THE EFFECTIVE INTEGRATION OF MARKET RESEARCH AND DATA DRIVEN MARKETING

This paper outlines how market research and data driven marketing can effectively be brought together. Both are essential tools to meet the challenges that marketers face when trying to understand their market and to service their customers. The pros and cons of both methods are outlined followed by an approach to data fusion. Several methods are available to integrate market research and data driven marketing while retaining the fundamental, essential anonymity of individuals. These methods are illustrated through examples within businesses.

THE CHALLENGE FACING MARKETERS

The challenge facing marketers is how to gather and utilise various sources of information to understand their markets and customers. The marketer then needs to provide products and services that meet the needs and expectations of their target markets. Marketers ideally want to pursue long-term sustainable relationships with their customers, rather than simply driving short-term acquisition initiatives. Doing this requires a deeper understanding of both the customers' behaviours patterns and their attitudes.

Market research and data driven marketing are both disciplines that support the marketer in pursuit of these goals. Market research will typically allow the marketer to understand more about their customers' motivations, feelings and attitudes. Conversely, data driven marketing will expose customers' actual behaviour patterns and characteristics.

The challenge then becomes how the two can be effectively integrated. This challenge is faced by organisations, some of which have taken the step to integrate the disciplines into a single department. This is often under the leadership of a market researcher (perhaps deemed to be 'more strategic' than their data driven marketing colleagues). Although this brings the disciplines together it does not necessarily lead to effective integration. Although data fusion is not a new technique it has been used across many different industries, the adoption of data fusion for market research and data driven marketing has been somewhat limited. [i]

This paper describes the two disciplines, their strengths, drawbacks and potential when used together; potential that is often left untapped by organisations.

THE PROS AND CONS OF MARKET RESEARCH AND DATA DRIVEN MARKETING

The advantages of market research

Market research provides information that is very difficult for companies to collect in an effective manner through normal interactions with their customers. It provides insight into reasons for customers acting in certain ways, and perceptions they may have about a company or its products and services. This may be useful during product and service development, or prior to new product launches by ascertaining attitudes to various scenarios, or directly testing response to various offers. Information can be gathered from customers on their usage of products and services from other providers or competitors, and the share of wallet between these respective suppliers.

Behavioural information, such as media usage, can also be gathered. This can be invaluable to guide the marketer in deciding on effective advertising to distinct market segments. Information can also be gathered reflecting how customers think they will behave in the future. Companies can also learn about people who are not currently their customers, including prospects, and customers who have recently defected.

A useful approach is to ask, 'Can I obtain this information from anywhere else, before I go and ask people through traditional market research methods?' The availability of market data varies from country to country. This paper does not cover the sources available of publicly and privately available information.

THE DRAWBACKS OF MARKET RESEARCH

Market research does have some disadvantages, namely:

- Sample size
- Sampling approach
- Customer response accuracy
- Customer response reliability
- Timeliness of results
- Implementation of findings

Sample size

Typically, only a small sample of customers is targeted. Furthermore, specific customer responses are not provided back to companies to enable them to identify who responded in certain ways. For example, a company may learn from market research that 20% of customers intend to behave in a certain way. However, they do not know which 20% of their customer base will behave in that way.

Sampling approach

Carrying out a sampling approach can lead to bias that is difficult to quantify let alone rectify. For example, if the research carries an incentive the type of incentive will introduce a bias within the response base. As more quantitative research is now carried out on-line, this brings with it its own biases.

Customer response accuracy

The accuracy of customer responses may also be questionable. A customer's perception of their own behaviour or activity may not actually represent reality. For example, customers may inaccurately estimate their own level of spend or frequency of usage. If such results are grouped together with responses that are more 'accurate' it is impossible to tell which ones were exaggerated.

Customer response reliability

A customer's indication of future behaviour may not in fact eventuate. For example, after recent electricity reforms, the number of customers indicating their intention to switch providers was far higher than reality.

Responses may also be skewed by recent events. For example, a customer experiencing recent poor service from a company may be more likely to have a negative impression than if poor service event had not occurred immediately prior to the research.

Good market research is screened and cleaned for obvious inaccuracies such as, every answer having the same response and cross-validation across questions.

Timelines of results

Due to the typical time taken to carry out market research and compile results, findings may be out of date by the time they are re-presented back to the client. Of course, this will not necessarily be an issue

for all market research, as some topics retain their currency for longer than others. As digital continues to grow as a channel for marketers, this is rapidly reducing as an issue for market researchers.

Implementation of findings

As marketers have no way of knowing who the customers are that hold the range of attitudes and perceptions researched, it reduces the effectiveness of the research. Ideally the marketer wants to translate the research into actionable strategies, among these, carrying out relevant targeted communication or actions.

Whilst the emergence of on-line communities and social media monitoring is countering many of the issues encountered with traditional research methods, the need for integration remains.

THE ADVANTAGES OF DATA DRIVEN MARKETING

Data driven marketing accesses information about individuals' transactional, behavioural, and profile characteristics. These data are usually the most powerful and reliable for predicting future behaviour. This information can be used to discover patterns in behaviour that can then be used in a predictive manner for customers who share similar characteristics.

This information is typically available for all customers, and can be readily augmented with data from other external sources such as Census data. The fact that the data is available for all customers, not just a sample, essentially removes the problem of sample bias.

Data derived from data driven marketing is often more accurate for measuring customer behaviour. This is particularly the case for transactional data where precise usage can be measured.

Data driven marketing is usually cost effective, with the cost of data collection typically being minimal. This data is usually required for other business operations, and so may already be available.

THE CHALLENGES OF DATA DRIVEN MARKETING

Data driven marketing also has some disadvantages, namely:

- Lack of information on attitudes and perceptions
- Data held in different systems and difficult to extract
- Data completeness and accuracy

Lack of information

The main limitation of data driven marketing is the lack of information pertaining to customers' attitudes and perceptions.

Data useful for data driven marketing is usually only available for people who are current or past customers. This makes data driven marketing of less use for prospective customers.

Data held in different systems and difficult to extract

The way institutional data is stored by a company often makes it difficult to access or extract. Data may also be spread across multiple disparate systems, and linking all the information together and associating it with a single customer can prove highly challenging.

Data completeness and accuracy

The data itself is frequently inaccurate or incomplete. This can be particularly the case in large organisations with many people entering data, for example a large call centre. Even with the best training and the latest software, information may be missing, entered in the wrong fields, mis-typed, transposed etc. This can lead to substantial effort in cleaning the data prior to any analysis being carried out. This may prove impractical or impossible, leading to records being discarded, or being allocated default values.

There are a number of approaches for dealing with missing data. Data may be available from external data sources, for example, Census data for estimating household income where this data is missing for an individual. Another approach is to estimate values from aggregated data. For example, if postcode is unavailable and this provides a link to other variables, then use the town which links to data that is an aggregation of a number of postcodes.

INTEGRATING THE DISCIPLINES

Benefits

The benefits of using market research and data driven marketing together are far greater than either in isolation.

The ideal scenario is to match individual customer responses from market research to their actual behaviour gained from data mining. A link can be made between customers' attitudes, views and thoughts and how these translate into observed behaviours. This can provide insights, not only to that customer's behaviour, but also that of other customers sharing similar characteristics.

Integration can help overcome the weaknesses of each method alone, for example, market research combined with data driven marketing could identify patterns that enable a company to identify likely prospects, or predict customer defection.

Approaches to Integration

The following section describes the number of scenarios where market research and data driven marketing are used together to enhance, validate or drive the other discipline. This is followed by a case study to illustrate how consumer anonymity can be preserved.

ENHANCEMENT

Market research data can be used to enhance data driven marketing. This could include information such as radio preferences, magazines read, favourite TV programmes etc that will aid the marketer in identifying and targeting customers effectively.

Typically, data driven marketing will identify unique groups of customers. Common groups surveyed include high value customers, new customers, inactive customers. Samples of customers can then be selected appropriately from these groups, and details provided to a market research company. These customers can be surveyed, responses gathered, and the results returned. The responses are then used to categorise members in each of the customer groups.

This approach is potentially more useful to marketers than either method alone as it includes a combination of data from both their own database and market research. This combination provides insights into the actions resulting from different customer attitudes.

CASE STUDY ONE – AN INSURANCE COMPANY SEGMENTATION

A segmentation of the insurer's customers demonstrates how data driven marketing can be enhanced with market research information.

The segmentation was initially developed using customer specific data matched with Census and Quotable Value NZ (property ownership) data. The data driven marketing company then provided a sample of customers to the market researcher. This sample contained name, address and segment number (1-5) information.

Using its Financial Services Monitor, the researcher provided additional information about customers within each segment including share of wallet, activities, attitudes and feelings.

To further assist the insurer in understanding customers within each segment, the data driven marketing company carried out video interviews of typical customers within each segment.

VALIDATION

Market research techniques can be used to validate or confirm that data driven marketing techniques have correctly classified unique groups of customers.

This method is similar to that applied above. The difference in this case is that the market research is used to validate the results derived from the data driven marketing. Market research would be undertaken on a sample of customers from each of the customer groups, testing that some of the key features identified by the data driven marketing hold true.

One of the disadvantages of market research previously mentioned was the potential inaccuracy of customer responses. Data driven marketing techniques can be used to validate such responses.

For example, market research was undertaken to ascertain customers' spending patterns with a large retail organisation. This same data was then validated using actual transaction data for the customers surveyed. The results indicated a considerable discrepancy between the two methods, suggesting that customers may not always be the best judge of their own behaviour (Figure 1).

Purchase Size Distribution – Observed Behaviour

Purchase Size Distribution – Market Research Answers
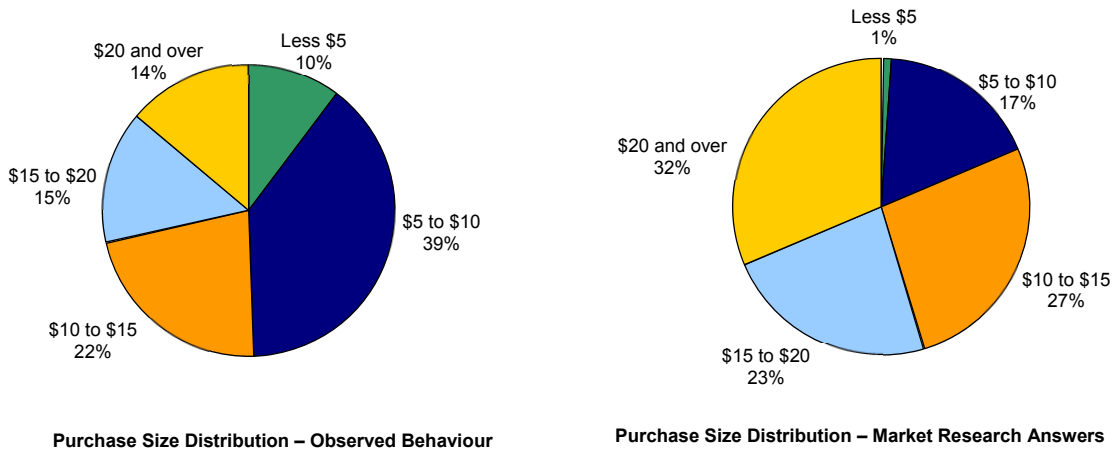
Figure 1: Comparison of market research responses to actual transaction data to determine customer spending patterns at a large retail organisation.

Analyses such as this can be used to decide whether to consider responses to other questions for those respondents whose behaviour was shown to be different from their market research response. This may help eliminate potential bias from respondents who tend to exaggerate. This approach may also be helpful in determining what information is best derived from market research, and which from data driven marketing.

Using one discipline to drive the other

Market research can be used to initially formulate distinct customer groups, and these decisions supported by data driven marketing techniques. For example, market research can be used to ascertain customers' commitment or loyalty towards a company, and these responses used to initially categorise groups of customers. A list of customers and the group they belong to could then be used to carry out further data driven marketing. Transactional or behavioural data then adds to the initial market research derived customer groups.

Market research is carried out on only a sample of the entire customer base. If the base is sampled randomly, 'average' customers may comprise a large portion of the customer base, and so the results will strongly represent these customers.

A better approach is to use data driven marketing to identify groups of customers illustrating behaviour that is of interest. Samples of customers falling into these interest groups can be extracted and provided to a market research organisation. This technique ensures those interest groups of customers, rather than 'average' customers are over-sampled (Figure 2).
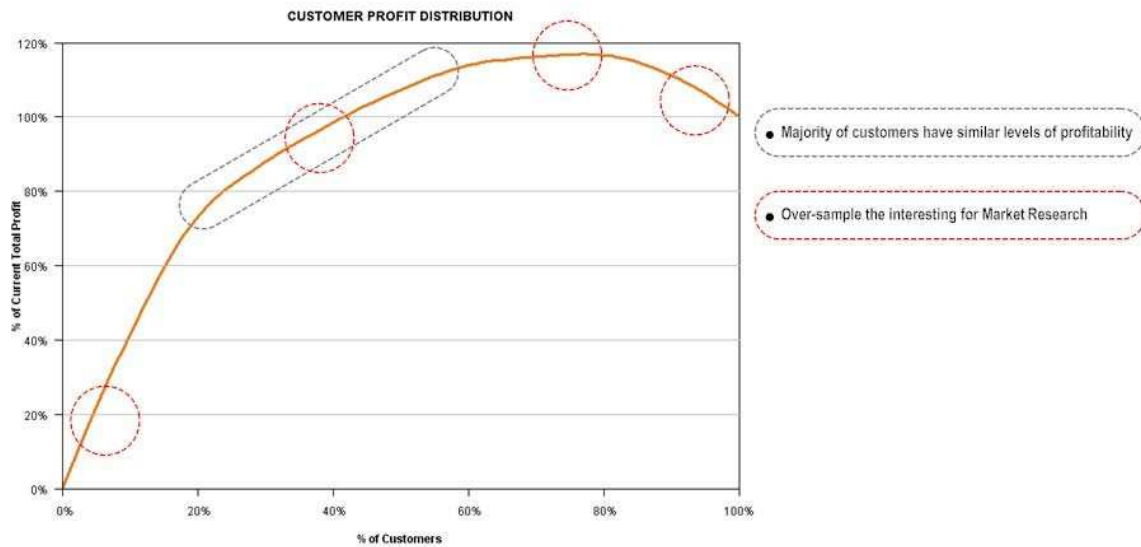

Figure 2: illustration if over sampling customers with unusual levels of profitability to ensure Market Research effectively profiles the entire base.

INTRODUCING DATA FUSION

Direct linkages can be found between the market research and database in order to fuse data into a single source. For example, both the research and database have age information. Although this is a form of data fusion, practitioners will get mixed results by just adopting this approach. In doing this direct linkage, the assumption is made that the variables chosen as the link variables, are good predictor variables. In the example, age may be a good linkage variable for a respondent saying they are interested in a new product, for example, life insurance, but would likely be poor predictor for an attitudinal response such as Net Promoter Score (NPS).[ii]

Thus, the fusion process needs to go through the following steps to ensure that the link created is optimal.

- Formulate a plan for how each of the research variable(s) are going to link through to the database
- Carry out research with good direct variables (held in the database) and indirect linkage variables (variables that can be included that are available on other data sets e.g., home ownership) included in the questions
- Merge the research, customer, product and transactional data together
- Model the research variable(s) against the data based variables. When there are multiple research variables, then the weights would need to be adjusted
- Check the models for validity
- Check the weighting of the variables
- Check the sample weightings and remove bias via post stratification
- Export the models into database language. For example, convert the statistical model into SQL

The following formula illustrates the statistical approach for the fusion of the research and data based information:

- Where is the vector (or matrix in the case of multiple research questions) of dependent research responses
- Is the matrix of independent direct and indirect data based variables
- Are coefficients for the independent variables in the model and are the errors within the model

This is used for simple data fusing; where there is an amalgamation of multiple research variables, simulated annealing is required, which is another topic in itself.

Watch out for over-fitting

Due to the large number of potential independent variables and few observations in the research variables, over fitting of the data is not just a potential issue, but one that will occur if not kept in check. This is due to the fact that the number of observations in the research may be small, for example 2,000 responses, and yet the number of variables that could be created from the data based information is immense (sometimes over 10,000 variables where rich transactional data is available in industries such as FMCG, telecoms, banking). Thus, building in over fitting technologies, such as 10-fold cross validation or bootstrapping, is required throughout the process.

The application of post weighting

Research undertaken may have been conducted with a well balanced quota sampling approach, such as age and gender. However, age and gender may not be related to the bias response. For example, people that dislike their service provider are less likely to respond (which is almost always the case). The service organization is collecting NPS via research with a well balanced age/gender quota.

Once the research data is linked to the database and applied across it, it can be seen that customers with low NPS's have definite differentiation in data based variables, for example the number of transactions per month. The sample population needs to be re-weighted to better reflect the bias in the non-response and hence ensure that the linkages that are created via the model are robust. The flow on from this is that different sampling frames will be created in the future, leading to the creation of more accurate assessments of the research variables.

It is important not to overlook the importance of sparse data for inclusion in the fusing model; sparse data can be valuable. For example, there are few respondents that had been turned down for a credit card. This could be a significant event that reduces the NPS. It would be advantageous to over-sample people that had been turned down for a credit card rather than leave this variable out of the model as the data was sparse.

THE ANSWER LIES IN INTEGRATION

Market research and data driven marketing can be used together to build understanding of a company's market and customers. Using this approach, neither type of information drives the solution more than the other. Most importantly, customer specific behavioural and transaction data should be combined with customer specific market research data within the constraints of the country's Market Research Code of Practice and Privacy Acts. Whilst the detail of these laws may vary by country, they are fundamentally there to protect the privacy of individuals and the principles apply across geographic boundaries. This process is illustrated in the following case study.

CASE STUDY TWO – PROTECTING CUSTOMER PRIVACY

This is an example of how data driven marketing and market research data can be used jointly to generate customer insight.

The key to carrying out this form of integration is to ensure that it is not possible to identify any individual using personal information. However, it is still necessary to know an individual's responses in order to link it to other customer specific data.

To achieve this, the company created a new unique ID for each customer. This new ID was not related to the old ID number, and the company kept the key to translating one to the other (Figure 3).
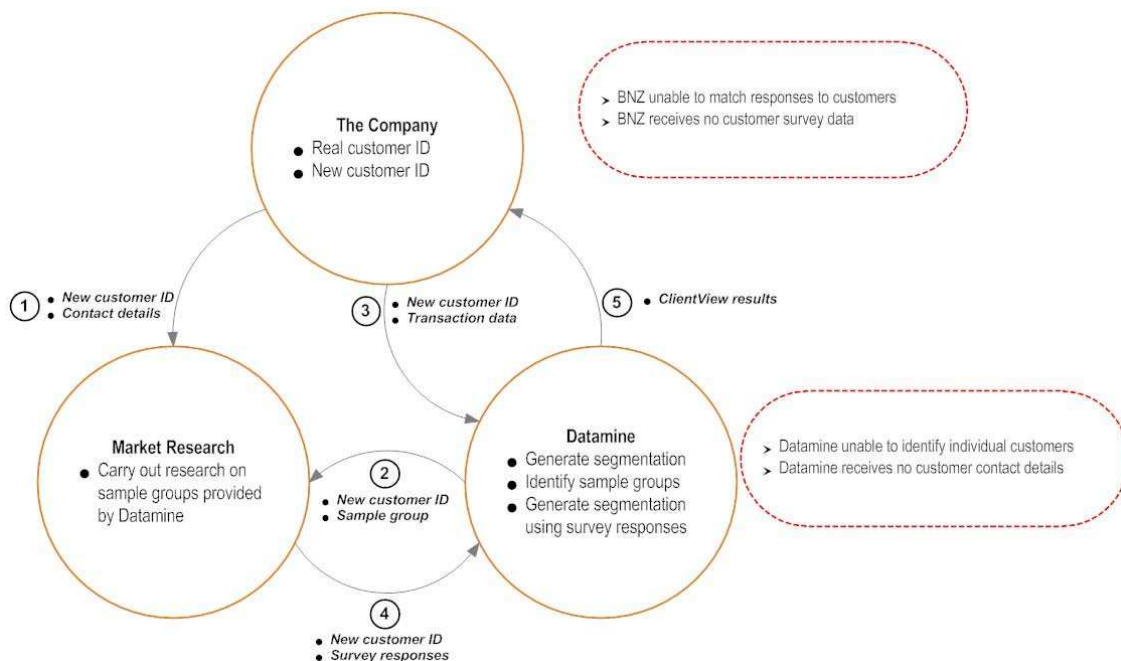
Figure 3: The flow of data between the Company, Datamine and the Market Research company necessary to carry out the Segmentation analyses and retain the anonymity of individual customers.

The Company provided Datamine with new customer numbers, and information necessary to carry out the segmentation analysis. No name or address information was provided so Datamine could not identify individual customers. The segmentation analyses enabled groups of customers to be identified, and these customer numbers were provided to the market research company.

The Company also provided the market research company a list of new customer numbers, and their associated name and address details. The market research company, after carrying out further research on the customers identified by Datamine, sent back the individual responses for each of the identified customers to Datamine.

Datamine then used the specific responses, such as share of wallet, and using the new customer number, matched these responses to other behaviour from the Company's database. This enabled further segmentation analyses to be produced which included competitor information.

Using this method, the customers' privacy was protected at all times. By the Company creating new customer numbers, and providing these to Datamine and the market research company, meant that customer transaction details could not be linked back to an individual. In turn, the Company were only provided the segmentation results, no individual market research responses.

The greatest benefit to the Company was that this method enabled them to identify penetration by competitors within various groups of their customer base.

IN SUMMARY

Both market research and data driven marketing have their own important part to play as essential tools for marketers. However, the greatest power can be extracted from these tools if they are used in combination, either one supporting the other, or used in tandem.

In order to achieve this, individual customer responses need to be linked back to their behavioural data using data fusion. This needs to be undertaken while ensuring that the anonymity of individuals is protected and following robust methodologies. Organizational barriers to integration need to be broken down so that teams from different disciplines work together to create optimal solutions for marketing.

It is worthwhile and beneficial for organisations to bring these important marketing disciplines together, providing an integrated approach to customer insight.

BRIEF AUTHOR BIOGRAPHIES

Paul O'Connor

Paul founded Datamine in 1995 and has been the catalyst in Datamine's fast growth into a credible business analytics company. Pauls qualifications include a BSc and a diploma in Operations Research and Statistics from Victoria University of Wellington NZ.

Sally Carey

Sally is Director of Datamine Ltd, a New Zealand based analytics consultancy that moves its clients beyond guesswork. Sally has over 25 years of B to B and B to C marketing and using quantitative approaches for business decision making. Sally has an MBA from Bradford University (UK) and is a Fellow of the Institute of Direct & Digital Marketing (UK).

Sally believes that extraordinary results are achieved by a combination of analysis and intuition, and have been referred to by some clients as magic.

Keywords: *churn, retention, profiling, modelling, predictive model, performance, data mining, direct marketing, customer insight*

[i] Wikipedia, http://en.wikipedia.org/wiki/Data_fusion, accessed August 2011
[ii] Net Promoter, http://www.netpromoter.com/np/calculate.jsp, accessed September 2011