Promotions: Effects, Measurement, and ROI

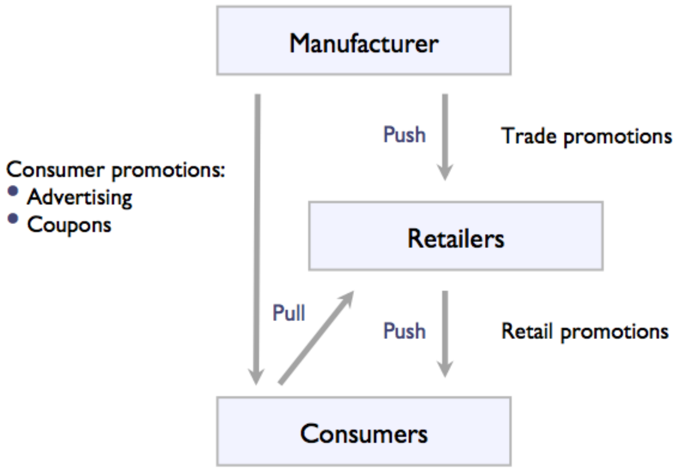
37105 Data Science for Marketing Decision Making
Günter J. Hitsch
The University of Chicago Booth School of Business

2017

Overview

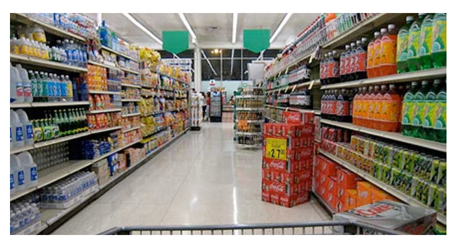
1. Overview and institutional details
    - Retail and trade promotions
    - Effects on consumer behavior
    - Promotion planning process
2. BDI/CDI analysis
3. ROI — the return on marketing investment
4. ROI-based promotions management and lift factors/promotion multipliers
    - Economics of trade promotions
    - Forward buying
5. Account and store specific marketing
    - Bayesian hierarchical models

Promotions: Overview



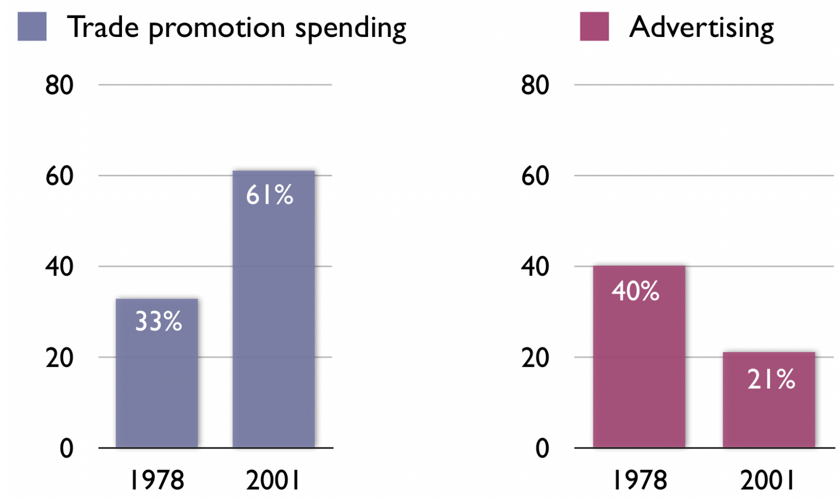
Retail promotions

- TPR's — temporary price reductions
- Display advertising
  - End-of-aisle display
  - In-aisle display
  - On-shelf display
  - Front-of-the-store display
- Feature advertising
  - Inserts in local newspapers
  - Circulars distributed around store
- Often, TPR's are accompanied by display and/or feature advertising



Trends in trade and consumer promotions: 1978-2001

Large changes in spending (percent of firm's marketing budgets):



- What caused this dramatic change in how CPG products are promoted?

Promotions: Patterns and sales effects

Analyze store-level data (Nielsen RMS scanner data)

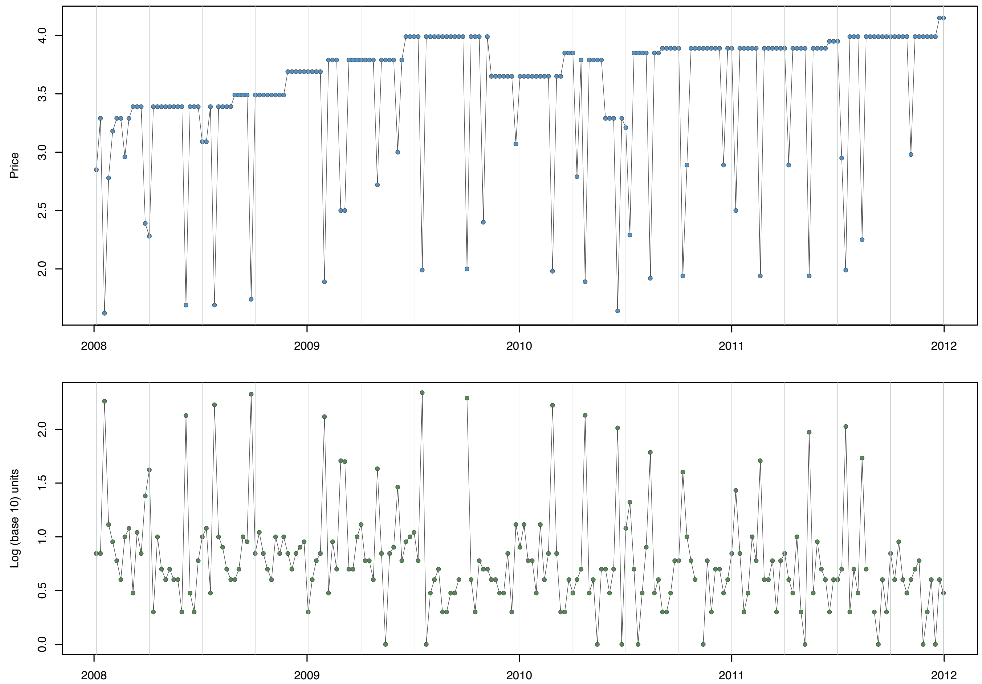
- UPC level
- Weekly

Base prices and promoted prices

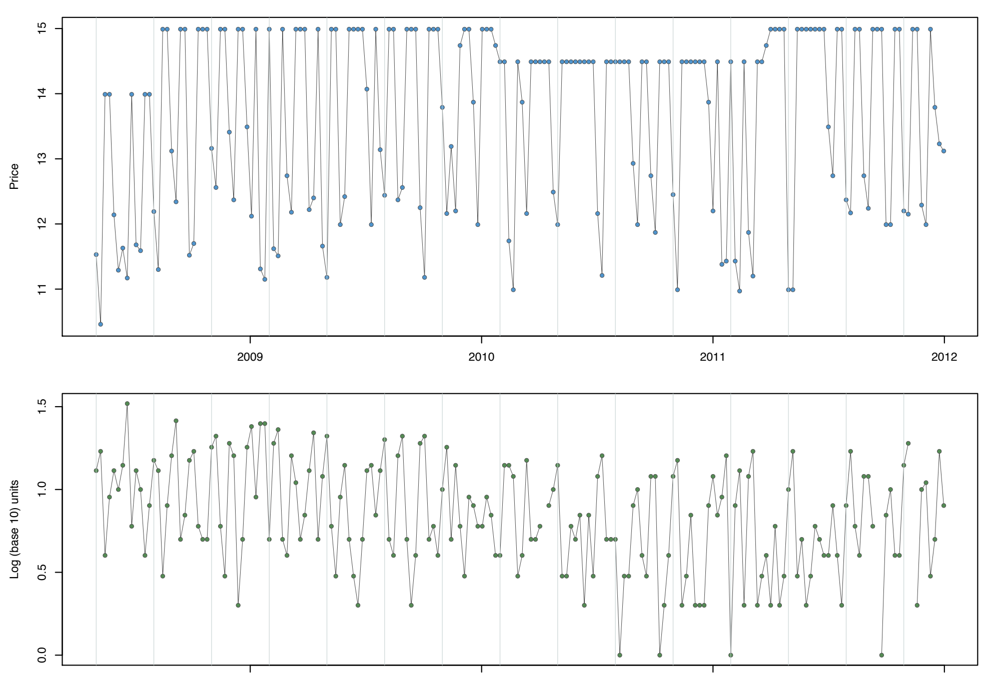
Volume increase: Note that volume is plotted in $\log_{10}$ (log base 10) units. Hence,

$$\Delta \log_{10}(Q) = 1 \Leftrightarrow \Delta Q = 10$$

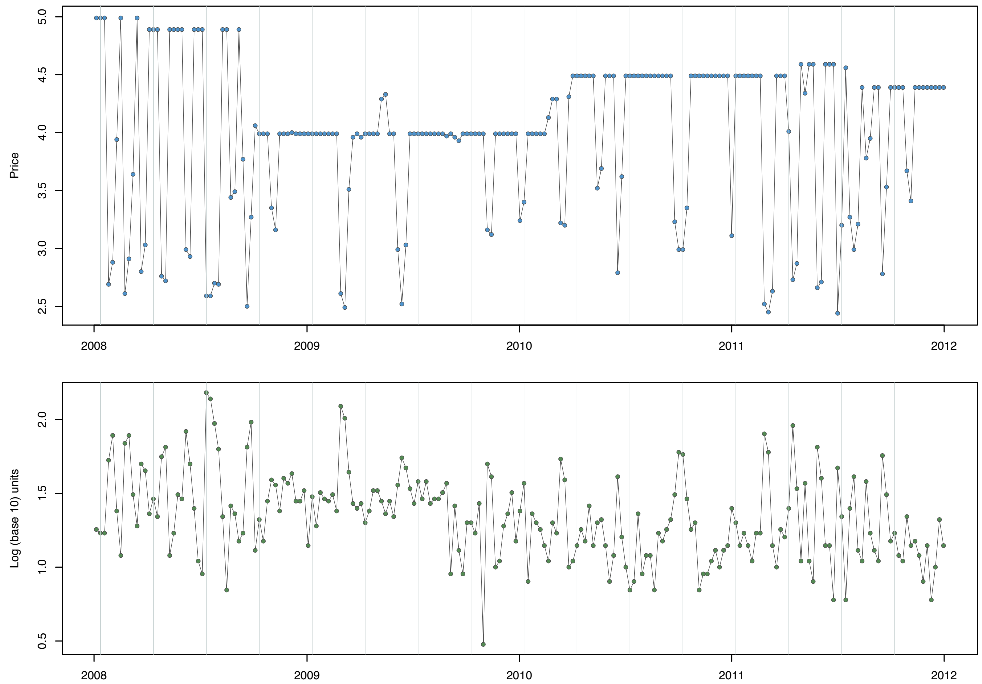
Kellogg's Raisin Bran, 20 oz



Tide Liquid, 100 oz



Kellogg's Frosted Mini Wheats, 24 oz



Sales effects of promotions

- Sales spikes at the time of promotional events are common across many brands and categories
- Possible sources of the increase in sales volume:
  - Increase in category consumption
    - Increase in number of consumers
    - Increase in household consumption of consumers already purchasing in category
  - Brand switching
  - Stock-piling (purchase acceleration)
- For many brands and categories:
  - Strong brand switching effects
  - Marketing research during the last decade also points to strong stock-piling effects

Stock-piling

- What is stock-piling?
  - Buying from deal to deal
  - Household or consumer buys several units or a large pack size when a product (brand) is promoted
  - Consumer holds inventory of the product at home
  - Consume off inventory until the next deal
- This works for storable products, such as ketchup or frozen meals
- Key question
  - Is stock-piling a good or a bad thing for profitability?

Stock-piling may enable price discrimination

- Example
  - Two groups of consumers with different willingness to pay (WTP)
    - Group 1 (price sensitive) WTP = \$2.69
    - Group 2 (insensitive) WTP = \$3.99
- **Inter-temporal price discrimination** is feasible if
  - Group 1 has a low opportunity cost of holding inventory and a low search cost (monitors price promotions, reads feature ads, ...)
  - Group 2 has a high search cost and does not monitor deals
- Implications
  - Group 2 consumers buy mostly at the regular price
  - Group 1 consumers buy at the deal price, and then consume off inventory until the next deal

Negative effects on profitability of stock-piling

- Households may be similar in their willingness to pay and change the *timing* of a product purchase when a product is promoted
  - Sales increase today is borrowed from tomorrow – cumulative purchase volume over the long-run is unaffected
  - Volume could have been sold at higher average price

Overall profitability effect of stock-piling

- In practice, we expect some price discrimination potential due to consumer heterogeneity but also changes in purchase timing due to promotions with little or no volume effect

- Can we estimate and evaluate the magnitude of the effects?
- Extremely hard and largely unsolved problem in marketing
  - Academic efforts since the late 1990's — topic in our Advanced Quantitative Marketing Ph.D. class

- Example
  - Use Homescan and observe that consumers only buy when a product is promoted
  - Is this evidence for stock-piling?

Other effects of promotions on demand

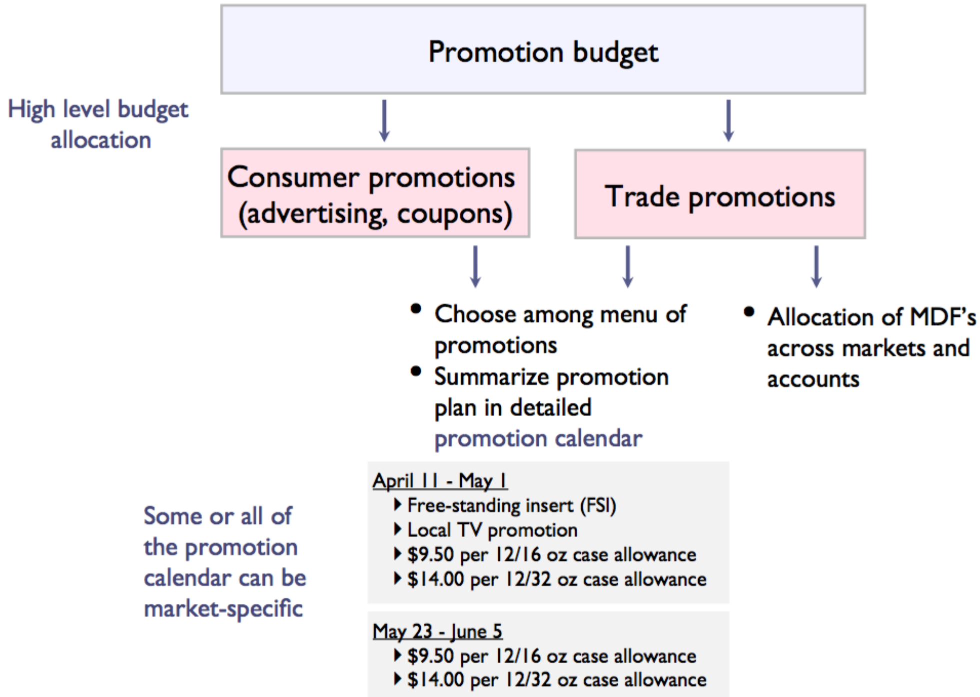
- Do promotions lead to increased consumption of a product in future, i.e. create customer loyalty?
  - Yes, according to a large body of marketing research
  - Background reading: Dubé, Hitsch, and Rossi (2010): "State Dependence and Alternative Explanations for Consumer Inertia," RAND Journal of Economics, 41(3), 417-445
- Do promotions become less effective if used more frequently?
  - Almost certainly yes
  - But difficult to incorporate into current promotion planning methods
- Do promotions degrade the brand franchise in the long run?
  - Unresolved

Institutions: Trade promotions

- Trade promotions are incentives given by the manufacturer to the retailer
  - Incentives to promote products to consumers

Off invoice allowance	Example: $10.00 off a 12/64 oz case of Tide, May 2 - June 6
Display allowance	
Advertising allowance	
Scan back	Deal only on cases sold during promotion period
MDF's (market development funds)	"Street money" — negotiated between manufacturer and retailer
Slotting allowance	Get (or keep) product on shelves

Manufacturer planning process



Retailer planning process

- Choice of products or categories to promote
- Considerations
  - Allowances or MDF's available for a product
  - Cannibalization of other products in the category
  - Does the product increase store traffic?
- Event planning
  - Timing of promotions
- Loss leaders
  - Products that increase store traffic if promoted
  - May increase store profits even if category profits decrease due to the promotion

Promotions management: BDI/CDI analysis

- Major task in promotion planning:
  - Allocate promotion budget across markets and accounts (retailers)

- Frequently used approach for promotion planning: BDI/CDI analysis
  - BDI — brand development index
  - CDI — category development index
  - Goal of the indices is to measure the "strength" of a brand or category across markets or accounts

Let's discuss the implementation of BDI/CDI analysis and its value (or lack thereof)

Definition

$$\text{BDI}_m = 100 \cdot \frac{\text{\% of product's total U.S. sales in market } m}{\text{\% of U.S. population in market } m}$$

$$\text{CDI}_m = 100 \cdot \frac{\text{\% of category's total U.S. sales in market } m}{\text{\% of U.S. population in market } m}$$

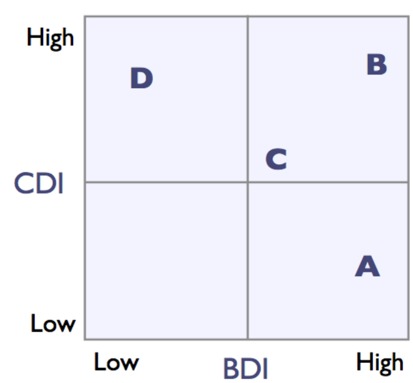
- Notes
  - Can be applied to accounts (or any other units) instead of markets
  - Substitute any other country for the "U.S."

BDI calculation: Example

Market	Sales (mill.)	Population (mill.)	Share %	Population %	BDI
A	40	40	20	13.3	150
B	10	5	5	1.7	300
C	50	75	25	25	100
D	100	180	50	60	83.3
Total	200	300	100	100	100

- Note the difference between the BDI and the share of total sales in each market
  - BDI measures brand "strength" while normalizing for market size

BDI/CDI analysis — quadrant analysis



Spending on promotions in markets A-D:

Increase, decrease, or keep unchanged?

BDI/CDI analysis and the returns on promotional spending

"Increase promotion spending in market D to grab a larger share of the category"

Is this a sensible proposal?

- Consider two scenarios:
  - Low BDI in market D is a consequence of insufficient trade spending in market D
  - Our brand doesn't sell well in market D, and historically trade spending has been less effective than in other markets

$\Rightarrow$ An increase in trade spending may make sense in scenario 1. but not in 2.

Fundamental problem with BDI/CDI analysis

- BDI and CDI measures inform us of *current* market situation
- BDI and CDI measures tell us nothing about the *return* on increasing or decreasing promotional spending

ROI — return on marketing investment

- Central concept in modern marketing analytics based on data science
- The ROI measures the return from an incremental marketing investment, such as increased trade or consumer promotional spending
- Notation: I is marketing investment spending (in dollars) and is $\pi(I)$ the corresponding profit net of I

$$\text{ROI} = \frac{\Delta\pi - I}{I} = \frac{(P-c)\cdot\Delta Q - I}{I}$$

- More concise:

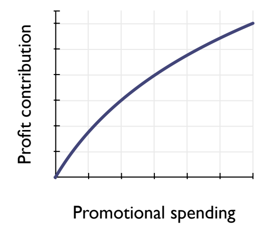
$$\text{ROI}(I) = \frac{d\pi(I)}{dI} - 1$$

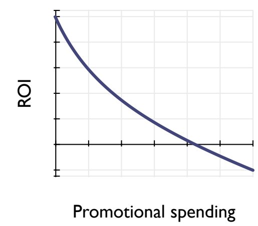
  - ROI $=$ total incremental profit from a *small* increase in I

- ROI provides a direct link between marketing and shareholder value
  - Do promotions increase or decrease profits?
  - What types of promotions are most profitable?

Making resource allocation decisions using ROI's

- Assumption: Profit increases in I, but at a diminishing rate $\Rightarrow$ ROI smaller for large than for small levels of I
- Consider multiple marketing investments, say I_1 and I_2
- Optimal allocation subject to budget constraint: If I_1^* and I_2^* are positive, then ROI$(I_1^*) =$ ROI(I_2^*)
  - Otherwise, reallocate from low to high ROI investment
- Optimal allocation without budget constraint: ROI$(I_i^*) =$ cost of capital for any investment i (unless ROI$(I_i^*) <$ cost of capital at $I_i = 0$, in which case the investment is never profitable)





Calculating ROI's

- Central difficulty: What is the *incremental* profit, $\Delta\pi$, generated by a marketing investment?
- Data-driven approach to measure the ROI:
  - Use data on sales (quantities) and marketing investments, prices, etc.
  - Estimate the sales volume for given marketing investments
- As ROI's are not directly observed, ROI-based marketing is impossible without data and statistical tools to analyze the data
  - That's why the ROI revolution in marketing occurred only in the last decades

Managing promotions using ROI's

- Key concepts
  - **Incremental effect** on sales generated by a promotion, ΔQ
  - Industry terminology: Promotional *sales lift* or *multiplier*

$$\textit{lift/multiplier} = \frac{\textit{sales } (Q)}{\textit{baseline sales}}$$

  - Baseline sales: Sales volume Q that would have occurred in the absence of a promotional event
- The incremental effect on volume and the promotional sales lift/multiplier are directly related

$$\Delta Q = \text{lift} \cdot \text{baseline sales} - \text{baseline sales}$$

Predicting incremental volume/lift using regression analysis

- Use the log-linear demand model and account for the effect of price and the characteristics (attributes) of a promotion

$$\log(Q) = \alpha - \eta \log(P) + \sum_{k=1}^{K} \gamma_j \texttt{promo\_attribute}_k + \ldots$$

  - "$+\ldots$" includes competitor prices and promotion attributes, fixed effects, etc.
- Frequently used promotion attributes:
  - Promotion event indicator (Assignment 2)
  - Data on feature and display advertising
- Example

$$\log(Q) = \alpha - \eta \log(P) + \gamma_F \texttt{feature} + \gamma_D \texttt{display} + \ldots$$

Measurement of promotion attributes

- Feature and display data frequently available to retailers based on in-house data/promotion calendar
- Captured only for some stores/chains in the Nielsen RMS data

Categorical vs. continuous measurement

- At store level, `feature` and `display` are indicator (dummy) variables
- At chain (or market) level, `feature` and `display` are numbers between 0 and 1
  - Percentage of stores where a product is featured or on display, e.g. $\texttt{feature} = 0.9$ and $\texttt{display} = 0.75$
  - Necessary because even in the same chain not all stores may display or feature a product
- Also used: Store-size weighted `feature` and `display` measures
  - Example: `%ACV_feature` and `%ACV_display` are the percentage of ACV on feature or display, i.e. percentage of stores that feature or display a product, accounting for differences in store size

Calculating lift factors from the regression estimates

- Use `predict`, or a simple formula
- Keep competitor variables and all other controls (fixed effects, time trend, etc.) constant. Calculate
  - Baseline (non-promoted) sales
  - Sales under a specific promotion event with percentage TPR r, `feature`, and `display`

$$\begin{aligned}
\log(Q) &= \alpha - \eta\log(P) + \gamma_F \cdot 0 + \gamma_D \cdot 0 + \dots \\
\log(Q') &= \alpha - \eta\log(P') + \gamma_F \texttt{feature} + \gamma_D \texttt{display} + \dots \\
\Rightarrow \log(Q') - \log(Q) &= -\eta\left(\log((1-r)P) - \log(P)\right) \\
&\quad + \gamma_F \texttt{feature} + \gamma_D \texttt{display} \\
\Rightarrow \text{lift} = \frac{Q'}{Q} &= \exp\left(-\eta(1-r) + \gamma_F \texttt{feature} + \gamma_D \texttt{display}\right)
\end{aligned}$$

- Illustrates how to interpret the coefficients on `feature` and `display`. Example: $r = 0$, $\texttt{feature} = 0$, and $\texttt{display} = 1$:

$$\Rightarrow \text{lift} = \frac{Q'}{Q} = \exp(\gamma_D)$$

Example

Estimate price, feature, and display effects at the retail account level

```
fit = lm(log(sales_units) ~ log(price) + feature_percentage
                              + display_percentage,
         data = promo_DT[account == retailer_name])
```

```
Coefficients:
                   Estimate Std. Error t value Pr(>|t|)
(Intercept)         9.52123    0.08944 106.451  < 2e-16 ***
log(price)         -1.84318    0.45032  -4.093 9.74e-05 ***
feature_percentage  0.28531    0.08925   3.197  0.00196 **
display_percentage  0.83410    0.17653   4.725 9.14e-06 ***
```


Economics of trade promotions

- Next table: Main principles of the economics of trade promotions
- See Promotion-ROI.xlsx
  - Use Excel formulas or script in R (or any other suitable language) to automatically
    - Calculate the event lift factor from the regression estimates and the event structure
    - Calculate the event ROI from the lift factor and the margin and cost data
  - Allows for scalability
    - ROI's corresponding to different events
    - ROI's across accounts or markets
- Purchase acceleration: Based on assumption !!!
  - Predicting volume due to stock-piling is frontier topic in marketing research
  - No industry implementation yet

	Account A			Account B		
Regression estimates						
Intercept	9.5212	9.5212	9.5212	10.0888	10.0888	10.0888
log(price)	-1.8432	-1.8432	-1.8432	-1.8974	-1.8974	-1.8974
feature_percentage	0.2853	0.2853	0.2853	0.0000	0.0000	0.0000
display_percentage	0.8341	0.8341	0.8341	1.0695	1.0695	1.0695
Event structure						
Price (% change)	-0.15	-0.15	-0.15	-0.15	-0.15	-0.15
Feature percentage	0.0	0.0	1.0	0.0	0.0	1.0
Display percentage	0.0	0.7	0.7	0.0	0.7	0.7
Lift from event	1.35	2.42	3.22	1.36	2.88	2.88
Base price	1.20	1.20	1.20	1.20	1.20	1.20
Promoted price	1.02	1.02	1.02	1.02	1.02	1.02
Margin ($)	0.55	0.55	0.55	0.55	0.55	0.55
Margin at promoted price	0.37	0.37	0.37	0.37	0.37	0.37
Units						
Baseline units	9,751	9,751	9,751	17,032	17,032	17,032
Incremental units (model prediction)	3,406	13,839	21,628	6,152	31,983	31,983
Total units	13,157	23,591	31,379	23,184	49,015	49,015
% incremental units purchase acceleration	25	25	25	25	25	25
Incremental units net of purchase acceleration	2,554	10,379	16,221	4,614	23,987	23,987
Manufacturer promotion P&L						
Incremental contribution	945	3,840	6,002	1,707	8,875	8,875
Variable cost (baseline units)	1,755	1,755	1,755	3,066	3,066	3,066
Fixed payment cost	0	2,800	3,800	0	4,750	6,500
Event cost	1,755	4,555	5,555	3,066	7,816	9,566
Event gross contribution	-810	-715	447	-1,359	1,059	-691
ROI	**-46.2**	**-15.7**	**8.0**	**-44.3**	**13.6**	**-7.2**

Forward buying

- Abraham and Lodish (*Harvard Business* Review, 1990) discuss a particular and highly important cost of trade promotions
- An analysis of trade promotions for all brands in 65 different product categories suggests only 16 percent of the promotions are profitable
- Why? — See economics of trade promotions on the next slide
- Lesson
  - Need to include foregone contribution from forward buying as an additional event cost in ROI calculation
- Question
  - What contracts or incentives can a manufacturer use to prevent forward buying by the retailer?

	Cases	Gross dollars
Baseline (Sales that would have occurred during the four-week promotion period even without the promotion)	400	$4,000
Incremental sales to consumer •Due to one week of feature •Due to 50% of stores with three weeks of display and price reduction •Due to 50% of stores with four weeks of price reduction only Total	 100 250 80 430	 $1,000 $2,500 $800 $4,300
Ten weeks of forwarding buying by retailers	1,000	$10,000
Total sales during promotion	1,830	$18,300
Cost of promotion ($18,300 x 15% discount)		$2,745
Cost of incremental sales (Promotion cost divided by total incremental sales)		$0.64 (=2,745/4,300)

Assumptions:

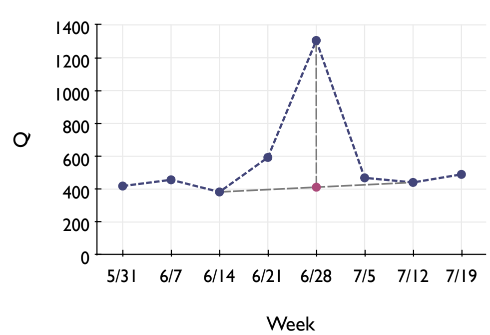
- Four week period
- Base sales are 100 cases per week
- List price is $10 per case
- 15% off-invoice offered for four weeks

$$\text{ROI} = \frac{\text{margin} \cdot 4300 - 2745}{2745}$$

Seasonality

- Many products and categories exhibit seasonal demand patterns trends
  - Christmas
  - 4th of July — beer, soft drinks, and ketchup

- Promotions often occur during periods of high demand due to seasonality
  - Implication for lift estimates if we do not control for seasonality?



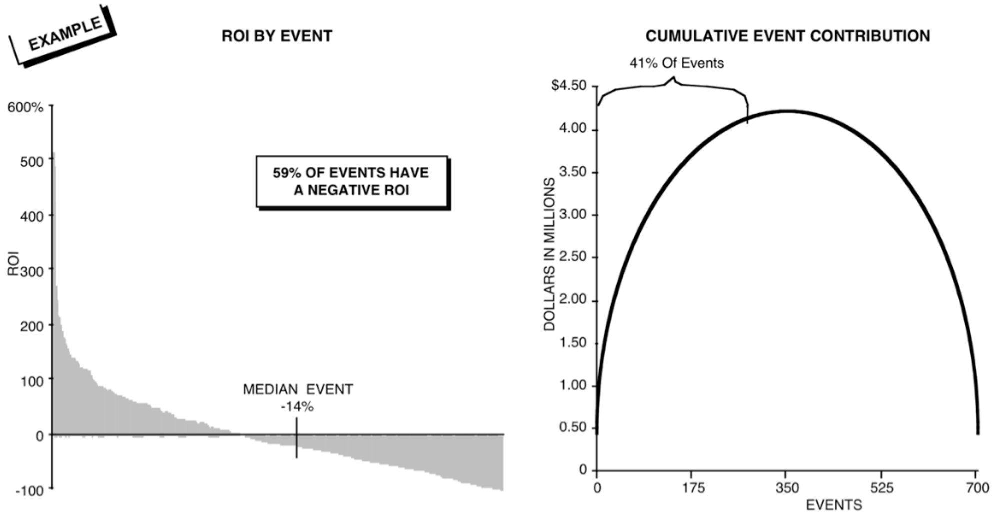
Controlling for seasonality in a demand model

- Dummy variables for specific calendar events or time of the year
  - Dummy variables for one or two weeks before the 4th of July
  - Christmas holiday season dummy
  - Month or week dummies — preferred solution with big data
- Also used:
  - Include average regional temperature as independent variable

The state of the art: Account and store specific marketing

- Manufacturer decisions
  - Promotional fund allocation across retail accounts
- Retailer decisions
  - Whether to run the same or different promotions across stores or zones
- Should these fund allocation decisions be uniform across accounts or stores?
- The Booz Allen Hamilton study (next slide) shows the rationale for engaging in account or store specific marketing

WHAT IS NEEDED IS PROFIT CALIBRATED EVENT-BY-EVENT UNDERSTANDING—THE RESULTS TYPICALLY ARE EYE-OPENING: FOR EXAMPLE, AT ONE CLIENT:

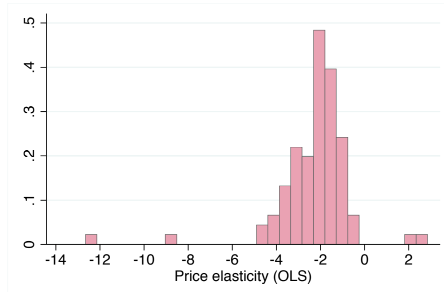


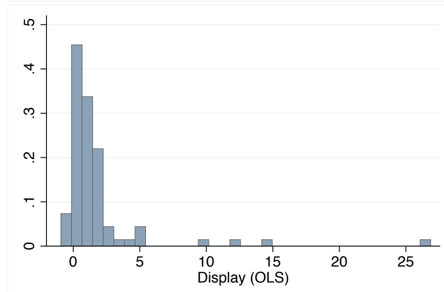
State of the art industry practice

- Manufacturers
  - Trade promotional fund allocation based on account level ROI's
- Retailers
  - Choose from menu of retail promotions based on store (or market zone) level ROI's
- Implementation
  - Predict ROI's for promotions across accounts or stores (zones) to make promotional budget allocation decisions
  - Requires demand estimation at each account or store

Account specific demand estimates: Regression analysis

- Histogram shows regression estimates for a branded cheese product from 88 accounts
  - 52-68 sales, price, and display observations per account
  - Price elasticities
  - Display coefficients
- Problem:
  - Some large price elasticities and display coefficients
  - Some positive price elasticities
- Why?





$\exp(14) = 1202604$

Bayesian hierarchical models

- Central problem:
  - Insufficient number of observations to reliably estimate demand models at the account or store level using regression analysis
- State of the art approach to obtain individual (account or store level) demand parameters: Bayesian hierarchical models
  - Employed in the industry by IRI and consulting companies such as IBM DemandTec
  - Implementing Bayesian methods requires a lot of computing power—hence these methods were not practical for data driven marketing until about 15-20 years ago
  - See *Bayesian Statistics and Marketing*, by Rossi, Allenby, and McCulloch

Bayesian hierarchical models: Central ideas

Goal:

- Obtain reliable parameter estimates for each account or store

$$y_j = X_j \beta_j + \varepsilon_j$$

  - β_j parameter vector for account j
- Problem: Regression estimates not reliable unless we have many observations

However:

- Typically we can reliably estimate the distribution of parameters across accounts (stores)
- Example: Distribution of parameters is multivariate normal

$$\beta_j \sim N(\mu, \Sigma)$$

  - μ mean parameter vector across all accounts
  - Σ variance-covariance matrix across accounts

Parameter estimates using Bayes' law

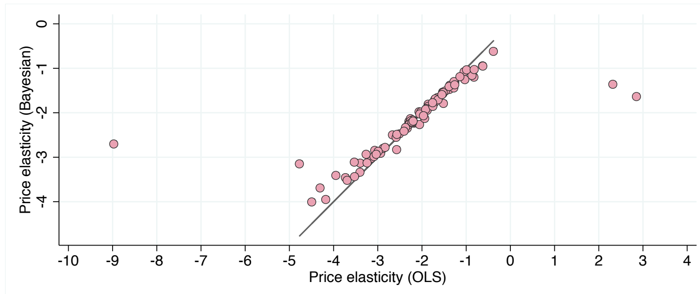
$$\beta_j = (\Sigma^{-1} + \sigma_j^{-2} X_j^T X_j)^{-1} (\Sigma^{-1}\mu + (\sigma_j^{-2} X_j^T X_j)\beta_j^{OLS})$$

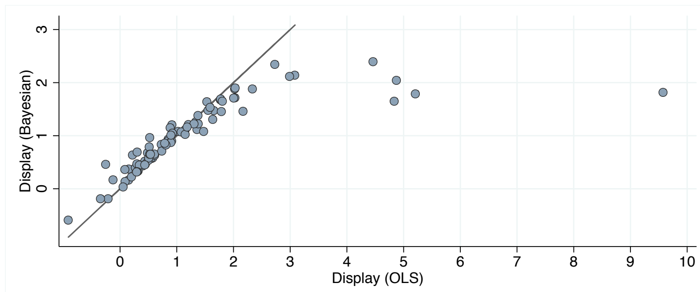
Σ^{-1}	inverse of variance of parameters across accounts
μ	mean parameter vector across all accounts
X_j^T	inverse of variance of regression estimates
β_j^{OLS}	regression (OLS) estimate

- Estimate is a weighted average that combines:
  - The mean (average) of the parameters across all accounts
  - The account level regression estimate (based on the OLS method)
- More weight is given to the OLS estimate if the variance of the estimate is small $\Leftrightarrow$ inverse of variance of parameters is large
- More weight is given to the mean of the parameters across accounts if the variance of the parameters across accounts is small $\Leftrightarrow$ inverse of the variance of parameters across accounts is large

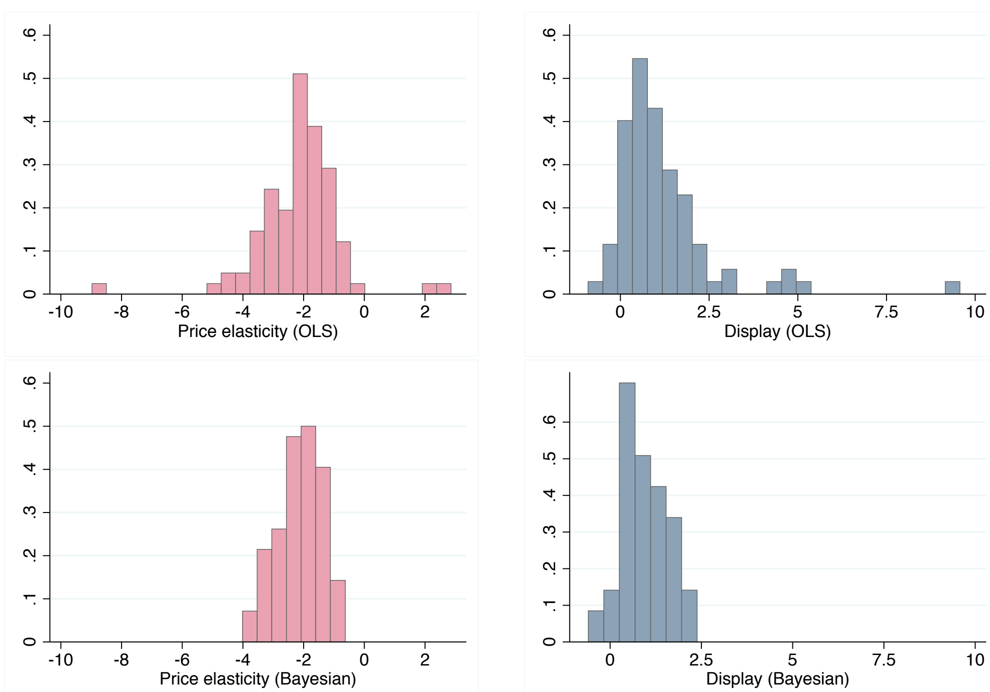
Bayesian shrinkage

- The Bayesian approach shrinks the original regression (OLS) estimates if the estimated variance of the regression parameters is large





OLS vs Bayesian shrinkage estimates



Bayesian hierarchical models: Summary

- Bayesian hierarchical models allow us to estimate price elasticities, promotion effects, etc., for individual cross-sectional units, such as accounts or stores
  - Works even if there are only few data points available for each unit
  - OLS estimate will automatically be shrunk towards the overall mean if the estimate is not precise
- Bayesian hierarchical models are now widely applied in various areas of quantitative marketing
  - Conjoint analysis
  - Household-level choice models

Summary

- Retail promotions and consumer behavior
- Trade promotions are complex incentives given to the retailer
- BDI/CDI analysis
  - Does not assess the profitability consequences of promotions
- Modern data-driven marketing: Resource allocation decisions based on the return of marketing investment (ROI)
- Lift factors measure the incremental sales due to a promotion
  - Estimation based on regression model
- Economics of trade promotions
  - Limit forward buying
- Increase overall ROI using account or store specific marketing
  - Implemented using Bayesian hierarchical models