

# BIG DATA AND ACCOUNTABILITY IN MARKETING

Oct. 16, 2014

Jean-Pierre Dubé

CHICAGO BOOTH 

Kilts Center  
for Marketing



*"Half the money I spend on advertising is wasted; the trouble is I don't know which half."*

John Wanamaker, US dept store merchant (1838-1922)

# Marketing Mix Modeling (1970s and 1980s)

- Econometric analysis of sales and marketing-mix time series data (usually POS)
- Estimates used to recommend marketing mix tactics
  - Advertising
  - Trade promotions
  - Pricing etc
- In newer digital media space, use simpler **attribution models** (although some companies now literally do use MMM for online data)
- Key: methods rely on observational data

# Findings from econometric models

- Meta-analyses of econometric studies
  - ▣ Seethuraman and Tellis (1991), Assmus, Farley and Lehmann (1984), Seethuraman, Tellis and Briesch (2013)
- Average SR ad elasticity ranges from **0.1** to **0.3**
  - ▣ Issues about data frequency and exact interpretation
- But, large implied ROI
- Have we solved Wannamaker's dilemma?
- Concerns about endogeneity in econometric studies...

# IRI and an attempt at *Advertising Accountability*

- RCT at IRI
- 1980s, IRI introduced Behaviorscan Testing (consumer level)
  - ▣ Static sample of 1800-3000 households
  - ▣ Between-household splits
- 1990s, IRI introduced the Matched-Market Test (store level)
  - ▣ Stores in approximately 32 markets
  - ▣ Between-geographic-market splits

# Was the econometric evidence too good to be true?

- Lodish et al (1995)
  - ▣ 389 split-cable experiments (new & mature products)
  - ▣ Each experiment had about 3,000 households
- Using a 20% significance level, less than 50% significant!
- Findings replicated with newer tests (Hu, Lodish and Krieger 2007)
  - ▣ Slight improvement in significance ... but using matched samples, not RCT
- Wannamaker's dilemma lives on?

# Accountability

- *The new data still suggest (as did the original data) that it is of great managerial interest to identify advertising effectiveness before launching advertising campaigns.*

*Hu, Lodish and Krieger (2007)*

- Agreed. But how?
  - Need larger samples (3,000 insufficient to get precision even at ridiculously liberal levels of significance)

# Sadly practitioners do not share this view

*Measuring the online sales impact of an online ad or a paid-search campaign—in which a company pays to have its link appear at the top of a page of search results—is straightforward: We determine who has viewed the ad, then compare online purchases made by those who have and those who have not seen it.*

M. Abraham, 2008. Harvard Business Review  
(president Comscore Networks)

- Looks like some of IRI advocates have abandoned RCT



# Testing

## One of the **Big** opportunities from Big Data

- Big Data are an opportunity for RCT
  - Create large samples for statistical power (**Volume**)
  - Low opportunity cost -- small portion of business exposed to unprofitable experimental conditions
  - Potential real-time testing (**Velocity**)

# RCT in the digital environment

- Lewis and Rao (2014)
  - ▣ 25 large field experiments with major US retailers and brokerages
  - ▣ A well-powered informative advertising experiment may require over 10 million person-weeks
- This may be too tall an order for most advertisers...
- Are Big Data really enough to solve Wanamaker's dilemma?

# RCT combined with Big Data ...

let's give it another shot

- Big Data also means lots of potential covariates
- Use covariates to reduce noise, (*post stratification* analog)

$$E[y_i | d_i, x_i] = \beta_0 + x_i' \beta + \gamma d_i$$

Addition covariates  
for each subject

Treatment effect

- In small samples,
  - ▣ Reduce variance
  - ▣ Reduce potential bias from imbalance in  $x_i$
- one company finding evidence of imbalance bias even with 20 million users!)

# RCT combined with Big Data ...

## heterogeneous treatment effects

- Big Data also means lots of potential covariates and methods to select the *ideal* subset of them

- Interact  $d_i$  and  $x_i$

$$E[y_i | d_i, x_i] = \beta_0 + x_i' \beta + \gamma_0 d_i + (d_i * x_i)' \gamma$$

- Problem:  $\dim(x_i)$  may be huge

- How do you make a decision based on this?

- Solution: use data-mining methods (e.g. Lasso) to select small number of factors

- Early findings seem very promising...