Variable Selection for Health Insurance Market Segmentation

Introduction.

Our position on basis variable selection for health insurance market segmentation is solidly based on the marketing research professional and academic literature, coupled with our extensive experience in market segmentation for health insurance organizations. This work has covered just about every constituency served by health insurance organizations:

- Employers.
- Members—commercial group, individual, Medicare, and Medicaid.
- Network physicians.
- Brokers.

Although the orientation of the analysis differs somewhat across these constituencies, basic principles apply to all. Through experience we have found the right connections between conceptual and analytical approaches to derive segments and implementation strategies that clients have found useful for improving marketing efficiency and effectiveness. Our general approach to market segmentation is consistent with those used by other sophisticated marketing research practitioners with our own special twists learned through experience, interpretation of the results of dozens of studies and the special needs of the health insurance industry.

Market segmentation is based on differences in expected behavior.

Nearly 50 years ago, Wendell R. Smith defined market segmentation as the process of identifying customer subgroups (current or potential) that respond or are expected to respond differently to marketing variables (benefits, price, promotion, distribution, etc.). This definition has been accepted by contemporary marketing thought leaders such as Phillip Kotler.

The term market segmentation has also been closely linked with the term target marketing. In fact, the terms 'target marketing' and 'market segmentation' are often used interchangeably. However, there is an important difference in emphasis. Market segmentation involves dividing the market into subgroups based on expected response to different marketing offerings while target marketing is the process of identifying the most attractive market segments and designing marketing strategies to reach them. The concept of target marketing is a refinement of the basic philosophy of marketing. It is an attempt by companies to relate the characteristics or attributes of their products more closely to customer requirements and to communicate this to members of the targeted segments. Kotler has noted that the process of target marketing has three distinct stages:

- **Market segmentation.** The overall market is divided into distinct groups of buyers who are likely to respond favorably to different product/service offerings and market mixes.
- **Market targeting.** The process whereby one or more of the market segments previously identified are evaluated and selected.
- **Product positioning.** Competition is likely to exist for identified market segments in 'niche' positions. Product positioning is the process whereby the product or service and all the other marketing mix elements are designed to fit a given place within a particular segment. The position may be defined by communications such as advertising rather than actual product differences.
The process of market segmentation must be conducted with an eye toward the ability of an organization to address the resulting segments. There is a danger that unrestrained segmentation may lead to an unsustainable range of product modifications and produce similar impediments to market diversification. Criteria for assessing potential segments include:

- **Profitable size** - The relative profit potential in a segment is directly related to the competitive strength and cost effectiveness of the company. Even a small market may be profitable if the company has competitive pre-eminence.

- **Accessibility** - A segment must be accessible through advertising, other promotional media, and distributive networks.

- **Self containment** - Preferably a product launched at a market segment should not take demand from another product in the company's portfolio.

- **Marketing mix response** - The market segment should be responsive to marketing and promotion efforts.

**Process of market segmentation.**

1. Specify Basis and Descriptor variables
2. Use pattern recognition techniques
   - K-means cluster analysis
   - Latent class analysis
   - Two Step cluster analysis
3. Identify solutions for further testing
4. ANOVA to test solutions
5. Iterate (evaluate to “best” solution)
   - Multiple discriminant analysis (MDA) to develop model to predict segment membership based on Knowledge Base measures

---

*Use blocks of related items. Should be based on some theory or logic regarding things that can be expected to drive coverage choices. Variables should not be thrown into the statistical meat grinder willy nilly due to the potential for solutions that are nothing more than artifacts of what we put in.*

*Factor analysis may be used to reduce the number of variables in a particular block of related Basis variables. However, the process should involve using the variables most highly correlated with each factor to it. You DEFINITELY SHOULD NOT use factor scores in clustering or LCA. Factor scores tend to remove unique variance and make it difficult to find “good” solutions.*

*None of the techniques are optimization techniques. There are as many possible solutions for a particular dataset as there are observations in that dataset. We should start with the application of statistical criteria to make sure that the solution is “meaningful.” Beyond that, the solution must be submitted to a “gut check” in terms of its possible usefulness.*
**Identifying basis variables.** This is obviously a key decision in the market segmentation process and the key driver of the resulting segments. Market segmentation requires experimentation with different sets of basis variables with the ultimate goal of arriving at a set of basis variables that produce segments that are:

- Identifiable with externally observable data and, thus, actionable.
- Expected to be relatively stable over a time line of several years.
- Efficiently reachable.

Given that market segments involve subgroups of the total market that are expected to respond differently to different marketing offerings (product, price, promotion, distribution) it makes sense that the variables that we use as a basis for segmentation be variables that are predictors of behavior. To this end, thinking in terms of decision-makers for employer group coverage, we have found that the best basis variables are:

- **Attitudes.** Specifically, we are talking about attitudes toward employees, the role of employee benefits, and employers’ perceptions of specific types of benefits and cost control measures.

- **Benefit (product feature) preferences.** We measure preferences (the perceived value of different product features and levels of those features) through the use of conjoint / discrete choice analysis to force meaningful tradeoffs. Preferences are simply special types of attitudes in that they are learned and include knowledge (what you know about the product / brand), behavioral predispositions (buy it / don’t buy it, etc.) and affective (how you feel about it, commitment / lack of commitment, etc.) components.

- **Past behavior.** History of switching, current health plan type and carrier, current benefit design, monthly premium and cost sharing arrangements with employees, etc.

In terms of the basis variables noted above, we normally get no argument about attitudes or past behavior. However, we frequently get questions about our use of product preference data for market segmentation. Some note that products will change over time and when they change the segmentation scheme will no longer be valid. However, we actually use product attribute preference data in our segmentation schemes. We do not use the terms HMO, PPO, POS, indemnity, etc. Rather, we use a conjoint or decompositional approach where we measure sensitivity to features such as premium, deductible, copay and co-insurance levels, various types of network restriction, savings accounts, etc. Certainly, one can find combinations of these attributes that represent HMOs, various forms of PPOs, POS products, high deductible health plans with savings accounts, etc. These attributes can also be combined to represent products currently not available in the market place. As noted above, preferences are nothing more than special types of attitudes. If it makes sense to include attitudes then it makes equal sense to include preferences.

Utilities from conjoint analysis are simply interval scaled data (usually falling within a range of +4 to -4). Attitudinal ratings are ordinal data (or interval scaled if you believe the rating items are equally spaced). Since utilities and attitudinal ratings are both represented as numeric values, they can be combined in any statistical model. During data preparation, we standardized all values so that no one item or group of items can skew the clustering algorithm due to an arbitrary scaling effect.
Reasons for using preference data along with attitudes as basis variables

- **Attitudes/needs alone are not the best discriminators.** We have found that attitudes and needs-based ratings questions alone do not always differentiate people as well as they do when coupled with conjoint utilities as basis variables. Respondents typically do not vary by more than one or two ratings points on most attitudinal items. Respondents also may try to avoid looking “cheap” or “uncaring” by providing more socially acceptable answers to some questions. However, combining attitudes regarding health care with objective measures of benefit preferences (via conjoint analysis) has proven to be very effective in identifying segments that differ significantly on attitudes, needs and preferences. The conjoint utilities help us associate price sensitivity and preferences for certain benefits with the underlying attitudes and needs that drive those preferences.

- **Simplified conjoint exercise is not demanding.** For market segmentation, we use a simplified conjoint exercise with just 5 or 6 key attributes (e.g. premium, deductible, max-out-of-pocket, coinsurance, etc.). Often, no more than six choice sets are needed to provide a reasonable estimate of the relative value of the benefits being tested. Therefore, the conjoint exercise does not significantly increase the overall length of the survey. A full battery of attitudinal and needs-based questions is asked, along with firmagraphic, usage and descriptive questions, without lowering the participation and completion rates for the survey. We have successfully implemented this approach with as few as 2 or 3 choice sets in circumstances where survey length became an issue.

- **Conjoint results also provide marketing insights.** Besides aiding the segmentation process as basis variables, the conjoint utilities provide insights into each segment that allow marketers to target specific products or benefits messages to the appropriate employers.

Descriptor variables used to profile segments.

In addition to the basis variables used to derive segments, we use descriptor variables to profile and better understand segments. These variables include items that are often used for a priori segmentation, such as industry, number of employees and the like. Although not used as bases for segmentation, we typically find that segments derived as described above usually differ in terms of descriptor variables as well. For example, one segment may primarily be made up of larger firms; another is made up heavily of small firms, etc. This, of course, reflects the fact that there may be a correlation between firm size, attitudes and preferences. These correlations provide the basis for the predictive modeling discussed below.
Predictive modeling provides basis for operationalizing segmentation results.

Once a segmentation result has been finalized and accepted, it is necessary to find ways to put it to work, e.g. classify people in databases into different segments for mailings (different messages to members of different segments), provide a tool for brokers and sales reps that let’s them classify prospects into different segments and tailor their message accordingly, etc. (see graphic below).

The problem is that the above tasks must be completed without the detailed attitudinal and preference data used to develop the segmentation solution. Utilities, for example, are only captured through survey research. They cannot be observed or directly captured by a short classification tool. We sometimes use stated importance ratings as surrogates for conjoint utilities in the predictive model. However, it is difficult to use anything remotely similar to a conjoint exercise (such as a simple choice set or rating of a product profile) as a substitute for conjoint utilities because it takes too much time and energy to get feedback from the people you need to classify.

Conjoint utilities measure personal preferences for the benefits being tested and those personal preferences can cut across all demographic groups. In consumer products and other industries, preferences may directly relate to demographic or psychographic characteristics, but we don't find this pattern in health insurance. In the health insurance industry, there is some correlation between demographics / firmagraphics and utilities (e.g. low income people might be more price sensitive), but these correlations are usually small. When we place attitudes into the mix, we find a synergy between attitudes and utilities that produce more recognizable segments and predict segment membership with acceptable precision. The predictive model is developed using the survey data in a process where we use observable data (e.g. firm size) and or a limited set of attitude questions, depending on the intended use of the predictive model, to predict segment membership.