Pricing and Promotions Analysis at *Molson Coors Brewing Company*

1 The Problem

Marketers in most business-to-consumer markets share a common problem: how to assess the impact of pricing and promotion actions across multiple product lines in a competitive and volatile environment? For brewers, this problem is particularly challenging because customers can and will switch back and forth across brands and product lines in response to shifts in price, season, advertising, preferences, and other factors. As a consequence, decisions relating to one product can have significant impacts on other products, and it is not sufficient to consider each product in isolation. Here's a sample of the types of questions that arise:

1) If you drop the price on one item, what is the effect on sales all of your other products? Which products are competing inside your line-up?
2) Is a direct price drop more or less effective than other promotions, for example, bonus package size offerings or pack promotions?
3) What are the effects of the myriad other types of promotional and advertising actions -- aisle locations, cross-selling promotions, loyalty program bonuses, and so on?
4) How can effects be isolated when there are seasonal and other systematic shifts in taste?
5) Are their differences in responses across different channels or outlets within channels? Are there regional variations?
6) Many consumers will remain loyal to a brand or type of beverage; for example ales versus lagers, but migrate across package sizes or manufacturers. Is it possible to model this type of market structure?

Experienced marketers use a combination of intuition and analysis to feel their way through such questions, but there are always nagging questions about the relationship between decisions and outcomes. Swings in demand for products can result from a multitude of actions by companies and their competitors. How can you separate out the effects of particular actions? Could different decisions have produced better performance?

2 A Big Data Solution?

Managers at *Molson Coors Canada* (*MCC*), a global brewer, knew that they had an underutilized asset that might help them meet the decision-outcome measurement problem. Point-of-sale data capture and other technologies had been streaming detailed sales volume records data into 'data warehouses' for many years. In addition, contemporaneous data on promotions and other marketing initiatives were available and might be matched up with the sales data. In principle, it should be possible to analyze data like this and remove some of the guesswork, but the level of data detail and complexity is significant. Aside from the large number of brands under consideration, each brand could be sold in multiple different package and container formats. Large volume brands could have as many as fourteen different stock-keeping-units (SKU) varying across cans and bottles of different sizes in multiple pack-sizes. Promotions and pricing actions might apply to a particular SKU or might apply across an entire product line.

Typical data consisted of several years of weekly sales by outlet for most company products at the SKU level of detail plus weekly sales for a selection of important competitor SKUs. For example, one analysis involved weekly sales of 206 beers through 1,314 stores over 130 time periods spanning 2010 to 2012). The raw sales volume data contained over 35 million observations, and there were additional data for prices over all periods and promotions and other actions at selected times.

Experienced analysts will recognize this situation as calling for some sort of associative model – perhaps a multi-equation econometric model in which each equation links sales for a product to pricing and promotion decisions for all products. However, there is an immediate problem with this approach: if each equation contains five parameters for fitting time trends and seasonality, five parameters representing specific actions for the product and five more for each of the other products, the resulting model will be a multivariate multiple regression with hundreds of equations and thousands of parameters per equation. For example, a model for two hundred SKU's would require estimates for 200,000 parameters. Moreover, price responses are rarely linear – any effective model would need to reflect non-linearities in a natural way. Brand clustering and other factors would require constraints on groups of parameters that would further expand the size of the problem. Even with the impressive computational capacities available today, this is not a realistic estimation problem to attempt on a routine basis.

Successful resolution of the challenge required the product knowledge of *MCC*, the expertise in strategic and tactical pricing of *Pricing Solutions Limited*, Toronto, and the knowledge of mathematical modeling for pricing and revenue management of a team at Queen's University.

The key challenge was to devise a modeling approach that could reveal useful information while limiting the estimation problem to a much smaller number of parameters. Based on experience with other problems of this type, we knew that a practical limit was in the order of five hundred parameters – a massive reduction from the hundreds of thousands of parameters for a full model. Our solution was a three-level *Nested Multinomial Logit* (3-MNL) consumer choice model. These models aim to estimate the desirability of each of a set of choices through a *utility function* linked to prices, promotions, and other factors. These utilities are then used to estimate the probability of each choice for an 'average' consumer. The nest structure permits modeling some of the complex brand and product interactions. For example, the probability of choice of a cluster of products or brands is estimated and then the conditional probabilities of products within that cluster and of specific SKU's within the cluster/product nest. This approach also allows selective avoidance of the well-known 'independence of irrelevant alternatives' found with non-nested MNL models. In such models virtually identical products are treated independently when they should be treated as substitutes. In the 3-MNL model, items inside nests can be interpreted as partial substitutes, and the strength of the substitution effect can be gauged by a nest parameter analogous to a correlation.

Product choice probabilities can be directly interpreted as percentages of sales of the observed products (called 'incidence rates'). Regular fluctuations in incidence rates attributable to seasonal factors (e.g. beers popular in summer or over holidays) were handled with a low order trend model using trigonometric functions and adjustments for calendar effects. This also permitted use of the model as a forecasting tool since seasonally-adjusted incidence rates could be used to disaggregate projections of total demand into SKU-level demand estimates.

This model still required an equation for each product, but the much-needed reduction in the number of parameters is an inherent property of MNL models. Essentially, these models assume that effects of pricing and promotional actions are distributed pro-rata across competing nests, sub-nests, or products. For example, if you drop the price of a SKU and thereby increase demand, the reductions across other SKU's will be proportional to their market shares before the price drop. This approximation reduces the sensitivity of the model to the vagaries of real market shifts, but it is a good approximation in many cases, and the corresponding reduction in size of the model is dramatic.

MNL models have been used for many years in studies of consumer choices among several alternatives; for example, choices of transportation mode for commuters. The novel aspect of the project described here is the size of the model, even after the reduction in parameters. In fact, we were not able to find any commercially available statistical package capable of fitting an MNL model with as many as two hundred product choices. (One special-purpose package allowed a little over one hundred products, but we were warned of convergence problems in models of that size.) The central mathematical problem was maximization of a nonlinear, mixed-integer likelihood function over several hundred variables. The eventual analysis was accomplished with a custom-built combination of MySQL database extraction, C++ programming, and custom built linear and nonlinear programming algorithms. Final summarization of results was accomplished with a user model custom-built by *PSL* that allowed Revenue Managers at *MCC* to assess and plan different scenarios.

The first project was undertaken in Canadian markets, which differ across regions because of provincial variation in liquor laws and distribution channels. For example, beer can be sold in corner stores and groceries in the province of Quebec but only through government controlled beer or liquor stores in Ontario. Consequently, each region presented different modeling challenges. Success in Canada has meant that other Molson Coors country teams have begun the process of building similar models for their markets, which have additional variations.

As with most model-building exercises, this was not an automated process. The number of possible model structures grows exponentially in the number of parameters and equations and, with two hundred equations and several hundred parameters, the model design challenge was formidable. The product knowledge of the *MCC* and *PSL* teams was instrumental in narrowing down the range of model structures to consider, but there was much trial-and-error required. (This and similar problems have motivated a doctoral research project at Queen's seeking machine learning and other heuristic methods to aid in the development of large scale models.)

3 Results

There were many detailed results from the model that were of interest to *MCC.* We'll illustrate the types of results obtained with one model for sales of selected *MCC* and competing brands through stores operated by the Liquor Control Board of Ontario (LCBO). For reasons of confidentiality, we conceal the true identity of products under consideration.

1) **The Consumer Choice and Total Sales Model**: The 3-level nested multinomial logit model, when combined with a total sales prediction model, can estimate incidence rates and total sales with reasonable accuracy. Average absolute errors in estimation of incidence rates of about 0.14% can be anticipated. Prediction errors were tracked both inside the sample of time periods used to fit the model (the 'calibration period') and on a holdout sample of the final 14 weeks in the available data.

   The consumer choice model, combined with a model of total sales, allows estimating the sales at the level of individual products. The average error of the total sales model is under 6% both on the calibration data and the holdout data.

   The charts of actual and predicted incidence rates (see Figures 1 and 2 below) show that incidence rates and significant seasonal shifts were successfully reproduced by the model. It is possible to achieve almost perfect fits to any data with a sufficient number of parameters, but it is important to note that this model is condensing the information contained in 35 million observations to a few hundred parameters. Also, error checks on a holdout sample gave good evidence that the model was not 'over-fitting'.

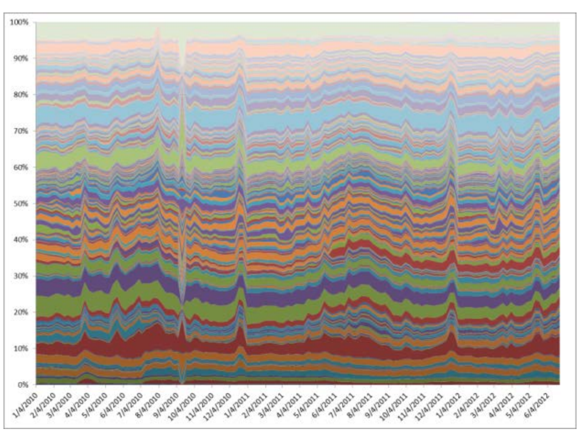

Figure 1: Actual Incidence Rates

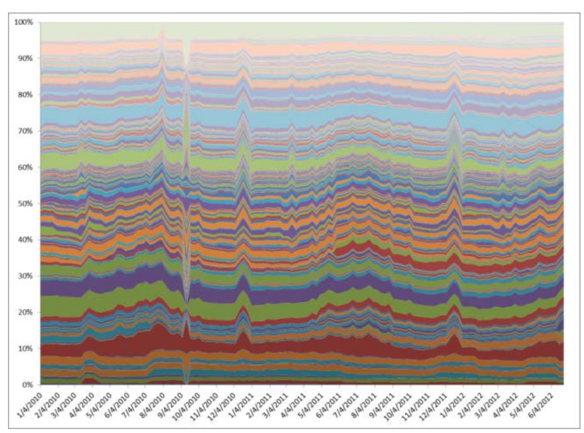

Figure 2: Predicted Incidence Rates

2) **The Nest structure** of the model reflects choice structures that were discovered by extensive experimentation with different model variants in consultation with product experts. Examples of findings were:
   a) Comprehensive experimentation with one strongly interacting nest revealed a robust and meaningful substitution pattern between certain European brands sold at LCBO.
   b) Several other European brands were found to interact with particular Canadian brands.
   c) Substitution patterns were identified among selected large domestic brands, whereas other large brands appeared to form independent nests.

3) **Price Sensitivities** obtained by the model represent the 'instantaneous' change in incidence rate corresponding to a one dollar increase in price per litre from current levels. These are not the same as conventional price elasticities but can be interpreted in much the same way when comparing products on similar scales. Sensitivities to promotions and other 'on/off' effects are interpreted similarly. The report excerpt in Figure 3, below, illustrates the kind of information obtained (brand, pack-size and promotion codes have been altered for reasons of confidentiality). For example, brand 100 in pack size 5, revealed a 0.741% drop in incidence rate (-0.741) per dollar price increase in price and a 0.769% increase in incidence rate corresponding to promotion type 1. The ability to rank order products according to price sensitivity and gauge the effects of promotions on specific products was viewed as particularly valuable.

		Sensitivity to			Sensitivity to promotions				
Brand Code	Pack Size Code	Price	#Stores	#Products	1	2	3	4	5
100	5	-0.741	0.000	0.000	0.769	0.000	0.000	0.000	0.0
117	11	0.000	1.301	0.000	0.705	0.000	1.618	0.000	0.0
159	23	-0.227	1.399	0.000	0.597	0.000	0.000	0.000	0.0
59	20	-0.370	0.000	0.441	0.461	0.455	0.000	2.264	0.0
178	17	0.000	0.000	0.000	0.296	0.000	0.000	2.310	0.0

Figure 3: Excerpt from Price Sensitivity Report

4) **Seasonality** was estimated at the level of brand types and included holiday affects as well as slow varying periodic patterns (with periods of one year and half-a-year). For example, it will be no secret that a particularly strong holiday effect was observed for Irish brands on St Patrick's Day, and seasonality patterns were quite strong for certain other brands but not observed for others.

4 Conclusion

The Price and Promotional Planning Model has been used to enhance the accuracy and rigor of the weekly, monthly and annual Revenue Management processes at Molson Coors Canada. The Revenue Management team had historically made decisions based on rules, norms and models built from years of experience. The new Price and Promotional Planning Model has enabled *MCC* to transition to a more data driven promotional management process.

Every week the business is confronted with decisions about price adjustments, promotions, incremental retail display, and other activities. Having a comprehensive model has enabled the team to consider and evaluate more scenarios surrounding each decision. The Model enables much more comprehensive analysis and consideration of complex issues such as promotion effectiveness, interaction with or cannibalization of other *MCC* brands, potential competitive reaction, or the effect of holidays and events and the different effect they have on the many different *MCC* brands and formats.

The annual planning process has benefitted from the use of the Model as well. Investment decisions can be made after considering and evaluating multiple alternatives. The baseline plan can be saved as a basis for comparison throughout the year as conditions change and actions are taken in the market. Competitive price changes can be quickly input in the model and impacts and responses can be diligently considered and updated and saved.

The legacy processes relied heavily on the skills of the team members. The new Price and Promotional Planning Model enables effective knowledge management and has enabled *MCC* to leverage the skill of the team, build on success, and standardize and document processes leading to more consistent results.

We feel that this approach is applicable for any consumer packaged goods for which reliable data on historical price, volume, and promotional activities are available. This highlights the need for business teams to maintain excellent records of promotional and other activities and offers a structured pathway for continuous improvement as the quality and variety of data improves over time.