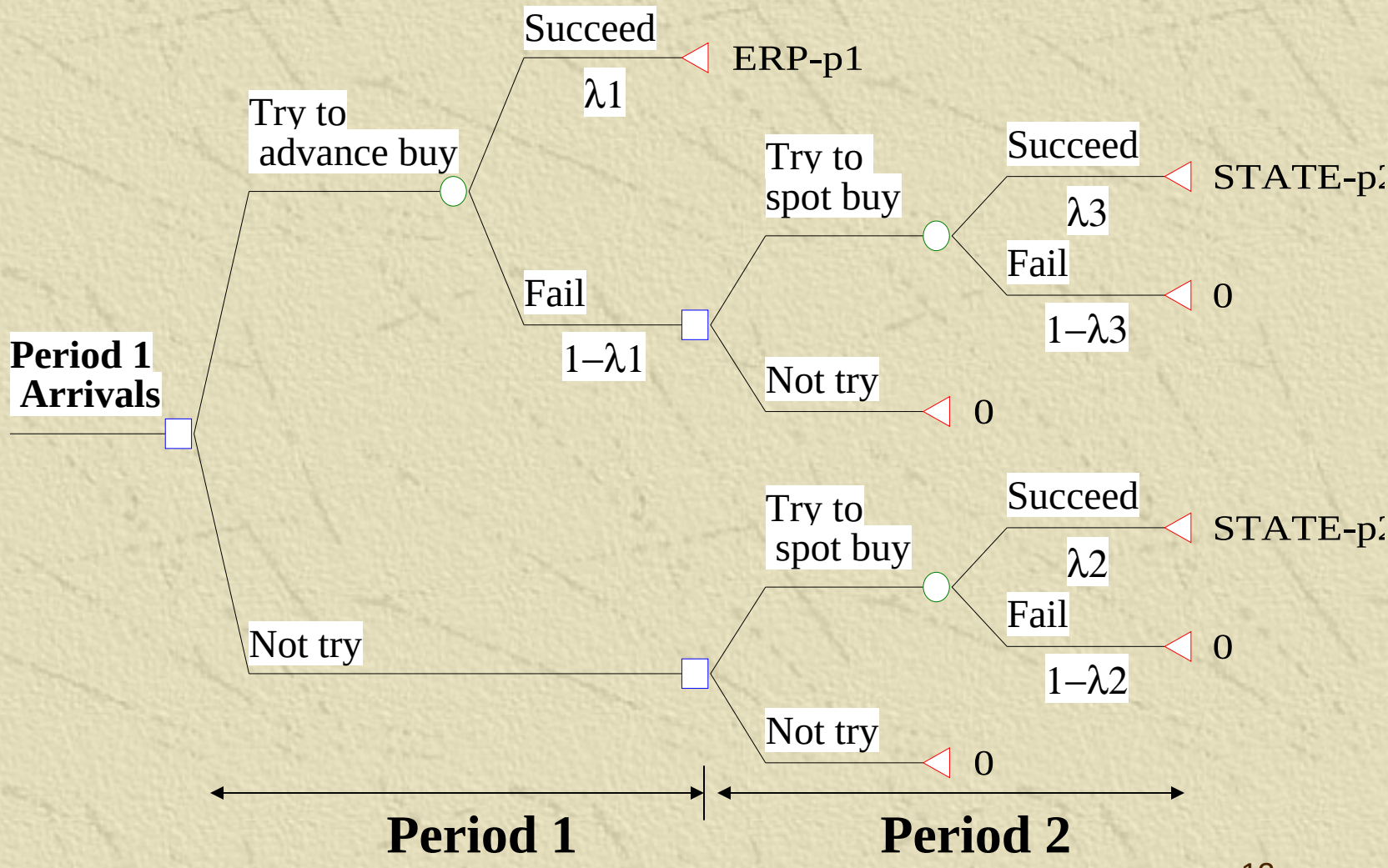
- Complex issue … but general foundation (buyer uncertainty)
- EXAMPLE: 100 potential ticket buyers
- Future valuations unknown
- Suppose $5 or $2 with equal likelihood
- Profits
  - Spot sell at $2; profit = $2 × 100 = $200
  - Spot sell at $5; profit = $5 × 50 = $250
  - Maximum advance price = $2 × ½ + $5 × ½ = $3.50
  - Advance sell at $3.5; profit = $3.50 × 100 = $350
  - Increases profits ($350-$250)/$250 = 40%

Not Price Discrimination

- Complex issue … but general foundation (buyer uncertainty)
- EXAMPLE: 100 potential ticket buyers
- Future valuations unknown
- Suppose $5 or $2 with equal likelihood
- Profits
  - Spot sell at $2; profit = $2 × 100 = $200
  - Spot sell at $5; profit = $5 × 50 = $250
  - Maximum advance price = $2 × ½ + $5 × ½ = $3.50
  - Advance sell at $3.5; profit = $3.50 × 100 = **$350**
  - Increases profits ($350-$250)/$250 = 40%
  - **1st degree price discrimination:**

    Profit = $2 × 50 +$5 × 50 = $350

Impact of Capacity Constraint on Buyers



Empirical Work

- As Illustration
- Test Assumptions (Conditions?)
  - Is there state dependency??
  - Does technology exist to advance sell??
  - Will some consumers pay more at spot??
  - Are consumers heterogeneous in period 1??
- Calibration (to make decisions)
  - Find specific advance price for a situation
  - Find specific spot price for a situation
  - Find quantity limits for situation

Conclusions

- Whether articles contain numeric data IS IRRELEVANT
- Research objective dictates desirability of data (numeric real, numeric synthetic or non-numeric)
- Assumptions SHOULD often substitute for data
- give equal scrutiny to data collection - regardless of whether researcher influences or not.
- rather than data, focus on whether research compelling evidence for conclusions

Why use data?

- OPINION vs. FACT
- Observations (or data) pivotal role for publication
- Fuel Popper “falsification” engine of science
- Data foundation decision-support systems
- Sophisticated analysis more efficiently extracts information from data
- Approaches to data sometimes have conflicting objectives

Troubling developments

- leading journal refused to accept just theory
- conflicts over sources of numeric data
- data determines classification
- mining large databases (exploratory?)
- research agendas around databases

Why Numeric Data Matters?

- Signal by author?
- Signal -- author aspires to applicability
- Signal – author concerned with realism
- Signal – author has velleity for practicality
- Signal – author respects empiricism
- MORE IMPORTANT: appropriate or inappropriate?
- MORE IMPORTANT: how data are used
- data sometimes only illustrative
- data sometimes reveal discovery
- focus on answering important research questions

Appropriate Data and Synthetic Data

- Different objectives
  - tools for decision-makers, qualitative strategies, identify or eliminate courses of action, extract information, produce hypotheses, improve theories, test theories
- different objectives require different criteria
- Example – Synthetic Data Appropriate
- Example – Synthetic Data Inappropriate
- Proofs avoid tacit & hidden conditions

The Absence of Data

- Numeric Data Sometimes Irrelevant
  - Compare 2 statistical estimators
  - Determine optimal response given conditions
  - Consistency between stylized facts
  - Deduce the testable implications from assumptions
  - produce testable hypothesis
  - decisions of what to measure & relationships
  - observations should be means to an end
- Focus SHOULD BE ON research question
- Research demarcation-- theoretical & empirical non-sensible

Substituting Assumptions for Data

- All research starts with assumptions (conditions!)
  - Conditions when research valid
  - Assumptions forfeit claims of generality
  - Explicit vs. Shared (Implicit) vs. hidden in procedures
  - Data and assumptions are substitutes
  - more assumptions - lessens need for data
  - fewer assumptions - need more data or weaker conclusions
  - limitations: assumption generality, data representativeness, future reality
- meticulous observations important but …
  - Seek information NOT observations per se.
  - Hope to reveal previously unknown information
  - previously overlooked actions and alternatives
  - value of observations = value of extracted information NOT quantity of observations
  - collecting a unique dataset not inevitably a contribution
- Avoid marginalizing the role of non-numeric data
  - non-numeric observations can reveal more than numeric
  - BUT traditional statistical tools focus on numeric data
  - nothing sacrosanct about numeric data
  - we are unable to extract information that is not there
  - sometimes, assumptions will drive the conclusions - data is a distraction

Data as Deception

- Truly empirical – findings from data
  - Sometimes assumptions within analysis dictate findings
  - finding originated from postulated model not data
  - assumptions often embedded within a complex estimation
  - assumptions: variables excluded, a concave response function, a specific prior distributions
  - implicit assumption of independence -> independence, optimal behavior, non-optimal prices
  - Excluding & including goodwill -> very different results
  - Implicit assumption firms profitable -> diverse marketing strategies can be successful
  - In litigation, found Logit models favor plaintiffs, regression favors defendants
  - falsification difficult if not impossible
  - Not argument for complexity – need parsimony
- instead, explicit assumptions are a virtue of research
  - sometimes punish being transparent
  - necessary balance given past research --- data/assumptions substitutes
  - assumptions impact all research results
  - clever ways extract information from data/ no assumptions unable to do so
  - unable to extract absent information
  - Data source of information NOT interesting sideshow (or deceptive distraction).

The Need for Data

- legitimate reasons for skepticism of theory
  - unrelated to presence of numeric data
  - ultimate contribution unclear
  - Unable to improve prediction -- Unable to improve decisions
  - Faulhaber and Baumol (1988) the record of theoretic adoption spotty
  - Ehrenberg (1995) -- consider empirical generalizations separate from theory (to discover)
  - Cobb (1928) argues that advances from accretion of data NOT excogitation of theory.
- Theory can hinder predictions.
  - Theory can restrict set of admissible models (i.e., consistent with the theory)
  - Does theory beat trial-and-error model (from unrestricted class)
  - Atheoretical Models often outperform (e.g., time series, neural network models, models of random walks) often out-predict theoretical models.
  - Theory can hinder invention – discourages pursuit of clever atheoretical ideas
  - Do serendipity, creativity and necessity play greater roles?
  - Existing theories may might discourage trial of atheoretical alternatives
- Theories should be ultimately testable and be ultimately rigorously verified
  - avoid casual excuses for not collecting data
  - Olaus Romer crudely measured speed of light in 1676, Albert Michelson precisely w/19th century technology
  - we should be humble about our difficulties
  - theory, application and technology are so intertwined -- impossible to tell which came first
  - we must hope theory, empirical research and application are synergistic
  - This argument independent of numeric data

Primary, Secondary and Good Data

- Reviews fault sample, response rate, weak manipulations, confounds
- procedures (e.g., survey, experimental, quasi-experimental)
- Sometimes ignore procedures for secondary data collection
- Data collection procedures always impact research
- Primary data w/recognized defects often BETTER than secondary data w/unknown collection
- Gives researchers the wrong incentives.
- Encourage squeezing more information from secondary datasets
- Discourages measuring new variables
- Consequently -- more incremental
- Should avoid different standards for primary data collections
- not whether defects exist but how (or if) defects impact findings
- impossible to extract information absent from the data
- variable only important because we give it ambiguous variance
- variable only important because we overlooked other variables

Good Data

- Consistency.
  - Different error specification – same findings
  - Sub-samples/larger same -- same results
  - Observations same underlying realities
  - Not inference but predictions to future--external validity (Griliches 1985)
- Inclusiveness
  - must have information to answer the central research question
  - data must be more than interesting sidebar
  - data not be a distraction
  - findings should be almost imperious to assumptions
  - capable of refutation
- Reliability
  - Others must be able to reproduce
  - deductive research often has this property