Leveraging Primary Market Research Synchronized with Data Mining Models

Thomas T. John, Ph.D.

Lead Statistician, KS&R, Inc.

Adjunct Faculty, Syracuse University

C.R.Rao Advanced Institute of Mathematics, Statistics & Computer Science

INTERNATIONAL CONFERENCE ON BUSINESS DATA MINING

School of Management Studies, University of Hyderabad

December 30, 2008

Overview

- Introduction
- Choice Models
- Applications of Choice Models
- Segmentation/Classification
- Business Need for a “Bridge"
- Possible Approaches
- Discussion

Introduction

- Sophisticated data mining models built & utilized by businesses
  - Clustering/classification schemes
  - Neural/social networks
  - Based on existing & evolving databases
  - Leverages past behavioral patterns
- Businesses also conduct many market survey research (to hear the “voice of the customer")
  - Purchase preference/choice models & Demand forecasting
  - Market sizing/price optimization
  - Brand assessment/tracking
  - Customer equity/loyalty models
  - Customer Loss/Churn drivers

Introduction - continued

- We will discuss possibilities of an integrated business strategy utilizing both:
  - Discrete Choice Models - based on survey research statistically designed for:
    - Product & Price Optimization
    - Demand Forecasting
  - Clustering/Classification Models - based on data mining built for:
    - Custom Market Segment Strategies
    - Customer Retention & Churn Reduction

Discrete Choice Model Theory - Overview

- Popularized by Prof. Daniel McFadden
  - 2000 Nobel Laureate in Economics
  - Seminal paper in *Frontiers in Econometrics* (1973)
- Random Utility Maximization (RUM)*:
  - Set of alternatives: $\mathcal{C} = \{\mathbf{X}_1, \ldots, \mathbf{X}_p\}$
  - An alternative $\mathbf{X}_i \in \mathcal{C}$: $\mathbf{X}_i = (X_{i1}, \ldots, X_{iq})$
    X_{ij} are attribute specifications of the i-th alternative.
  - For an individual k, utility for $\mathbf{X}_i$:

$$U_{ik} = V_{ik} + \epsilon_{ik} = \beta_k' \mathbf{X}_i + \epsilon_{ik}.$$

  - Given a choice set $C \subseteq \mathcal{C}$, we assume that an individual chooses an alternative $\mathbf{X}_{i^*} \in C$ if $U_{i^*k} > U_{ik}$, $\forall \mathbf{X}_i \in C, i \neq i^*$

* *Model specifics condensed/restated for simplicity*

Discrete Choice Model Theory - Overview - continued

- Recall $U_{ik} = V_{ik} + \epsilon_{ik} = \beta' \mathbf{X}_i + \epsilon_{ik}$.
  Assuming different distributions for ϵ gives different models:
  - $\epsilon_{ik} \sim$ Type-I Extreme $\Rightarrow$ MNLogit-model $\Rightarrow$

$$P(\{\text{Individual } k \text{ chooses } \mathbf{X}_i\}) = \frac{e^{V_{ik}}}{\sum_C e^{V_{ik}}}$$

  - $\epsilon_{ik} \sim$ Normal $\Rightarrow$ MNProbit-model $\Rightarrow$

$$P(\{\text{Individual } k \text{ chooses } \mathbf{X}_i\}) = \Phi(V_{ik})$$

- Taking the probabilities as % of time an individual will choose $\mathbf{X}_i \Rightarrow$ “Share of Preference" approach.
- Original assumption that an individual chooses an alternative $\mathbf{X}_{i^*}$ if $U_{i^*k} > U_{ik} \Rightarrow$ “First Choice" approach.

Choice Model Execution via Market Survey Research

- Proper experimental design for the choice exercise
  - Commonly used criteria is D-efficiency which for $N \times p$ design matrix $\mathbf{Y}$ is: $1/\left(N\left|(\mathbf{Y}'\mathbf{Y})^{-1}\right|^{1/p}\right)$.
- Proper sampling design for the survey (e.g., stratified)
- Model estimation (β_k for each individual k) via robust Hierarchical Bayes estimation:
  - Assume $\beta_k \sim \mathrm{Normal}(\alpha, \mathbf{D})$
  - Iterative estimation via MCMC/Gibbs Sampling (one of $\beta, \alpha, \mathbf{D}$ estimated conditional on current values of the other two)
  - α drawn from Normal distribution and $\mathbf{D}$ drawn from Inverse Wishart distribution. β estimated using Metropolis-Hastings Algorithm (assessing predictability of observed choices via MNL model).

Prediction From Choice Model - *A Case Study**

- Category Level Predictions

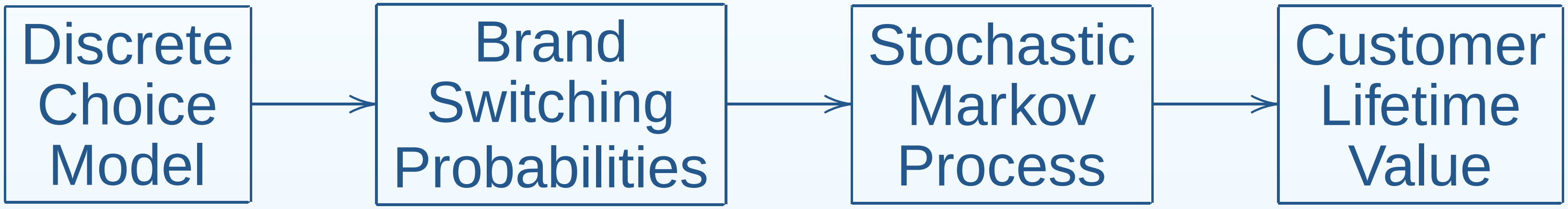
	Volume Share		Retail $ Share		WS $ Share	
	Actual**	Pred	Actual**	Pred	Actual**	Pred
Category A	35.4%	28.3%	20.6%	17.7%	18.9%	16.0%
Category B	20.5%	20.2%	12.6%	12.6%	11.0%	10.6%
Category C	4.8%	4.2%	5.2%	4.4%	5.3%	4.2%
Category D	16.3%	21.2%	17.6%	22.8%	18.6%	24.1%
Category E	8.0%	5.4%	16.8%	11.5%	18.8%	12.2%
Category F	13.7%	18.2%	24.3%	26.5%	24.1%	28.1%
Category G	1.0%	1.3%	2.7%	3.5%	3.0%	3.8%
Category H	0.2%	1.1%	0.2%	1.0%	0.3%	1.0%

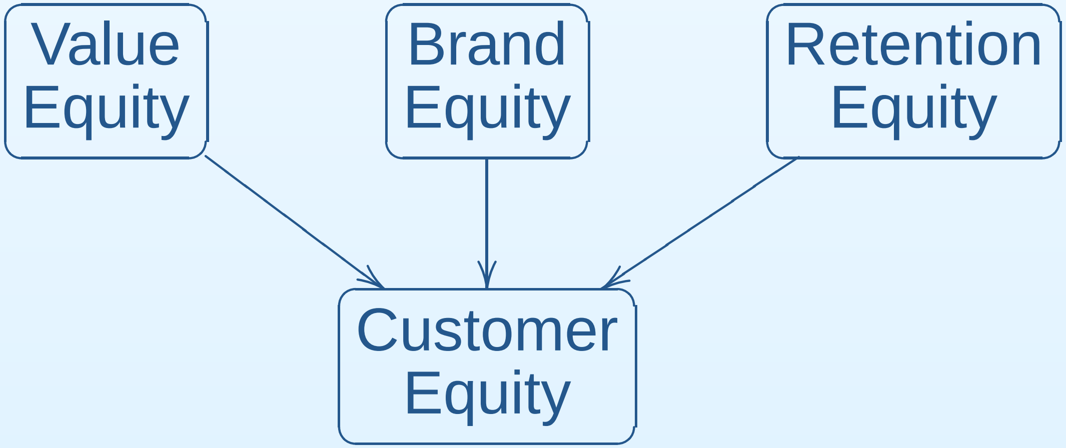
- Item Level Prediction Errors - Summary Statistics:

Mean	Std Dev	Min	Q1	Median	Q3	Max
0.27%	2.80%	-5.77%	-0.48%	0.25%	1.77%	5.28%

* *Small (n=589) US Consumer Study* ** *Actual from previous year's sales.*

Measuring Customer Equity/Lifetime Value From DCM

- Proposed by the book: *“Driving Customer Equity - How Customer Lifetime Value Is Reshaping Corporate Strategy"*, by Rust, Zeithaml, and Lemon (2000).
- 

- The drivers of customer equity may also be identified via structural models:
- 

Segmentation/Classification

- Segmentation divides a defined market into groups of customers/targets who are likely to exhibit similar behaviors related to purchasing behavior, loyalty, retention, etc.
- It’s purpose is to drive tailored approaches to these smaller groups (segments/clusters), via:
  - targeted marketing communication,
  - customized value propositions & offerings, and
  - customized delivery.
- The end objective of market segmentation is business performance driven - i.e., to improve win rates and/or, to reduce customer loss (“churn")

Segmentation/Classification - continued

- Market segmentation should also be operational ("implementable") in that it should enable target marketing/selling (or de-selection).
- First step is to be able to classify customers/prospects into defined segments.
- Further, the implications of the segments must be identified
  - How to communicate (targeted messaging, value proposition, etc.)
  - Find the best media for messaging
  - How to increase satisfaction through proper product/service delivery
  - Find the best channels for product delivery

Business Problem: “BI Silos”

- Most times, these models are used in “silos"
  - Database/CRM systems use mining models
  - Independently, marketing strategy teams use survey research
  - Lack of “knowledge transfer"
  - Resulting in many missed opportunities
- Such business intelligence “silos" also exist elsewhere in the business structure
  - Human resources management, especially of large corporations, could utilize data mining models together with employee survey research
  - Supply chain management could utilize data mining models leveraged against inventory databases

Viable General Approaches

- Survey sample stratification could include cluster distributions
  - Data mining modeler & survey sample designer need to communicate
- Capture/append DMM-relevant variables in survey research
  - This will enrich the analyses of the survey data
  - On the other hand, this may provide additional insights to DMM
- Incorporate intelligence from the other side into models as prior information/bayesian approach
  - e.g., Brand affinity or market sizes from survey research could be useful for improving the neural networks
  - e.g., Cluster patterns from data mining can aid in explaning/teasing out findings from survey research

Specific Approaches for DCM/Segmentation Models

- Segmentation $\rightarrowtail$ DCM:

  Conditional Logit Model:

  $$U_{ik} = V_{ik} + \epsilon_{ik} = \alpha' \mathbf{A}_k + \beta_k' \mathbf{X}_i + \epsilon_{ik}.$$

  where $\mathbf{A}_k$ denotes the k-th individual's characteristics. Some components of $\mathbf{A}_k$ could be from the segmentation scheme.

- DCM $\rightarrowtail$ Segmentation:
  - HB estimates β available for each individual. Averages of these estimates have been found to be very close to MSL estimates (Train, 2001).
  - Build cluster level models using such average β's. Incorporate these into segmentation strategies.

Integrated Business Performance Approach

- Such approaches can provide improved business performance.
- For churn/loss intervention approach, offer products/services that will help to retain them
  - Determined from both segmentation & choice models
  - Highly tuned to individual’s past behavior and indicated preference
- For a loyal customer, offer products/services that will improve their satisfaction as well as increase profit
  - Again, determined from a comprehensive view and highly tuned to the individual
  - Enables higher profit generation (e.g., leverage likelihood to spend more, customer looking for sophisticated products)

Choice Based Mining Models?

- Consider a Market Basket Analysis scenario (e.g., Super market shelf displays, online store webpages)
- Focusing on a specific category (e.g., laundry supplies), include product attributes relevant to product choices
  - Brand, Price, Weight, etc.
  - Adjacency to other categories
  - Distance from the store front/home page
- Other predictors may also be considered
  - Time series based: e.g., seasonality
  - Advertisement intensity
- Prior Information
  - Product popularity/market share
  - Individual specific purchase frequency
- Based on these, predict utility for the product and purchase likelihood

Discussion

Questions?