



Multiple Comparisons

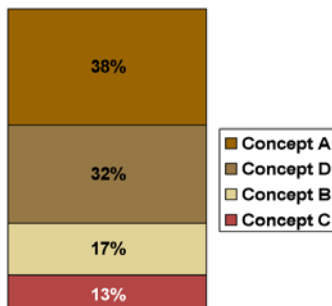
Many of our projects involve the evaluation of multiple positioning concepts or alternative versions of a spot ad. It is often impractical for one person to evaluate (or compare) more than two of them.

Nevertheless the client needs to know how all of them stack up against each other.

Fortunately, we've long had a tool in our toolbox that permits just that. It's been around for 40 years, so don't let other suppliers pass it off as the latest & greatest. The procedure is generally known as the Bradley-Terry-Luce model after the psychologists who created it.

The starting point is counts of the number of times each member of a pair was preferred to the other, plus the ties. An iterative procedure then creates a model of what the (first) preferences would have been if all had been seen and rated at once. How well can the model in turn reproduce the observed paired preferences? Usually very well.

With even four alternatives, you would have to report six different sets of two-at-a-time preferences. Do busy decision makers want to absorb all that and make a judgment call? Usually not. But a successful BTL model (not to be confused with our favorite sandwich) could state quite succinctly that the relative preferences overall are:



If Data Can Be Manipulated. . .

it will be. The British Royal Mail conducts mystery shopping experiments to test the accuracy and speed of delivery. Top managers earn bonuses for good scores. And they were able to find out the identity of the supposedly anonymous testers and provide them with exceptional service (so reports *The Economist* magazine). It was their bad luck to get busted.

As we know, customer satisfaction research is a huge industry, and (too) many firms use the results to reward and punish staff. The "rational" course of action is to try to manipulate surveys to look good. It is incumbent on any compensation-related research to include every conceivable safeguard against such manipulation. How many of us have had a car salesperson get down on his/her knees to beg for "all 5s" on the inevitable follow-up survey? It happens.

Significance

Anyone using quantitative research needs to understand inferential statistics. Alas, these are difficult lessons to learn and to remember. This brief note hardly substitutes for a focused seminar or university class, but we will mention a few pointers.

In lay terms, a declaration of statistical significance means only that there is a real difference between an observed statistic (such as a percentage) and either another observed statistic or an hypothesized value. The "real" difference could be teeny-tiny. It could be of no managerial relevance. But we can't chalk it up to noise alone.

(More Significance)

How much of a difference is important or managerially relevant is not a statistical question. A good quant researcher will have some perspective, but don't ask us to make a call on the basis of raw data alone. There should be a larger business context of risks, expectations, and hypotheses that have to be evaluated alongside the pure statistics.

"Significance" is not an attribute of a whole study. Significant differences can be found in very small samples, and can be elusive in very large samples. But please don't ask if a whole study is significant or not, or what sample size is needed to get there.

You should shudder—as we do—when looking at someone else's crosstabs where there are dozens of Tests of Significance (TOS) run on every page. By golly, you'll see a couple differences or so on every page that are significant at the .05 level. Such wholesale mining for significance is asking for trouble.

Prediction Markets

From time to time we check in with InTrade to see what the betting line is, so to speak, on interesting matters. Remember that in prediction markets, participants stand to profit if their predictions are correct (unlike media pundits). The "consensus" as we write is that:

- * the DJIA will close around 10,250 at the end of November.
- * the average national price for regular unleaded gas will be \$2.50 per gallon or more at the end of the year.
- * the odds of Ford filing bankruptcy by year-end are barely 1%.
- * the odds of Twitter announcing an IPO by year-end are around 2%.

Speaking of Pundits

Oft-cited research by UC-Berkeley Professor Philip Tetlock in his book "Expert Political Judgment" found a miserable record of accuracy among political pundits. One generalization is that accuracy has a negative relationship with celebrity!

Satisfaction Matters

We have never been shy about critiquing the satisfaction research industry (see previous page), but have no quarrel with the idea of treating customers decently. A potentially far-reaching study published in the October *Journal of Marketing Research* by Anderson and Mansi builds the business case for doing so.

Using an independent database of customer satisfaction with 150 large firms, the authors establish that satisfaction correlates negatively with the cost of financing through debt offerings. That's right. Taking a variety of other factors into account, firms with superior customer satisfaction paid less in the bond market.

This effect falls outside the normal boundaries of marketing ROI, but a firm's financing costs are often significant.

The authors concede that the mechanism by which satisfaction translates into higher credit ratings is only a matter of speculation. Some literature argues that high levels of satisfaction lead to more stable cash flow—which investors value.

It would be a challenge to prove such a connection for a specific firm (like yours), but the likelihood that marketing ROI extends to financing is an important lesson.

