

Improved Data Will Produce Better Models

*By Jim Wheaton
Principal, Wheaton Group*

Original Version of an article that appeared in the October 14, 1996 issue of "DM News"

How many angels can dance on the head of a pin? This is a question that inspired a furious theological debate many centuries ago.

We database marketers are fixated on our own pins and angels with our constant wrangling about modeling techniques. A by-product of this debate is the fantastic claims made by proponents of these techniques.

My favorite is a promotional piece that I received this summer from a neural network software company. The brochure began by noting that conventionally selected rental lists generate a 1% to 2% response, and it ended with the case study of a client that pulled 63% via the company's neural net product.

Unfortunately, such claims are pabulum for the gullible. I am not questioning the veracity of the 63% response rate. I am confident, however, that this extreme result was due to a unique combination of list, product and offer – and not to the neural net product itself. There is only so much variance in the data, and the unfortunate truth is that new modeling techniques are not going to drastically improve segmentation power.

While we focus on our pins and angels, however, a quiet revolution is taking place that centers around improved data rather than modeling techniques. Smart target marketers are realizing that the real key to more powerful models is increased and more accurate information – the input to the modeling process. In other words, to fracture a quote from Bill Clinton, "It's the data stupid!"

One important development is the co-op databases of catalog purchase history. Abacus, the granddaddy of them all, has succeeded in combining customer transaction history from over 600 catalogers into one massive database. Other players, such as Smartbase and Z-24, are formidable competitors.

A senior catalog executive recently relayed to me the success he has had with one of these co-op databases. His co-op of choice consistently has identified profitable groups of purchasers from customer segments that his own models have classified as poor performing.

The reason for this success is improved data rather than revolutionary modeling techniques. The co-op is able to identify individuals who are buying from other catalogs – although not recently from our senior executive's – and are therefore likely to be responsive.

The power of co-op databases is that they contain what I call “money where your mouth is” lifestyle information. Based on actual catalog purchase behavior, individuals identify their life-stages – parents of newborns, bicycling enthusiasts, and the like. Such information is critical to targeted promotions.

A limitation of these co-op databases is that they only represent the catalog-buying portion of U.S. households, and do not contain all of the catalog purchases for even this minority. The challenge is to extend the coverage in terms of both names and transactions by finding additional sources of “money where your mouth is” lifestyle information.

One candidate is retail databases. We have a client, for example, that maintains detailed purchase information on eleven million customers. This client tracks item-specific transactions, totaling several billion dollars, on about 45 diverse merchandise groups such as jewelry, sporting goods and electronics. Imagine the targeting possibilities of a co-op database comprised of four or five large, non-competing retailers, totaling – say – 70 million households.

Besides driving retail-specific promotions, this data would allow the catalog co-ops to close their purchase gaps. The catalog data, of course, would do the same for the retailers. Throw in some clever marketing and creative, and the result would be innovative combined offers; for example, purchase Product A from Retailer X and get 50% off an order from complementary Catalog Y.

Another promising source is magazine subscriber files. The hundreds of niche titles that populate newsstands throughout the country are excellent vehicles for lifestyle information.

With complete lifestyle profiles gathered from multiple industries, the targeting possibilities would be endless. The trick, of course, is combining these disparate data sources in such a way that it is advantageous to all participants. This is a challenging yet not impossible task. After all, who would have thought five or ten years ago that hundreds of catalogers would willingly and profitably share data?

To summarize, the success of catalog co-op databases is evidence that today’s key to dramatically enhanced target marketing lies more with improving the data – the input to the modeling process – rather than with the modeling process itself. “Money where your mouth is” lifestyle information, compiled from multiple sources of purchase information, would result in astonishingly accurate promotions.

Jim Wheaton is a Principal 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.