Creating Retail Decision Support Systems using Consumer Transaction Data

Alan Montgomery
Associate Professor
Carnegie Mellon University
Tepper School of Business

e-mail: alanmontgomery@cmu.edu
web: http://www.andrew.cmu.edu/user/alm3

University of Hong Kong, 19 July 2005

© 2005 by Alan Montgomery, All rights reserved.
Outline

• Introduction
• Quantitative Approaches to Pricing
• Pricing Decision Support Systems
• The Determinants of Price Elasticity
• Micro-Marketing Pricing Strategies
• Experimental Results
• Conclusions
Introduction

Retail Pricing using Transaction Data
Retail Pricing and Promotion
Collecting Transaction Data

Why do companies collect this information?
Category Management
The Next Generation of Transaction Systems

- RFID has the potential to “revolutionize” the marketplace

  - Examples:
    - Electronic pricing environments (e.g., e-commerce sites and electronic shelf labels) enable price experimentation in real-time

- What happens to price, product, promotion, and placement?
Quantitative Approaches to Pricing
Goal

• How can we better exploit the *installed base* of data:
  - Store level scanner data

• To make these types of decisions at *low cost*:
  - Everyday pricing
  - Zone/store pricing decisions
  - Planning the promotional calendar
  - Trade promotions
Weekly Movement and Price of TropPrem64

![Graph showing weekly movement and price of TropPrem64 over a period from 01/02/1988 to 10/02/1991. The graph displays fluctuations in movement and price over time.]
Movement vs Price of TropPrem64
Movement vs Price of TropPrem64
Movement vs Price of TropPrem64

Chicago #6

Chicago #58
Statistical Demand Models

- Relate movement of each product to its price changes
- Consider prices of other products within the category
- Estimate the effects of feature ads, in-store displays, and shelf-tags

\[
\ln(q_{bst}) = \alpha_{bs} + \mu_{bs} \ln(x_{bst}) + \sum_{i=1}^{M} \eta_{bis} \ln(p_{ist}) + \varphi_{bs} f_{bst} + \delta_{bs} d_{bst} + \varepsilon_{bst}
\]

- quantity
- total store volume
- expenditures
- own + cross price effects
- out-of-store features/ads
- in-store deals/ads
Evaluating This Approach

Advantages
- Can “learn” about consumers based on their past behavior
- Leverages data warehouse
- Summarizes complicated behavior
- Easy to use

Disadvantages
- Complicated to build
- Forecasts can be wrong
- Can conflict with our intuition
- How do we “prove” the models are correct
Profitability of TropPrem64 at Pittsburgh #637 (Cost=$2.40)
Decision Support Systems for Retail Pricing

Massive Datasets and Massive Decision Problems
Goal

• How can we better exploit the *installed base* of data:
  - Point-of-Sale Data (Scanner/Loyalty programs)

• To make these types of decisions at *low cost*:
  - Everyday pricing
  - Zone/store pricing decisions
  - Planning the promotional calendar
  - Trade promotions
Decision Support Systems must be...

- Simple
- Robust
- Easy to control
- Adaptive
- As complete as possible
- Easy to communicate with

See Little (1970, 1979)
Pricing Decision Support

Market Simulation Model

Decision Variables

<table>
<thead>
<tr>
<th>Brand Description</th>
<th>Carton Price</th>
<th>Feature</th>
<th>In-Store Display</th>
<th>Expected Movement</th>
<th>Wholesale Cost</th>
<th>Profits</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 TropPrem64</td>
<td>$2.89</td>
<td>No</td>
<td>Yes</td>
<td>12248</td>
<td>$1.75</td>
<td>$217.27</td>
</tr>
<tr>
<td>2 TropPrem96</td>
<td>$3.79</td>
<td>Yes</td>
<td>No</td>
<td>21139</td>
<td>$3.32</td>
<td>$103.01</td>
</tr>
<tr>
<td>3 TropReg64</td>
<td>$2.29</td>
<td>No</td>
<td>No</td>
<td>3566</td>
<td>$1.49</td>
<td>$44.64</td>
</tr>
<tr>
<td>4 MinMaid64</td>
<td>$2.24</td>
<td>No</td>
<td>No</td>
<td>8459</td>
<td>$1.67</td>
<td>$75.52</td>
</tr>
<tr>
<td>5 Store64</td>
<td>$1.79</td>
<td>No</td>
<td>No</td>
<td>9106</td>
<td>$1.20</td>
<td>$83.96</td>
</tr>
</tbody>
</table>

Category Profits: $524.41

- Dynamically forecast movement, revenue, profit
- Manipulate price, display, advertising, wholesale cost
- Produce forecasts for store, chain, zone level
- Provide multi-week planning horizon
- Measure both acquisition and wholesale costs
- Manage promotional calendars
- Optimizer to suggest best pricing strategy (either by groups or all)
Pricing DSS

- Detect price response using historical transaction data
- Forecast movement, revenue, profit in real-time
- Produce weekly forecasts at the chain, zone, and store level
- Manipulate price, feature, display, and wholesale cost in an interactive environment
- Change prices for groups of products
- Provide a multi-week planning horizon in order to manage promotional calendars
- Work with incomplete information
- Coordination across categories and stores
- Integrate information from many sources
- Scalability
- Recommend price strategies
Problems

- **Data**
  - Historical data from the warehouse
  - Current inventory
  - Anticipated prices from promotions/competitors
  - Continually changing inventories/product assortments

- **Modeling**
  - How to model 200-10,000 SKUs per category?
  - What about 100-2,000 stores in the chain?
  - Over 300 categories per store?
  - Seasonal patterns

- **Inference**
  - Prevent model from making bad predictions
  - Need to consider promotional calendar
  - Making Optimization Decisions
Illustrating our Information Flow in our Pricing DSS

1. Data Warehouse
2. Data Preparation
3. Modeling
4. Price Selection/Optimization
5. Price Staging
6. Retail Pricing
7. Stores
8. Consumers

Competitor Channel
<table>
<thead>
<tr>
<th>Company</th>
<th>Location</th>
<th>Website</th>
<th>Founded</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACNielsen</td>
<td>New York, NY</td>
<td>Acnielsen.com</td>
<td>1923, 1995*</td>
</tr>
<tr>
<td>Applied Predictive Technologies</td>
<td>Arlington, VA</td>
<td>predictivetechologies.com</td>
<td>1999</td>
</tr>
<tr>
<td>DemandTec</td>
<td>San Carlos, CA</td>
<td>Demandtec.com</td>
<td>1999</td>
</tr>
<tr>
<td>I2</td>
<td>Dallas, TX</td>
<td>i2.com</td>
<td>1988</td>
</tr>
<tr>
<td>Evant</td>
<td>San Francisco, CA</td>
<td>nonstop.com</td>
<td>1994</td>
</tr>
<tr>
<td>KhiMetrics</td>
<td>Scottsdale, AZ</td>
<td>khimetrics.com</td>
<td>1993, 2000*</td>
</tr>
<tr>
<td>Knowledge Support Systems</td>
<td>Florham Park, NJ</td>
<td>kssg.com</td>
<td>1993</td>
</tr>
<tr>
<td>Manugistics</td>
<td>Rockville, MD</td>
<td>Manugistics.com</td>
<td>2001*</td>
</tr>
<tr>
<td>Marketmax</td>
<td>Wakefield, MA</td>
<td>marketmax.com</td>
<td>2003*</td>
</tr>
<tr>
<td>Maxager Technology</td>
<td>San Rafael, CA</td>
<td>maxager.com</td>
<td></td>
</tr>
<tr>
<td>Metreo</td>
<td>Palo Alto, CA</td>
<td>metreo.com</td>
<td>2000</td>
</tr>
<tr>
<td>ProfitLogic</td>
<td>Cambridge, MA</td>
<td>profitlogic.com</td>
<td>1984, 2001*</td>
</tr>
<tr>
<td>Retek</td>
<td>Minneapolis, MN</td>
<td>retek.com</td>
<td>1986, 1996*</td>
</tr>
<tr>
<td>Zilliant</td>
<td>Austin, TX</td>
<td>zilliant.com</td>
<td>1998</td>
</tr>
</tbody>
</table>
Our Modeling Contributions

- Hierarchical Bayesian Models
  - Estimation of product X store level models
  - Use of Monte-Carlo Markov Chain methods for estimation and inference
- Real-time Inference
  - Use of simulated estimates incorporates uncertainty over parameter estimates and prior information. Avoids approximations commonly used in practice.
- Informative Priors
  - Shrinkage estimators allow borrowing of information across stores to improve estimates
  - Allow managers and analysts to incorporate prior information about parameter values and optimal prices
The Determinants of Price Elasticity
Why do price elasticities vary?

- Do all stores have the same price response profiles?
- What explains these differences?
- How do you customize a pricing strategy that appeals to a store’s trading area?

Implementing a Micro-Marketing Strategy
Are there really differences across stores and brands?

Legend:
- Very Price Sensitive
- Moderate Price Sensitive
- Not Price Sensitive
What explains these differences? Demographics + Unique Store Profiles

Legend:
- Few College Educated Adults
- Moderate Education Levels
- Highly Educated Areas
Summary of Results

+ Elderly      Can devote more time to price search and shopping
- Education    Higher opportunity costs, less attention to shopping
+ Ethnic       Proxy for other causal factors
? Income       Dependent upon category
+ FamilySize   Larger share of disposable income on groceries, increased returns to search
+ Working Women Tighter constraints on household budget
- House Value  Fewer income constraints
- Competitor Dist. Isolated stores less price sensitive
+ Relative Volume Consumers self-select for location and convenience or price and assortment
Micro-Marketing Pricing

Developing Models for every Store x Item
Store-Level Strategies

• The previous results show that there are differences in how consumers respond to price changes across stores

• How do you cater to neighborhood store preferences?
  – Different product assortments
  – Store-level everyday pricing
  – Unique in-store promotions
  – Customized store features
Movement vs Price of TropPrem64

Chicago #6

Chicago #58
Profitability of TropPrem64 (Cost=$2.40)

Profit vs Price

- Chicago #6
- Chicago #105
What is the problem with the usual regression approach?

Difficult to acquire reliable estimates for individual products at store-level

Model Dimension
Stores x Brands x Regressors
= 100 x 10 x 14 = 14,000 parameters

Data Dimension =
Stores x Brands x Weeks
= 100 x 10 x 156 weeks = 156,000 data points

Frustrated use of this data in industry!

© 2002 by Alan Montgomery
What can be done?

Two extremes:

- **Pooling**
  - Ignore all store differences

- **Individual Store Models**
  - Difficult to estimate

Our Solution:

- **Shrinkage**
  - Exploit commonalities across stores to improve individual store estimates

© 2002 by Alan Montgomery
Hierarchical Bayesian Setup

\[ y_{1,1} = X_{1,1}\beta_{1,1} + \varepsilon_{1,1} \]
\[ \vdots \]
\[ y_{s,m} = X_{s,m}\beta_{s,m} + \varepsilon_{s,m} \]
\[ \vdots \]
\[ y_{S,M} = X_{S,M}\beta_{S,M} + \varepsilon_{S,M} \]

\[ \beta_{s,m} \sim N(\bar{\beta}_m, V_\beta) \]

How do we set our prior?

Quantity
Elasticities
Expenditures, Price, Feature, Deal
Shrinkage estimates

• We are exploiting commonalities across the stores to improve the estimates

• Our approach incorporates several new theoretical developments:
  - Shrinkage estimation using Monte Carlo Markov Chain methods
  - General approach to estimation
  - Incorporation of informative priors to specify similarities across stores, model structure, and information about price solution
Visualizing the Priors Effect: Shrinkage toward the Prior
Variation in Tropicana Own Price Elasticity

© 2002 by Alan Montgomery
Setting Store-Level National Brand/Store Brand Price Gaps

Expected Profits for Store # 6

Expected Profits for Store # 58
Effects of a reduction in the price multiplier

![Graph showing the effects of a reduction in the price multiplier. The graph plots the price for Minute Maid 64 oz over 50 weeks. There are three lines representing the base price, a 10% everyday price reduction, and the effects of in-store displays and features. The graph demonstrates how the price varies over time with and without the reduction.]
# Expected Profits from Micro-Marketing Pricing Strategies

<table>
<thead>
<tr>
<th>#</th>
<th>Description of Pricing Strategy</th>
<th>Expected Profits</th>
<th>Expected Increase</th>
<th>% Change in Expected Profits</th>
<th>Prob[Expected Increase &gt;0]</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Uniform Prices across all Stores</td>
<td>$3,330,900</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(11,900)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Optimal Uniform Strategy</td>
<td>$3,344,100</td>
<td>$13,200</td>
<td>+.4%</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(11,600)</td>
<td>(1,700)</td>
<td>(.1)</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Optimal Micro-Marketing Strategy</td>
<td>$3,459,000</td>
<td>$128,100</td>
<td>+3.9%</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(19,400)</td>
<td>(18,800)</td>
<td>(.6)</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Optimal Micro-Marketing Strategy</td>
<td>$3,481,600</td>
<td>$150,700</td>
<td>+4.5%</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>with constraints at the Chain-level</td>
<td>(20,900)</td>
<td>(20,200)</td>
<td>(.6)</td>
<td></td>
</tr>
</tbody>
</table>

Note: The standard deviation of the posterior are given in parentheses below the posterior means.
Effects of Price Changes from an Optimal Pricing Strategy

Store Changes in Everyday Price Multipliers

Optimal Uniform Chain Strategy

Product
Findings

• Every store is different and has its own price response profile
  - We can identify the price profile of a store using historical data
  - The most important determinants are store demographics (to a lesser extent competitive characteristics)

• Micro-marketing presents a rich environment for store-level pricing
  - Do not simply increase all prices up or down
  - Manage the price gaps between the brands to encourage substitution towards more profitable baskets or products

• Can recommend better or optimal pricing strategies
Experimental Results
An Experimental Approach

• A simple and reliable method (albeit costly) to find a better pricing strategy is through experimentation:
  - Divide 86 stores into 3 treatments: Control (leave prices unchanged), EDLP (decrease prices by 7%), and Hi-Lo (increase prices by 7%)
  - Measure change in profits and movement and compare them to control group
Experimental Results

Pricing Experiment Results

- EDLP: +3%
- Control: -18%
- Hi-Lo: +17%
EDLP vs Hi-Lo

• Clearly Hi-lo is much more profitable
  - Why was Dominick’s hesitant to implement these results?

• What is driving these results is price sensitivity
  - If products were more price sensitive than EDLP would work
    (price decreases would dramatically increase sales)
  - What determines price sensitivity?
Elasticity Based Zone Assignments

<table>
<thead>
<tr>
<th>Price Sensitivity</th>
<th>Low Price Zone</th>
<th>Medium Price Zone</th>
<th>High Price Zone</th>
</tr>
</thead>
<tbody>
<tr>
<td>High Price Sensitivity</td>
<td>2</td>
<td>14</td>
<td>4</td>
</tr>
<tr>
<td>Medium Price Sensitivity</td>
<td>6</td>
<td>25</td>
<td>11</td>
</tr>
<tr>
<td>Low Price Sensitivity</td>
<td>1</td>
<td>10</td>
<td>10</td>
</tr>
</tbody>
</table>
Conclusions
Findings

• The Retail Market is moving towards integrated supply-channel and demand-based pricing solutions
• Many challenging modeling problems have been addressed, however there are still much to be done
• Strong potential for increased profitability and efficiency that can benefit both the retailer and the consumer