

Improve Use of Quantitative Methods

*By Jim Wheaton and Cynthia Baughan Wheaton
Principals, Wheaton Group*

Original Version of an article that appeared in the August 26, 2002 issue of DM News

It is remarkable how many direct marketing companies have failed to fully embrace sophisticated quantitative techniques to better target prospect and customer promotions. Several years ago, a large study found that over fifty percent of catalogers still use RFM Cells rather than statistics-based predictive models. Perhaps this is understandable for “mom and pop” organizations. Interestingly, however, this percentage was only marginally lower for large catalogers.

Unfortunately, direct marketers also display a corresponding lack of sophistication in setting up test designs. Many organizations are unschooled in the basics of experimental design. Some companies, for example, still believe that fifty responders is always an adequate amount to make test panel comparisons. Invariably, they are unpleasantly surprised when informed about the degree of reliability to which this translates.

Even worse, some direct marketers run test panels of 5,000 regardless of the anticipated response rate, even those whose typical prospecting rates are just a fraction of a percent. Often, these companies make important decisions based on fifteen to thirty responders.

So, the broader question is why there is such resistance by direct marketers to embrace quantitative approaches when making prospect and customer circulation decisions. There are, we believe, several basic reasons:

“Seat-of-the-Pants Direct Marketers”

Many direct marketing operations are run by seat-of-the-pants circulation professionals, and a number by entrepreneurial founders or founder-families. These companies have found a viable niche, and have successfully grown and even thrived using homegrown approaches. Within such organizations, methods of operation tend to become rigidly codified over the years. Often, this is even true for companies that have evolved into large entities.

One cataloger, for example, grew from a niche operation to a \$250 million retail/direct hybrid. However, throughout this explosive increase in scale, the organization clung to its traditional, homegrown, cell-based method of selecting names for mailings. This was a convoluted process that required several days of effort on the part of the circulation manager for every promotion, and at least an equal amount by its service bureau.

Interestingly, a back-end analysis by a consultant indicated that, in reality, this cataloger's complex selection strategy generally boiled down to very simple "de facto" criteria. House multi-buyers, for example, would be promoted if one of the following conditions were met:

- Bought from any one of several core titles within the past 36 months; that is, if a purchase was made from any of the books, then the customer would receive all of the books.
- Had a company credit card.
- Hit against an outside rental list.
- Had an Abacus score within a certain range.

Despite this clarification, the cataloger clung to its convoluted selection strategy. The thinking was that the traditional method had been successful, so why risk trying something else.

Data Miners Who Alienate

To exacerbate matters, many direct marketers have had terrible experiences with data miners who have failed to deliver on their promises. As a result, these companies have become extremely resistant to sophisticated analytical techniques. Frequently, for example, we hear that, "We tried regression modeling and it didn't work all that well." Or, "We built a model and it looked good for awhile, but then it fell apart."

The problem generally lies with the types of individuals who are building the models. Our industry is filled with analysts who understand the mechanics of building regression-based predictive models. However, a much smaller number have the experience and ability to approach projects as insightful data detectives and savvy business people.

Unfortunately, building a potent model that will stand up over time generally requires acute data insight and an understanding of the mechanics of direct marketing. Consider, for example, that the following key decisions in any predictive modeling project require "in-the-trenches direct marketing" experience rather than an advanced degree in statistics:

- What mailings/drops should comprise the analysis file?
- What ratio of responders versus non-responders should be used, and should it vary by mailing/drop?
- Should a single or multiple-model strategy be employed?
- If a multiple-model strategy is used, how should their identities be determined?
- What is/are the appropriate dependent variable(s); that is, response, gross sales, net sales, gross margin, or what?
- How should missing values be handled, especially with continuous potential predictors?
- Should outliers be eliminated or capped?

Unfortunately, most companies underestimate the importance of “in-the-trenches” direct marketing knowledge when interviewing potential data miners. Hence, they tend to hire pedigreed, advanced-degree statisticians who then proceed to make every rookie mistake in the book.

Ironically, many of the best direct marketing analysts do not have an advanced degree in statistics. Instead, they are “data detectives” with years of direct marketing experience. Of course, seasoned data detectives with advanced statistics degrees are the ideal. Unfortunately, they are very rare.

Statistically-Unschooled List Brokers

Resistance to quantitative techniques is also seen in the many list professionals who lack statistical sophistication. Making test recommendations is a core function of the list brokerage business. Nevertheless, many practitioners have never had a statistics course.

Recently, a veteran list broker recommended test quantities of as few as 5,000 to a direct marketer with prospecting response rates as low as 0.25%. Unfortunately, the resulting thirteen responders on a list with a 0.25% response rate would be far from adequate to confidently read the results.

When confronted with a basic statistical formula for calculating appropriate test quantities, the broker refused to believe in its validity. She was skeptical when informed that, except for niche lists with very low rollout quantities, universe size generally has only a secondary impact in determining the appropriate test quantity.

For example, assume an expected response rate of 1%, and that we want to be 80% confident that the actual response rate will be at least 0.9%. With a rollout universe of 50 million, we need a test quantity of 6,984. With a universe of 50,000, we require a quantity of 6,129 – just 855 fewer names. Only when we get down to small universe quantities are the required test quantities markedly smaller. A universe of 10,000, for example, requires a test of 4,113.

One of the list broker’s final comments about the basic formula for calculating appropriate test quantities summed up the lack of statistical sophistication that is all too common in our industry, “If this formula is valid, how come I’ve never run into it before in my 20 years in the business?”

Final Thoughts

The bottom line is that there is a major opportunity in the direct marketing industry for improved analytical sophistication of underlying circulation strategies. And, it is the wise direct marketer who will take advantage of this fact.

Jim Wheaton and Cynthia Baughan Wheaton are Principals at Wheaton Group, and can be reached at 919-969-8859 or jim.wheaton@wheatongroup.com. The firm specializes in direct marketing consulting and data mining, data quality assessment and assurance, and the delivery of cost-effective data warehouses and marts. Jim is also a Co-Founder of Data University www.datauniversity.org.